



**NATIONAL  
QUALITY FORUM**

Driving measurable health  
improvements together

# Surgery Fall 2020 Measure Review Cycle

## Standing Committee Orientation

Amy Moyer, NQF Director  
Janaki Panchal, NQF Manager  
Karri Albanese, NQF Analyst  
Mike DiVecchia, NQF Project Manager

*January 12, 2021*

*Funded by the Centers for Medicare and Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001.*



## Housekeeping Reminders

- The CenturyLink web platform will allow you to visually follow the presentation
- Please mute your computer and dial into the call to participate
  - Dial **800-768-2983** and enter passcode **7445915**
- Feel free to use the chat feature to communicate with NQF Staff or the group
- To reduce feedback, please mute your line when you are not speaking
- We will do a Committee roll call once the meeting begins

If you are experiencing technical issues, please contact the NQF project team at [surgery@qualityforum.org](mailto:surgery@qualityforum.org)

# Welcome

## Project Team



**Amy Moyer, MS,  
PMP**  
**NQF Director**



**Janaki Panchal,  
MSPH**  
**NQF Manager**



**Karri Albanese,  
NQF Analyst**



**Mike DiVecchia,  
MBA, PMP**  
**Project Manager**



## Agenda for the Call

- Standing Committee Attendance and Introductions
- Overview of NQF, the Consensus Development Process (CDP)
- Overview of Roles of the Standing Committee, Co-chairs, Scientific Methods Panel, and NQF Staff
- Overview of the Measure Evaluation Process
- Overview of NQF's Portfolio of Surgery Measures
- Overview of NQF's Measure Evaluation Criteria
- Overview of Social Risk
- SharePoint Tutorial
- Next Steps

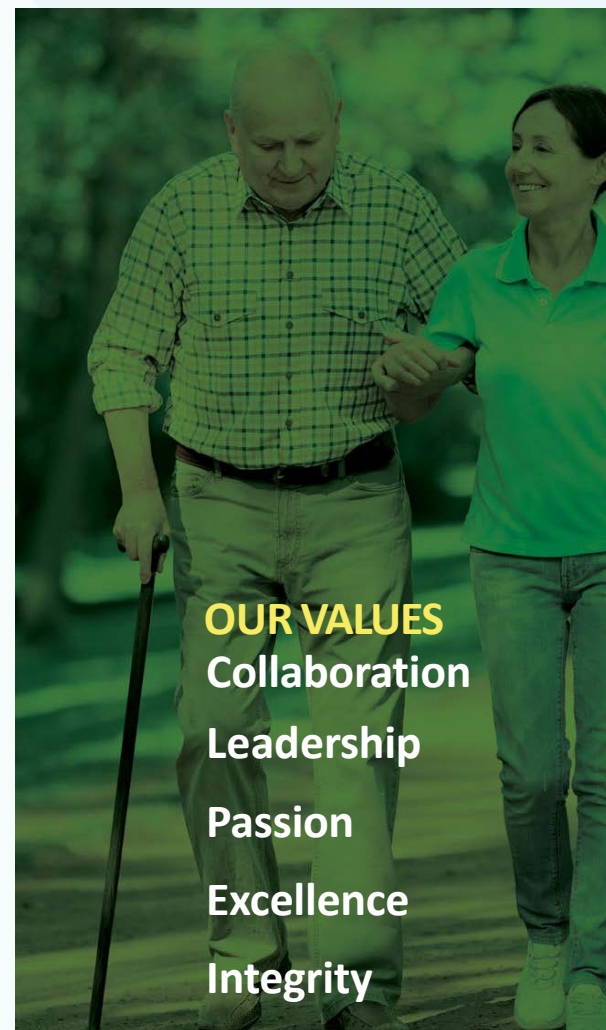
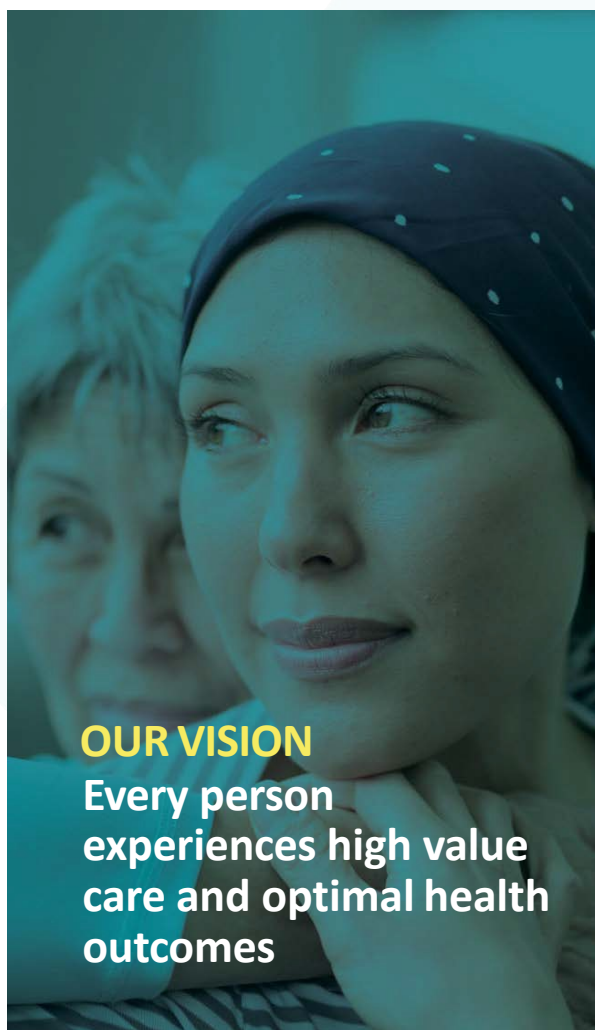
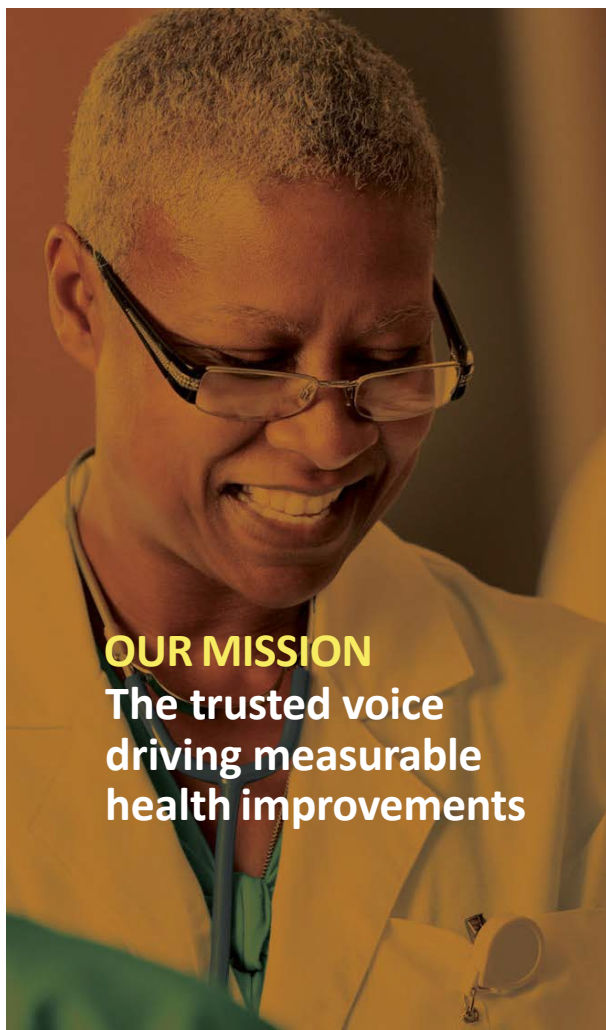
## Surgery Standing Committee

