

Scientific Methods Panel Spring 2022 Evaluation Meeting

March 22-23, 2022



Housekeeping Reminders – Day 1

This is a Webex meeting with audio and video capabilities:

Meeting link: https://nqf.webex.com/nqf/j.php?MTID=m640cebadd881642081024533fabd2638

Meeting number: 2337 905 5153

Password: MScAEvent

Optional Dial-In: 1-844-621-3956 and enter passcode [2337 905 5153]

- Please place yourself on mute when you are not speaking
- We encourage you to use the following features
 - Chat box: to message NQF staff or the group
 - Raise hand: to be called upon to speak
- We will conduct Scientific Methods Panel roll call once the meeting begins

If you are experiencing technical issues, please contact the NQF project team at methodspanel@qualityforum.org



Housekeeping Reminders

- Meeting breaks
- Voting Quorum
- Chat feature
- Raising hand
- Muting and unmuting
- If possible, do not speak on speaker phone
- Introduce yourself; we are transcribing the discussion
- Technical support

Day 1: Welcome, Introductions, and Disclosures of Interest

Welcome



NQF Scientific Methods Panel (SMP) Team

- Elizabeth Drye, MD, SM, Chief Scientific Officer
- Tricia Elliott, DHA, MBA, CPHQ, FNAHQ, Senior Managing Director
- Matthew Pickering, PharmD, Senior Director
- Poonam Bal, MHSA, Senior Director
- Mike DiVecchia, MBA, PMP, Director
- Hannah Ingber, MPH, Manager
- Gabby Kyle-Lion, MPH, Analyst



Scientific Methods Panel Members

- David Nerenz, PhD, Co-chair
- Christie Teigland, PhD, Co-chair
- J. Matt Austin, PhD
- John Bott, MBA, MSSW
- Daniel Deutscher, PT, PhD
- Marybeth Farquhar, PhD, MSN, RN
- •Jeffrey Geppert, EdM, JD
- Laurent Glance, MD
- Joseph Hyder, MD
- Sherrie Kaplan, PhD, MPH
- Joseph Kunisch, PhD, RN-BC, CPHQ
- Paul Kurlansky, MD

- Zhenqiu Lin, PhD
- Jack Needleman, PhD
- Eugene Nuccio, PhD
- Sean O'Brien, PhD
- Jennifer Perloff, PhD
- ■Patrick Romano, MD, MPH
- Sam Simon, PhD
- •Alex Sox-Harris, PhD, MS
- Ronald Walters, MD, MBA, MHA, MS
- ■Terri Warholak, PhD, RPh, CPHQ, FAPhA
- Eric Weinhandl, PhD, MS
- Susan White, PhD, RHIA, CHDA

Meeting Overview



Meeting Agenda – Day 1

- Welcome, Introductions, and Disclosures of Interest
- Evaluation Updates (Fall 2021 and Spring 2022 cycles)
- Process Overview and Evaluation Reminders
- Spring 2022 Measure Evaluations
- Break 30 minutes
- Measure Evaluations Continued
- Opportunity for NQF Member and Public Comment
- Next Steps
- Adjourn



Meeting Ground Rules

- No rank in the room
- Remain engaged and actively participate
- Be prepared, having reviewed the measures beforehand
- Base evaluation and recommendations on the measure evaluation criteria and guidance
- Keep comments concise and focused
- Be respectful and allow others to contribute
- Share your experiences
- Learn from others



Meeting Materials

Discussion Guide

- » A synopsis document of scientific acceptability content (i.e., reliability and validity requirement) for all complex measures in a measure cycle evaluated by the SMP members.
 - Each measure includes pertinent information from the submission, SMP reviewer feedback, related developer responses, and identification of measures that are pulled for SMP discussion.
 - Goal is to summarize and highlight priority information for SMP discussion, reduce developer burden from multiple submission materials requests, and target critical scientific acceptability questions/concerns
- » Appendix B: Additional information provided by measure developers

Background Materials

- » 2011 Testing Task Force Report
- » 2021 NQF Measure Evaluation Criteria and Guidance
- » SMP Measure Evaluation Guidance

Fall 2021 Evaluation Updates



Fall 2021 SMP Measure Evaluation Cycle Statistics

- 12 measures were evaluated by the SMP
 - Seven measures were discussed at the meeting
- Final results
 - Eight of 12 measures passed SMP and were evaluated by the respective Standing Committees
 - » Two of the eight measures were consensus not reach (CNR) by the SMP and were voted on by the Standing Committees
 - One did not pass Standing Committee evaluation



Fall 2021 SMP Measures Revoted on by the Standing Committee

NQF ID	Measure Title	SMP Decision	Standing Committee and Decision	Current Status
3667	Days at Home for Patients with Complex, Chronic Conditions	Reliability: Pass Validity: CNR	Primary Care and Chronic Illness Reliability: Accepted SMP's vote Validity: No Pass	Not recommended for endorsement
0689	Percent of Residents Who Lose Too Much Weight (Long-Stay)	Reliability: Pass Validity: CNR	Patient Safety Reliability: Accepted SMP's vote Validity: Pass	Recommended for endorsement