- William Gunnar, MD, JD (Co-Chair)
- Alex Sox-Harris, PhD, MS (Co-Chair)
- Ashrith Amarnath, MD
- Sherry Bernardo, CRNA\*
- Kenya Brown, LCSW-C
- Richard D'Agostino, MD\*
- TeMaya Eatmon
- Elisabeth Erekson, MD, MPH, FACOG, FACS
- Michael Firstenberg, MD, FACC, FAIM\*
- Linda Groah, MSN, RN, CNOR, NEA-BC\*
- Vilma Joseph, MD, MPH, FASA
- Miklos Kertai, MD, PhD\*
- Barbara Levy, MD, FACOG, FACS
- Jaime Ortiz, MD, MBA, FASA
- Shawn Rangel, MD, MSCE
- Kimberly Richardson\*
- Christopher Saigal, MD, MPH
- Rajdeep Sandhu, MD, MMM, FACS, FSVS\*
- Salvatore T. Scali, MD, FACS, DFSVS, RPVI
- Allan Siperstein, MD
- Kevin Wang, MHA
- Mark A. Wilson, MD, PhD\*

# Overview of NQF and the Consensus Development Process (CDP)



# The National Quality Forum – A Unique Role





## NQF Activities in Multiple Measurement Areas

- **Performance Measure Endorsement**

- 400+ NQF-endorsed measures across multiple clinical areas
- 15 empaneled standing expert committees including the Scientific Methods Panel

- **Measure Applications Partnership (MAP)**

- Provides recommendations to HHS on selecting measures for 19 federal programs

- **Advancing Measurement Science**

- Convenes private and public sector leaders to reach consensus on complex issues in healthcare performance measurement
- Examples include CMS-funded projects such as HCBS, rural issues, telehealth, interoperability, attribution, risk-adjustment for social risk factors, diagnostic accuracy and disparities

- **Other Measurement Work**

- Creation of action-oriented playbooks and implementation guides that include measurement frameworks and/or opportunities for organizations to measure progress on high-priority healthcare topics
- Conducts Strategy Sessions with stakeholders to identify measure gaps and opportunities



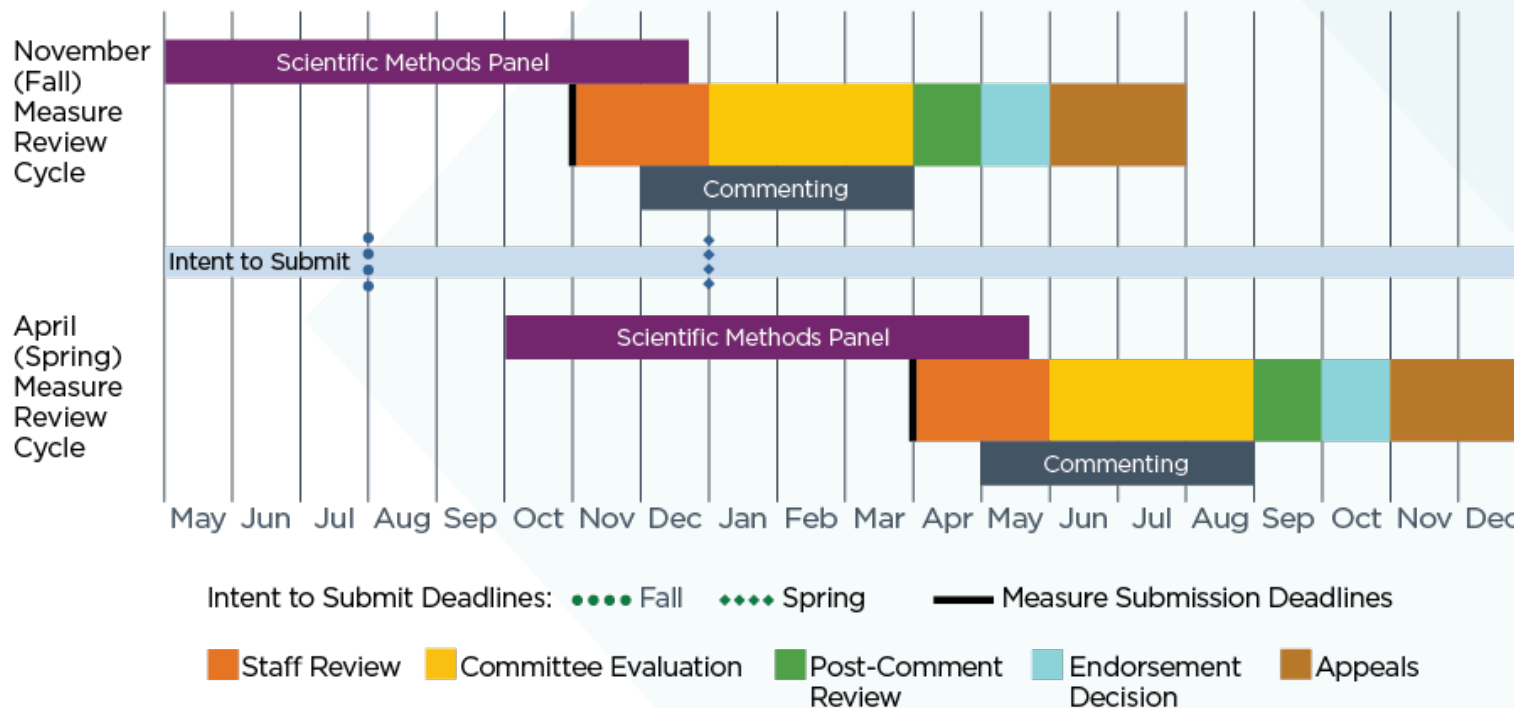
# NQF Consensus Development Process (CDP)

## 6 Steps for Measure Endorsement

- Intent to Submit
  - Scientific Methods Panel (SMP) if applicable
    - » Review of complex measures for scientific acceptability
- Call for Nominations
- Measure Evaluation
- Public Commenting Period with Member Support
- Measure Endorsement
  - Consensus Standards Approval Committee (CSAC)
- Measure Appeals

## Measure Review: Two Cycles Per Year

Consensus Development Process:  
Two Cycles Every Contract Year





## 14 Measure Review Topical Areas

- All Cause Admission/Readmissions
- Behavioral Health and Substance Use
- Cancer
- Cardiovascular
- Cost and Efficiency
- Geriatric and Palliative Care
- Neurology
- Patient Experience and Function
- Patient Safety
- Perinatal and Women's Health
- Prevention and Population Health
- Primary Care and Chronic Illness
- Renal
- Surgery

# **Overview of Roles of the Standing Committee, Co-chairs, Scientific Methods Panel, and NQF Staff**



## Role of the Standing Committee

### *General Duties*

- Act as a proxy for the NQF multi-stakeholder membership
- Serve initial 2-year or 3-year terms
  - Opportunity to renew for 2 additional years (4 cycles)
- Work with NQF staff to evaluate and endorse measures
- Evaluate candidate measures against the measure evaluation criteria
- Respond to comments submitted during the public commenting period
- Respond to any directions from the CSAC
- Refer to the [Standing Committee Guidebook](#) for more information



## Role of the Standing Committee

### *Meeting Participation*

- Meeting attendance
  - Must notify NQF staff in advance of meeting if unable to attend
- Quorum requirements
  - NQF Quorum=66% of active members
  - Committee recommendations can only be made with a quorum of Committee votes
    - » Not based on Robert's Rules of Order
  - Votes may be requested via email if quorum is not reached during the meeting
    - » Materials (i.e., transcripts upon request) will be sent to inform votes
  - Meetings may be cancelled (and rescheduled) if quorum not reached and vote is required
- Measure-specific disclosure of interest
  - Must be completed to participate in the measure evaluation discussion (each cycle)



## **Role of the Standing Committee**

### ***Measure Evaluation Duties***

- All members evaluate measures being considered for endorsement
- Evaluate measures against each criterion
  - Indicate the extent to which each criterion is met and rationale for the rating
- Make recommendations to the NQF membership for endorsement
- Oversee Surgery portfolio of measures
  - Promote alignment and harmonization
  - Identify gaps





## Role of the Standing Committee Co-Chairs

- Co-facilitate Standing Committee (SC) discussion with NQF staff
- Assist NQF in anticipating questions and identifying additional information that may be useful to the SC
- Keep SC on track to meet goals of the project without hindering critical discussion/input
- Represent the SC at CSAC meetings
- Participate as a SC member

## Role of Scientific Methods Panel

- The Scientific Methods Panel (SMP) was created to ensure high-level consistent reviews of the scientific acceptability of measures
- The SMP is charged with:
  - Conducting evaluation of complex measures for the Scientific Acceptability criterion, with a methodological focus on reliability and validity analyses and results
  - Serve in broad advisory capacity to NQF on methodologic issues, including those related to measure testing, risk adjustment, and measurement approaches
- The SMP review will help inform the standing committee's endorsement decision; SMP will not render endorsement recommendations



## Role of NQF Staff

- NQF project staff works with SC to achieve the goals of the project and ensure adherence to the consensus development process:
  - Facilitate SC meetings, ensuring that goals are met
  - Organize and staff SC meetings and conference calls
  - Guide SC through the CDP and advise on NQF policy and procedures; ensure NQF evaluation criteria are appropriately applied and process is followed
  - Review measure submissions and prepare materials for Committee review
  - Draft and edit reports for SC review
  - Ensure and facilitate communication among all project participants (including SC and measure developers)
  - Assist measure developers in understanding NQF criteria and process
  - Facilitate collaboration between different NQF projects



## **Role of NQF Staff**

### ***Communication***

- Respond to NQF member or public queries about the project
- Maintain documentation of project activities
- Post project information to NQF's website
- Work with measure developers to provide necessary information and communication for the SC to fairly and adequately evaluate measures for endorsement
- Publish final project report