Performance Metrics

Metrics	Fall 2017	Spring 2018	Fall 2018	Spring 2019	Fall 2019	Spring 2020	Fall 2020	Spring 2021	Fall 2021	Spring 2022*
Total number of complex measures submitted for evaluation by the SMP	8	21	39	47	20	21	24	29	12	13
Total Passed by SMP	5	7	25	30	15	16	20	23	8	5
Total Not Passed by SMP	3	13	10	11	4	3	2	2	2	1
Consensus Not Reached**	0	0	4	6	1	2	2	2	2	7
Percent agreement with Standing Committee ratings and SMP recommendations	4/4 (100%)	3/6 (50%)	24/25 (96%)	39/40 (97%)	14/15 (93%)	12/15 (80%)	18/19 (95%)	17/20 (85%)	6/6 (100%)	TBD

TBD: to be determined

^{*}Data for the Spring 2022 cycle are preliminary

^{**}These measures were sent to the Standing Committees

Spring 2022 Cycle Overview



Spring 2022 Evaluation Cycle Statistics

- 13 complex measures were assigned to the SMP
 - 10 were new measures
- 2 subgroups of 12 SMP members were each assigned 6 or 7 measures
 - » 5 measures passed reliability AND validity
 - » 7 measures were consensus not reached (CNR) on reliability OR validity
 - » 3 measures did not pass on reliability
 - » 1 measure did not pass validity and reliability
 - » 4 slated for re-vote
 - » 10 measures slated for discussion

- Reviewed Measures by Type
 - » 8 outcome
 - » 0 cost/resource use
 - » 1 composite
 - » 4 outcome: intermediate clinical outcome
 - » 0 PRO-PM
 - » 0 process
 - » 0 structure



Spring 2022 Measures Slated for Discussion

- Subgroup 1
 - Renal
 - » 1460**
 - Perinatal
 - » 0471e**
 - » 0716e**
 - » 3687e***
 - Patient Safety
 - » 2820***

- Subgroup 2
 - Renal
 - » 3679*
 - » 3689*
 - » 3694*
 - » 3695*
 - » 3697**

^{*}These measures will be re-voted on as consensus was not reached in the preliminary analysis done by the SMP.

^{**}These measures will be discussed because the developer submitted a response to the SMP evaluation. A re-vote could occur if the SMP decides it is warranted.

^{***}These measures were pulled for discussion by the SMP.

Process Overview and Evaluation Reminders



Overall Ratings

High (H)

- » Accountable entity level testing is required
- » A measure may be eligible for "HIGH," but the sampling method/results may warrant a "MODERATE" rating

Moderate (M)

- » The highest eligible rating if only patient/encounter level testing or face validity testing is conducted
- » A measure may be eligible for "MODERATE," but the sampling method/results may warrant a "LOW" rating

Low (L)

» Used primarily if testing results are not satisfactory or an inappropriate methodology was applied

Insufficient (I)

- » Use when the reviewer does not have sufficient information to assign a "HIGH," "MODERATE," or "LOW" rating
 - Examples: unclear specifications; unclear testing methodology, not conducting criteria required testing



Meeting Quorum and Achieving Consensus

- A meeting quorum is met with 66% of active SMP Members in attendance
- Achieving consensus is calculated from the percent of quorum members in attendance during a vote
- SMP scientific acceptability (i.e., reliability and validity criteria) evaluation results
 - » Pass/Recommended: Greater than 60% "Yes" of quorum votes (i.e., high + moderate ratings)
 - » Consensus not reached (CNR): 40-60% "Yes" of quorum votes (inclusive of 40% and 60%)
 - » No pass/Not recommended: Less than 40% "Yes" of quorum votes



Differences in Testing Requirements by Measure Type

- Health outcomes, intermediate clinical outcomes, cost/resource use, structure, process
 - » For both reliability and validity, NQF requires *EITHER* patient/encounter level testing *OR* accountability entity level testing for new measures
 - Both testing types are preferred, yet not currently required
 - Impacts rating, as described previously
 - Exception: face validity testing of the "computed measure score" for new measures is accepted at the accountable entity level
 - » If patient/encounter level validity testing is provided, we do not require additional reliability testing
 - In this case, use the rating you give for validity as the rating for reliability



Differences in Testing Requirements by Measure Type – *Instrument-based Measures (including PRO-PMs)*

- For reliability and validity, testing is required at both patient/encounter and accountable entity levels for initial endorsement evaluation
 - » Patient/encounter level testing *must be* conducted for reliability AND validity of the multi-item scales at the patient level
 - » Accountable entity level testing *must be* conducted for reliability AND validity testing of the actual performance measure at the level of analysis as defined in the measure specifications
 - Face validity testing of the "computed measure score" is accepted at initial endorsement evaluation in lieu of empirical accountable entity level validity testing