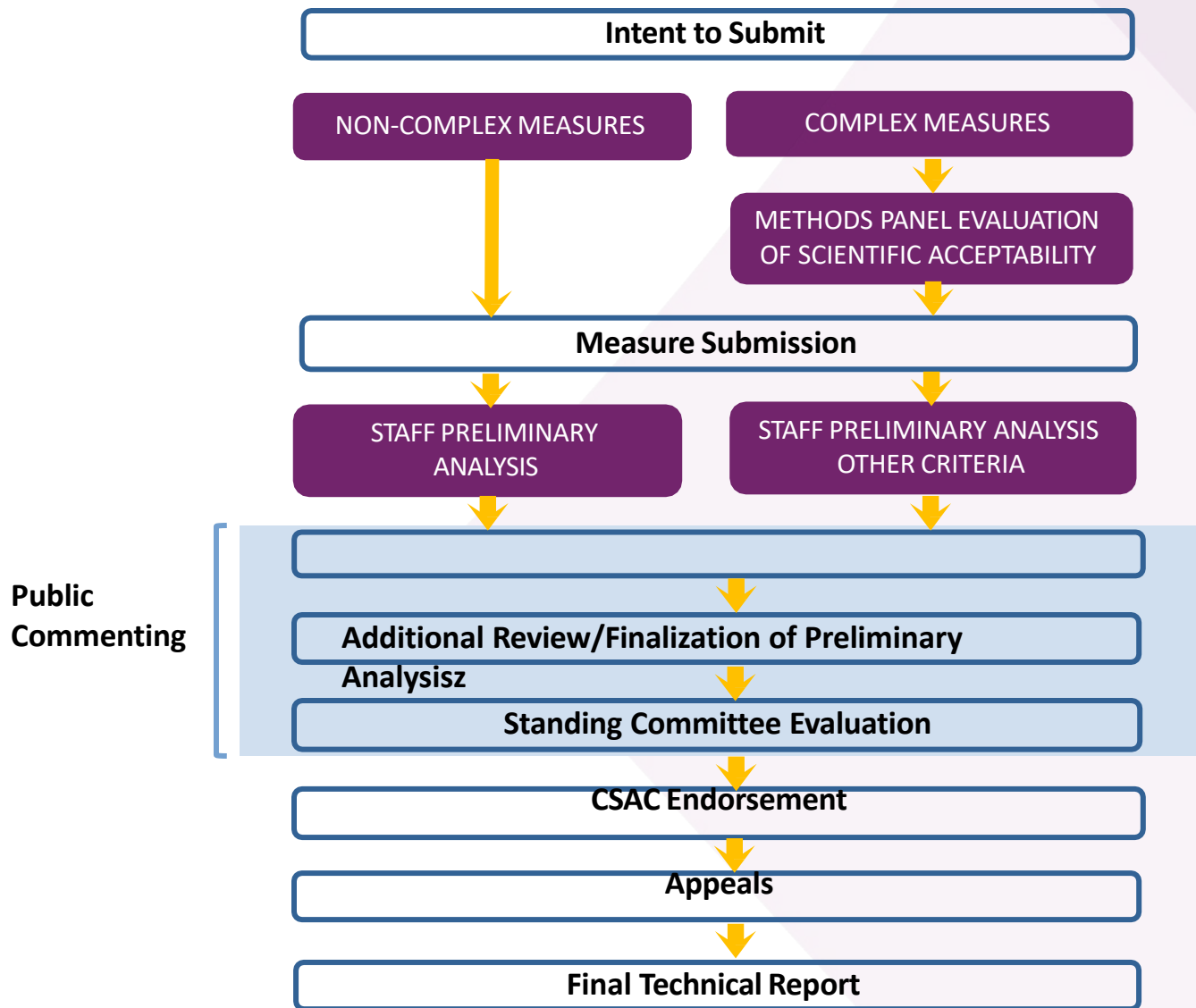
# Questions?<sup>1</sup>

# Overview of the Measure Evaluation Process



# Measure Evaluation Overview

# Measure Evaluation Workflow







## NQF Consensus Development Process (CDP) Measure Evaluation

### *Complex Measures*

- Outcome measures, including intermediate clinical outcomes
- Instrument-based measures (e.g., PRO-PMs)
- Cost/resource use measures
- Efficiency measures (those combining concepts of resource use and quality)
- Composite measures

- Process measures
- Structural measures
- Previously endorsed complex measures with no changes/updates to the specifications or testing



NATIONAL  
QUALITY FORUM

# **Complex Measures**

## **Scientific Methods Panel**

## Complex Measure Evaluation by the Scientific Methods Panel (SMP)

- Complex measures include composite, instrument-based (including PRO-PM), cost/resource, efficiency, and outcome (including intermediate clinical outcome) measures
- Complex measures are reviewed by the SMP when:
  - Newly submitted
  - Maintenance measures with updated testing
  - NQF staff requests (e.g., expert opinion needed to support review of testing, review of unfamiliar methodology)
- The SMP will provide evaluations and ratings of **reliability** and **validity** to the standing committees
  - Measures that did not get a "pass" for either reliability and validity during preliminary analyses are discussed at the SMP evaluation meetings, and are re-voted



## Post-SMP Evaluation

- All eligible measures reviewed by the SMP can be discussed by the Standing Committee
  - Standing Committee will evaluate and make recommendations for endorsement for:
    - » Measures that pass SMP review
    - » Measures where the SMP did not reach consensus
  - Measures that did not pass the SMP can be pulled by a standing committee member for further discussion
- Eligibility will be confirmed by NQF Staff and SMP co-chairs
- Measures that failed the SMP due to the following will not be eligible for re-vote:
  - » Inappropriate methodology or testing approach applied to demonstrate reliability or validity
  - » Incorrect calculations or formulas used for testing
  - » Description of testing approach, results, or data is insufficient for SMP to apply the criteria
  - » Appropriate levels of testing not provided or otherwise did not meet NQF's minimum evaluation requirements



NATIONAL  
QUALITY FORUM

# Measure Evaluation Standing Committee



## Standing Committee Measure Evaluation Process

- Standing Committee members are notified of the SMP evaluation results (if complex measures reviewed by SMP)
- Standing Committee members can pull failed measures for discussion (and re-vote for eligible measures)
- Any measure pulled by a Standing Committee member will be discussed
  - Request should be submitted with a brief rationale
- Some measures may be eligible for vote by the Standing Committee
  - Eligibility will be determined by NQF staff and SMP co-chairs



## NQF Process After Measure Submission

- NQF staff performs quality checks on measure submission
- Standing Committee members complete measure-specific disclosures of interest
- NQF staff creates a measure worksheet for each measure



## Committee Measure Evaluation Process

### **~3 week review period for Measure Worksheets:**

- Measure Information Form (MIF): describes measure and specifications (e.g., title, description, numerator, denominator)
- Preliminary analysis by NQF staff
- Committee preliminary ratings
- Member and public comments
- Information submitted by the developer
  - Evidence and testing attachments
  - Spreadsheets
  - Additional documents





## Committee Measure Evaluation Process<sup>1</sup>

- **Preliminary analysis (PA):** NQF staff will prepare a PA form and offer preliminary ratings for each criteria
  - The PA will be used as a starting point for the Committee evaluation
  - SMP will complete review of Scientific Acceptability criterion for complex measures
- **Individual evaluation:** Each Committee member will conduct an in-depth evaluation on all measures under review



## Committee Measure Evaluation Process<sup>2</sup>

- NQF staff compiles the Committee's comments and redistributes measure worksheet with summary of all members' preliminary evaluation
- **Lead discussants** are assigned to each measure for committee evaluation meetings
- **Measure evaluation and recommendations at the in-person/web meeting:** The entire Committee will discuss and rate each measure against the evaluation criteria and make recommendations for endorsement



## Committee Measure Evaluation Process<sup>3</sup>

- **Staff will prepare a draft report** detailing the Committee's discussion and recommendations
  - This report will be released for a 30-day public and member comment period
- **Post-comment call:** The Committee will re-convene for a post-comment call to discuss comments submitted
- **Final endorsement decision by the CSAC**
- **Opportunity for public to appeal endorsement decision**

# Overview of NQF's Surgery Portfolio

## Surgery Use Portfolio of Measures

- This project will evaluate measures related to Surgery conditions that can be used for accountability and public reporting for all populations and in all settings of care. This project will address topic areas including:
  - Perioperative safety
  - General surgery
  - Range of specialties including:
    - Cardiac
    - Cardiothoracic
    - Colorectal
    - Ocular
  - Orthopedic
  - Urogynecologic
  - Vascular surgery
- NQF currently has 57 endorsed measures within this topic area. Endorsed measures undergo periodic evaluation to maintain endorsement – “maintenance”.

## Surgery Portfolio of NQF-endorsed Measures

	Structure	Process	Outcome	Composite
Abdominal and Colorectal	-	-	1	-
Cardiac	3	5	16	6
General	3	-	2	-
Cross-Cutting (Inpatient)	-	-	2	-
Cross-Cutting (Outpatient)	-	-	2	-
Ocular	-	-	2	-
Orthopedic	-	-	4	-
Thoracic	1	-	1	1
Urogynecology/Gynecology	-	3	-	-
Vascular	-	2	6	-
Total	4	10	36	7



## Surgery Portfolio of NQF-endorsed Measures<sup>1</sup>

- 0114 Risk-Adjusted Postoperative Renal Failure
- 0115 Risk-Adjusted Surgical Re-exploration
- 0117 Beta Blockade at Discharge
- 0118 Anti-Lipid Treatment Discharge
- 0119 Risk-Adjusted Operative Mortality for CABG
- 0120 Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR)
- 0121 Risk-Adjusted Operative Mortality for Mitral Valve (MV) Replacement
- 0122 Risk-Adjusted Operative Mortality for Mitral Valve (MV) Replacement + CABG Surgery
- 0123 Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) + CABG Surgery



## Surgery Portfolio of NQF-endorsed Measures<sup>2</sup>

- 0127 Preoperative Beta Blockade
- 0129 Risk-Adjusted Postoperative Prolonged Intubation (Ventilation)
- 0130 Risk-Adjusted Deep Sternal Wound Infection
- 0131 Risk-Adjusted Stroke/Cerebrovascular Accident
- 0134 Use of Internal Mammary Artery (IMA) in Coronary Artery Bypass Graft (CABG)
- 0236 Coronary Artery Bypass Graft (CABG): Preoperative Beta-Blocker in Patients with Isolated CABG Surgery
- 0340 RACHS-1 Pediatric Heart Surgery Volume (PDI 7)
- 0354 Hip Fracture Mortality Rate (IQI 19)
- 0357 Abdominal Aortic Aneurysm (AAA) Repair Volume (IQI 4)





## Surgery Portfolio of NQF-endorsed Measures<sup>3</sup>

- 0359 Abdominal Aortic Aneurysm (AAA) Repair Mortality Rate (IQI 11)
- 0365 Pancreatic Resection Mortality Rate (IQI 9)
- 0366 Pancreatic Resection Volume (IQI 2)
- 0456 Participation in a Systematic National Database for General Thoracic Surgery
- 0465 Perioperative Anti-platelet Therapy for Patients undergoing Carotid Endarterectomy
- 0533 Postoperative Respiratory Failure Rate (PSI 11)
- 0564/0564e Cataracts: Complications within 30 Days Following Cataract Surgery Requiring Additional Surgical Procedures
- 0565/0565e Cataracts: 20/40 or Better Visual Acuity within 90 Days Following Cataract Surgery



## Surgery Portfolio of NQF-endorsed Measures<sup>4</sup>

- 0696 STS CABG Composite Score
- 0697 Risk Adjusted Case Mix Adjusted Elderly Surgery Outcomes Measure
- 0706 Risk Adjusted Colon Surgery Outcome Measure
- 0732 Surgical Volume for Pediatric and Congenital Heart Surgery: Total Programmatic Volume and Programmatic Volume Stratified by the 5 STAT Mortality Categories
- 0733 Operative Mortality Stratified by the 5 STAT Mortality Categories
- 0734 Participation in a National Database for Pediatric and Congenital Heart Surgery
- 1501 Risk-Adjusted Operative Mortality for Mitral Valve (MV) Repair
- 1502 Risk-Adjusted Operative Mortality for Mitral Valve (MV) Repair + CABG Surgery



## Surgery Portfolio of NQF-endorsed Measures<sup>5</sup>

- 1519 Statin Therapy at Discharge after Lower Extremity Bypass (LEB)
- 1523 Rate of Open Repair of Abdominal Aortic Aneurysms (AAA) Where Patients Are Discharged Alive
- 1534 In-hospital Mortality Following Elective EVAR of AAAs
- 1540 Postoperative Stroke or Death in Asymptomatic Patients undergoing Carotid Endarterectomy
- 1543 Postoperative Stroke or Death in Asymptomatic Patients undergoing Carotid Artery Stenting (CAS)
- 1550 Hospital-level Risk-standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA)
- 1551 Hospital-level Risk-standardized Readmission Rate (RSRR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA)



## Surgery Portfolio of NQF-endorsed Measures<sup>6</sup>

- 1790 Risk-Adjusted Morbidity and Mortality for Lung Resection for Lung Cancer
- 2038 Performing vaginal apical suspension at the time of hysterectomy to address pelvic organ prolapse
- 2063 Performing cystoscopy at the time of hysterectomy for pelvic organ prolapse to detect lower urinary tract injury
- 2558 Hospital 30-Day, All-Cause, Risk-Standardized Mortality Rate (RSMR) Following Coronary Artery Bypass Graft (CABG) Surgery
- 2561 STS Aortic Valve Replacement (AVR) Composite Score
- 2563 STS Aortic Valve Replacement (AVR) + Coronary Artery Bypass Graft (CABG) Composite Score
- 2677 Preoperative evaluation for stress urinary incontinence prior to hysterectomy for pelvic organ prolapse



## Surgery Portfolio of NQF-endorsed Measures<sup>7</sup>

- 2683 Risk-Adjusted Operative Mortality for Pediatric and Congenital Heart Surgery
- 2687 Hospital Visits after Hospital Outpatient Surgery
- 3030 STS Individual Surgeon Composite Measure for Adult Surgery
- 3031 STS Mitral Valve Repair/Replacement (MVRR) Composite Score
- 3032 STS Mitral Valve Repair/Replacement (MVRR) + Coronary Artery Bypass Graft (CABG) Composite Score
- 3294 STS Lobectomy for Lung Cancer Composite Score
- 3357 Facility-Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers



## Surgery Portfolio of NQF-endorsed Measures<sup>8</sup>

- 3493 Risk-standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA) for Merit-based Incentive Payment System (MIPS) Eligible Clinicians and Eligible Clinician Groups
- 3494 Hospital 90-Day, All-Cause, Risk-Standardized Mortality Rate (RSMR) Following Coronary Artery Bypass Graft (CABG) Surgery



## Fall 2020 Measures For Review

### ■ Maintenance measures

- 0117 Beta Blockade at Discharge
- 0127 Preoperative Beta Blockade
- 0134 Use of Internal Mammary Artery (IMA) in Coronary Artery Bypass Graft (CABG)
- 1550 Hospital-level risk-standardized complication rate (RSCR) following elective primary total hip arthroplasty (THA) and/or total knee arthroplasty (TKA)
- 1551 Hospital-level 30-day risk-standardized readmission rate (RSRR) following elective primary total hip arthroplasty (THA) and/or total knee arthroplasty (TKA)
- 3030 STS Individual Surgeon Composite Measure for Adult Cardiac Surgery
- 3031 STS Mitral Valve Repair/Replacement (MVRR) Composite Score
- 3032 STS Mitral Valve Repair/Replacement (MVRR) + Coronary Artery Bypass Graft (CABG) Composite Score



## Fall 2020 Measures Reviewed by the SMP

### **Passed Reliability and Validity**

- 1550 Hospital-level risk-standardized complication rate (RSCR) following elective primary total hip arthroplasty (THA) and/or total knee arthroplasty (TKA)
- 1551 Hospital-level 30-day risk-standardized readmission rate (RSRR) following elective primary total hip arthroplasty (THA) and/or total knee arthroplasty (TKA)





## Activities and Timeline

\*All timesET

Meeting	Date/Time
Orientation Call	January 12, 2021, 3:30-5:30 pm
Measure Evaluation Web Meeting 1	February 12, 2021, 9:00 am-5:00 pm
Measure Evaluation Web Meeting 2	February 16, 2021, 12:00-2:00 pm
Post-Comment Call	June 1, 2021 at 11:30 am-1:30 pm



NATIONAL  
QUALITY FORUM

# Questions?

# Measure Evaluation Criteria Overview



## NQF Measure Evaluation Criteria for Endorsement

**NQF endorses measures for accountability applications (public reporting, payment programs, accreditation, etc.) as well as quality improvement**

- Standardized evaluation criteria
- Criteria have evolved over time in response to stakeholder feedback
- The quality measurement enterprise is constantly growing and evolving—greater experience, lessons learned, expanding demands for measures—the criteria evolve to reflect the ongoing needs of stakeholders



## Major Endorsement Criteria (page 32 in the SC Guidebook)

- **Importance to measure and report:** Goal is to measure those aspects with greatest potential of driving improvements; if not important, the other criteria are less meaningful (**must-pass**)
- **Reliability and Validity-scientific acceptability of measure properties:** Goal is to make valid conclusions about quality; if not reliable and valid, there is risk of improper interpretation (**must-pass**)
- **Feasibility:** Goal is to, ideally, cause as little burden as possible; if not feasible, consider alternative approaches
- **Usability and Use (must-pass for maintenance measures):** Goal is to use for decisions related to accountability and improvement; if not useful, probably do not care if feasible
- **Comparison to related or competing measures**



## **Criterion #1: Importance to Measure and Report (page 34-42)**

1. Importance to measure and report - Extent to which the specific measure focus is evidence-based and important to making significant gains in healthcare quality where there is variation in or overall less-than-optimal performance.

1a. Evidence: the measure focus is evidence-based

1b. Opportunity for Improvement: 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

1c. Quality construct and rationale (composite measures only)



## Subcriterion #1a: Evidence (page 36-42)

### ■ Outcome measures

- Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.

### ■ Structure, process, intermediate outcome measures

- The quantity, quality, and consistency of the body of evidence underlying the measure should demonstrate that the measure focuses on those aspects of care known to influence desired patient outcomes
  - » Empirical studies (expert opinion is not evidence)
  - » Systematic review and grading of evidence
    - *Clinical Practice Guidelines – variable in approach to evidence review*

### ■ For measures derived from patient (or family/parent/etc.) report

- Evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Current requirements for structure and process measures also apply to patient-reported structure/process measures.



## Rating Evidence: Algorithm #1 (page 37)

- [Screen share Evidence algorithm]



# Criterion #1: Importance to measure and report

Criteria emphasis is different for new vs. maintenance measures

New measures	Maintenance measures
<ul style="list-style-type: none"> <li>Evidence – Quantity, quality, consistency (QQC)</li> <li>Established link for process measures with outcomes</li> </ul>	<p><b>DECREASED EMPHASIS:</b> Require measure developer to attest evidence is unchanged evidence from last evaluation; Standing Committee to affirm no change in evidence</p> <p>IF changes in evidence, the Committee will evaluate as for new measures</p>
<ul style="list-style-type: none"> <li>Gap – opportunity for improvement, variation, quality of care across providers</li> </ul>	<p><b>INCREASED EMPHASIS:</b> data on current performance, gap in care and variation</p>



## **Criterion #2: Reliability and Validity – Scientific Acceptability of Measure Properties (pages 42-54)**

**Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of health care delivery**

### **2a. Reliability (must-pass)**

2a1. Precise specifications including exclusions

2a2. Reliability testing—data elements or measure score

### **2b. Validity (must-pass)**

2b1. Validity testing—data elements or measure score

2b2. Justification of exclusions—relates to evidence

2b3. Risk adjustment—typically for outcome/cost/resource use

2b4. Identification of differences in performance

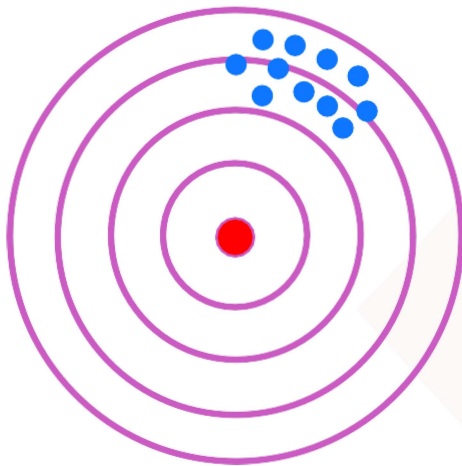
2b5. Comparability of data sources/methods

2b6. Missing data



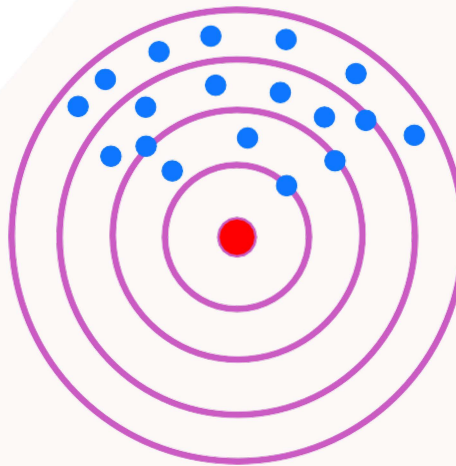
## Reliability and Validity (*page 44*)

**Assume the center of the target is the true score**



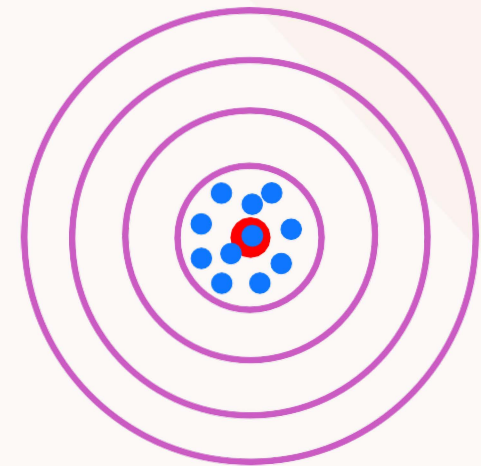
**Reliable  
Not Valid**

Consistent,  
but wrong



**Neither Reliable  
Nor Valid**

Inconsistent &  
wrong



**Both Reliable  
And Valid**

Consistent &  
correct



## Evaluating Scientific Acceptability – Key Points *(page 45)*

Empirical analysis to demonstrate the reliability and validity of the measure as specified, including:

- Analysis of issues that pose threats to the validity of conclusions about quality of care such as exclusions
- Risk adjustment/stratification for outcome and resource use measures
- Methods to identify differences in performance
- Comparability of data sources/methods



## Reliability Testing – Key Points (page 48)

- Reliability of the **measure score** refers to the proportion of variation in the performance scores due to systematic differences across the measured entities in relation to random variation or noise (i.e., the precision of the measure).
  - Example – Statistical analysis of sources of variation in performance measure scores (signal-to-noise analysis)
- Reliability of the **data elements** refers to the repeatability/reproducibility of the data and uses patient-level data
  - Example – inter-rater reliability
- Consider whether testing used an appropriate method and included adequate representation of providers and patients and whether results are within acceptable norms
- Algorithm #2



## Rating Reliability: Algorithm #2 (page 47)

- [Screen share Reliability algorithm]



## Validity Testing (pages 48-54)

- Empirical testing
  - Measure score – assesses a hypothesized relationship of the measure results to some other concept; assesses the correctness of conclusions about quality
  - Data element – assesses the correctness of the data elements compared to a “gold standard”
- Face validity
  - Subjective determination by experts that the measure appears to reflect quality of care
    - » Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.
    - » Requires systematic and transparent process, by identified experts, that explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.



## Rating Validity: Algorithm #3 (page 53)

- [Screen share Validity algorithm]





## Threats to Validity

- Conceptual
  - Measure focus is not a relevant outcome of healthcare or not strongly linked to a relevant outcome
- Unreliability
  - Generally, an unreliable measure cannot be valid
- Patients inappropriately excluded from measurement
- Differences in patient mix for outcome and resource use measures
- Measure scores that are generated with multiple data sources/methods
- Systematic missing or “incorrect” data (unintentional or intentional)

## Criterion #2: Scientific Acceptability

New measures	Maintenance measures
<ul style="list-style-type: none"> <li>Measure specifications are precise with all information needed to implement the measure</li> </ul>	<p>NO DIFFERENCE: Require updated specifications</p>
<ul style="list-style-type: none"> <li>Reliability</li> <li>Validity (including risk-adjustment)</li> </ul>	<p><b>DECREASED EMPHASIS:</b> If prior testing adequate, no need for additional testing at maintenance with certain exceptions (e.g., change in data source, level of analysis, or setting)</p> <p>Must address the questions regarding use of social risk factors in risk-adjustment approach</p>



## **Criterion #3: Feasibility** ***(pages 54-55)***

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement.

3a: Clinical data generated during care process

3b: Electronic sources

3c: Data collection strategy can be implemented



## Criterion #4: Usability and Use (pages 55-56)

Extent to which potential audiences (e.g., consumers, purchasers, providers, policymakers) 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.

### Use (4a) Must-pass for maintenance measures

**4a1: 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.

**4a2: Feedback by those being measured or others:** Those being measured have been given results and assistance in interpreting results; those being measured and others have been given opportunity for feedback; the feedback has been considered by developers.

### Usability (4b)

**4b1: Improvement:** Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

**4b2: Benefits outweigh the harms:** 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).

## Criteria #3-4: Feasibility and Usability and Use

### Feasibility

New measures	Maintenance measures
<ul style="list-style-type: none"> <li>Measure feasible, including eMeasure feasibility assessment</li> </ul>	NO DIFFERENCE: Implementation issues may be more prominent

### Usability and Use

New measures	Maintenance measures
<ul style="list-style-type: none"> <li>Use: used in accountability applications and public reporting</li> <li>Usability: impact and unintended consequences</li> </ul>	<b>INCREASED EMPHASIS:</b> Much greater focus on measure use and usefulness, including both impact and unintended consequences



## Criterion #5: Related or Competing Measures (pages 56-57)

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

- 5a. The measure specifications are harmonized with related measures **OR** the differences in specifications are justified.
- 5b. The measure is superior to competing measures (e.g., is a more valid or efficient way to measure) **OR** multiple measures are justified.



## Updated guidance for measures that use ICD-10 coding

- For CY2019 and beyond, reliability testing should be based on ICD-10 coded data.
- Validity testing should be based on ICD-10 coded data
- If providing face validity (FV), both FV of the ICD-10 coding scheme and FV of the measure score as an indicator of quality is required update

## eCQMs (Electronic Clinical Quality Measures)

- eCQMs must be tested empirically using the HQMF specifications. The minimum requirement is testing in EHR systems from more than one EHR vendor.
- Beginning Summer 2019, data element validation is required for all eCQMs (demonstration of score-level validation is also encouraged).
- For eCQMs based solely on structured data fields, reliability testing is not required if data element validation is demonstrated.
  - If data element testing is not possible, justification is required and must be accepted by the Standing Committee.
- A feasibility assessment (scorecard) is required to address the data elements and includes an assessment of the measure logic.



## eCQMs

- NQF staff technical review
  - Each submitted eCQM undergoes a technical review by NQF staff before going to the Standing Committee for evaluation.
  - For this technical review, NQF staff:
    - » Confirms that the measure uses the industry accepted eCQM technical specifications
    - » Determines if value sets have been vetted through the Value Set Authority Center (VSAC)
    - » Reviews the feasibility of each data element
    - » Confirms that the measure logic has been adequately unit tested using a simulated data set.
  - The technical review is included as part of the staff preliminary analyses within the measure worksheet.



# Questions?<sup>2</sup>

# Social Risk Overview



## Background

- NQF conducted a two-year trial period from 2015-2017. During this time, adjustment of measures for social risk factors was no longer prohibited
- The NQF Board of Directors reviewed the results of the trial period and determined there was a need to launch a new social risk initiative
- As part of the Equity Program, NQF will continue to explore the need to adjust for social risk
- Each measure must be assessed individually to determine if SDS adjustment is appropriate (included as part of **validity** subcriterion)
- The Standing Committee will continue to evaluate the measure as a whole, including the appropriateness of the risk adjustment approach used by the measure developer
- Efforts to implement SDS adjustment may be constrained by data limitations and data collection burden

*The Social Risk Trial is funded by the Centers for Medicare and Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001.*



## Standing Committee Evaluation

- The Standing Committee will be asked to consider the following questions:
  - Is there a conceptual relationship between the SDS factor and the measure focus?
  - What are the patient-level sociodemographic variables that were available and analyzed during measure development?
  - Does empirical analysis (as provided by the measure developer) show that the SDS factor has a significant and unique effect on the outcome in question?
  - Does the reliability and validity testing match the final measure specifications?



# Questions?<sup>3</sup>

# SharePoint Overview



# SharePoint Overview<sup>1</sup>

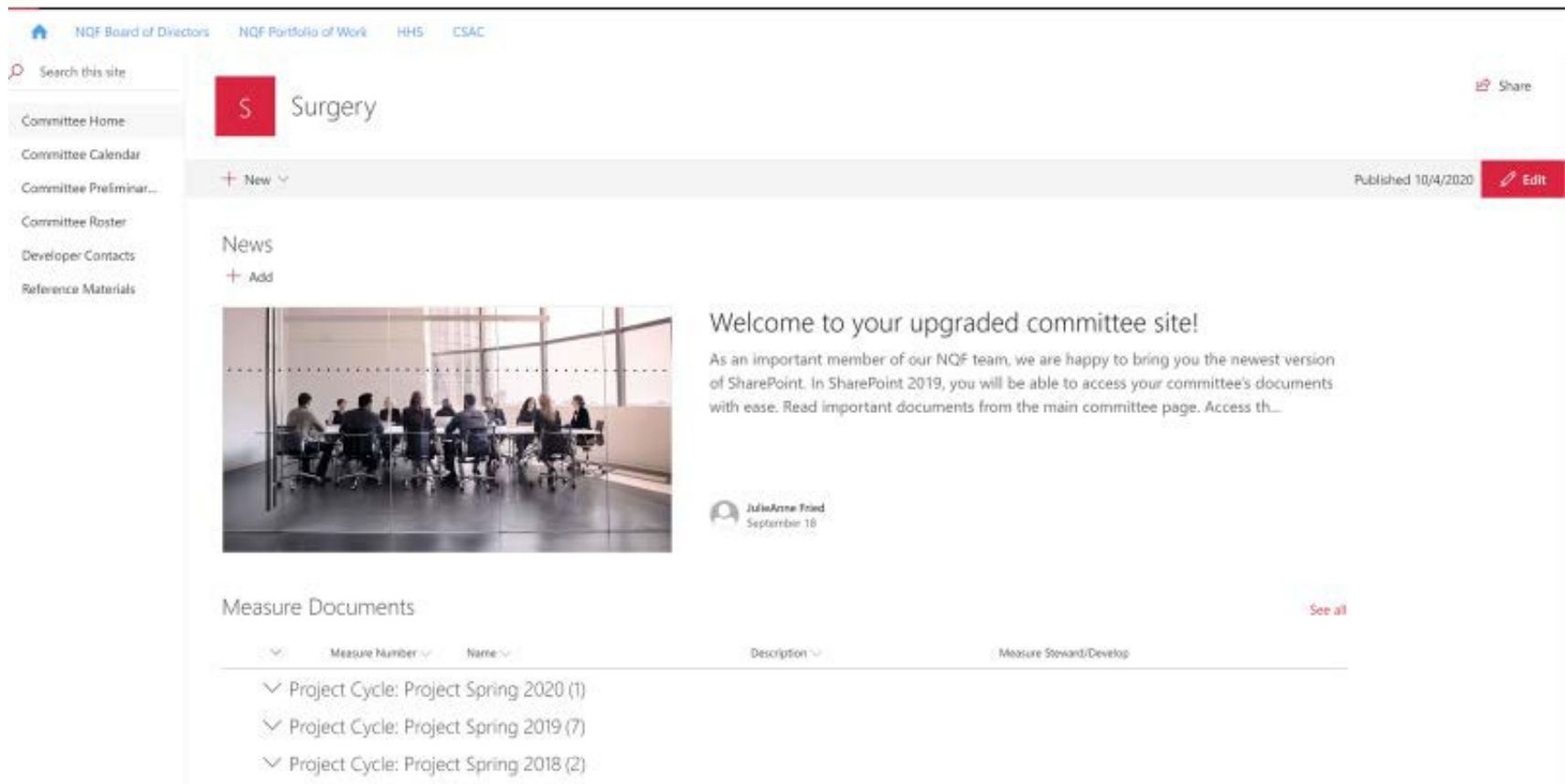
<https://share.qualityforum.org/portfolio/Surgery/SitePages/Home.aspx>

## Accessing SharePoint

- Standing Committee Policy
- Standing Committee Guidebook
- Measure Document Sets
- Meeting and Call Documents
- Committee Roster and Biographies
- Calendar of Meetings



## SharePoint Overview<sup>2</sup>



Home NQF Board of Directors NQF Portfolio of Work HHS CSAC

Search this site

Committee Home  
Committee Calendar  
Committee Preliminar...  
Committee Roster  
Developer Contacts  
Reference Materials

Surgery

+ New Published 10/4/2020 Edit

News  
+ Add

Welcome to your upgraded committee site!

As an important member of our NQF team, we are happy to bring you the newest version of SharePoint. In SharePoint 2019, you will be able to access your committee's documents with ease. Read important documents from the main committee page. Access th...

JulieAnne Fried  
September 18

Measure Documents See all

Measure Number	Name	Description	Measure Steward/Develop
✓	Project Cycle: Project Spring 2020 (1)		
✓	Project Cycle: Project Spring 2019 (7)		
✓	Project Cycle: Project Spring 2018 (2)		

# Next Steps



## Next Steps<sup>1</sup>

- Complete Measure-Specific DOIs
- Measure Worksheets will be shared with the Committee in January
- Preliminary Evaluation Survey due January 27, 2021
- Measure Evaluation Web Meetings
  - February 12, 9:00 am-5:00 pm ET
  - February 16, 12:00-2:00 pm ET



## Project Contact Info

- Email: [surgery@qualityforum.org](mailto:surgery@qualityforum.org)
- NQF phone: 202-783-1300
- Project page:  
[http://www.qualityforum.org/Surgery\\_2017-2018.aspx](http://www.qualityforum.org/Surgery_2017-2018.aspx)
- SharePoint site:  
<https://share.qualityforum.org/portfolio/Surgery/SitePages/Home.aspx>



# Questions?<sup>4</sup>

**THANK YOU.**

**NATIONAL QUALITY FORUM**

<http://www.qualityforum.org>