Differences in Testing Requirements by Measure Type – *Composite Measures*

- NQF provides specific guidance and definitions for "composite" measures
 - » Components of the composite measure should have their own properties of reliability and validity
 - » NQF does NOT consider multi-item scales in surveys/questionnaires as composites
 - » NQF does NOT consider multiple component measures without a single performance rate and multiple component performance rates as composites
- Accountability entity level reliability testing of the composite is required
- Demonstrating reliability of individual components alone is not sufficient to pass the criterion
- Accountability entity level validity testing is not required until maintenance
- Additional scientific acceptability subcriterion is required for composite measures
 - » Empirical analyses supporting the composite construction including the value of the components to the composite and the component aggregation and weighting consistency to composite quality construct



Testing and Evaluation Reminders

- All testing must align with specifications
 - » This is not a new requirement, yet NQF is more rigorously in upholding the requirement, particularly for level of analysis testing and minimum sample sizes
 - If multiple levels of analysis are specified, each must be tested separately
 - » NQF's requirements permit passing some or all levels of analysis for a measure
- Occasionally there are several performance measures included under one NQF number
 - » Each measure must be evaluated separately; some measures may pass and others may not pass



Additional Reminders

- Consideration for risk-adjustment is required for all outcome, resource use, and some process measures
 - » Inclusion (or exclusion) of certain factors in the risk-adjustment approach **should not** be a reason for not passing a measure
 - » Concerns with discrimination, calibration, or overall method of adjustment are grounds for not passing a measure
 - » In the absence of a risk adjustment for outcome, resource use, and some process measures, a strong rationale/data for excluding must be provided
- For all measures
 - » Incomplete or ambiguous specifications are grounds for not passing a measure
- Empirical validity testing is required at time of maintenance evaluation
 - » If not possible, a strong justification is required and must be accepted by the Standing Committee



Additional Reminders (continued)

- The SMP articulated additional **guidance** for submissions
 - 1. Provide greater detail when describing testing methodologies and results
 - 2. Provide more than one overall statistic when conducting signal-to-noise reliability testing
 - 3. Provide greater detail in description of construct validation describing:
 - » Hypothesized relationships
 - » Why examining hypothesized relationships would validate the measure
 - » Expected direction and strength of the association
 - » Specific statistical tests used, results, results interpretation, how the results related to hypothesis, and whether the results assist to validate the measure
 - Lack of #2 and #3 should not be grounds for not passing a measure



Standing Committee Complex Measure Evaluation

- All measures reviewed by the SMP can be discussed by the Standing Committees
 - » Standing Committees will evaluate and make recommendations for endorsement for:
 - Measures that pass SMP review
 - Measures where the SMP did not reach consensus (i.e., CNR)
 - » Measures that do not pass the SMP may be pulled by a Standing Committee member for further discussion and revote if it is an *eligible* measure



Committee Consideration of Measures that Do Not Pass the SMP

- Eligibility will be determined by NQF Staff and SMP co-chairs
 - » Measures that did not pass the SMP due to the following will not be eligible for revote by the Standing Committee:
 - Inappropriately applied methodology or testing approach 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



Measure Discussion Process

- Measures discussed by the SMP are determined during the SMP measure review activities
- Staff will briefly introduce the measure
- SMP member lead discussants will summarize key concerns
- Other SMP subgroup members are invited to comment
- Developers are given 2-3 minutes for an initial response, and may respond to SMP questions
- Discussions are opened to the full SMP and proceed by individual criterion
- Recused members cannot discuss measures where conflicts are identified



The Voting Process

- Voting is conducted synchronously, virtually, and confidentially via Poll Everywhere
- Voting occurs following each criterion discussion
- SMP subgroup members only vote on measures they were assigned
- Recused SMP members cannot vote for measures where conflicts are identified
- Subgroup voting results taken during the meeting are the official SMP vote
- Measures that are not pulled for discussion will pass in a consent calendar vote



Voting Test

Spring 2022 Measure Evaluation



#3689 First Year Standardized Waitlist Ratio (FYSWR)

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-0, M-10, L-0, I-0 Pass
 - Validity: H-1, M-5, L-4, I-0 Consensus Not Reached
- Lead Discussant: Joseph Hyder
- Measure Developer: University of Michigan Kidney Epidemiology and Cost Center
- Measure Steward: Centers for Medicare & Medicaid Services
- Discussion Guide page 18
- For SMP discussion:
 - Additional clarifying information from the developers.
 - Are there any concerns about the reliability and validity testing methodology, or the results?



#3694 Percentage of Prevalent Patients Waitlisted in Active Status (aPPPW)

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-5, M-3, L-0, I-2 Pass
 - Validity: H-2, M-4, L-3, I-1 Consensus Not Reached
- Lead Discussant: Zhenqiu Lin; Secondary Discussant: Eugene Nuccio
- Measure Developer: University of Michigan Kidney Epidemiology and Cost Center
- Measure Steward: Centers for Medicare & Medicaid Services
- Discussion Guide page 19
- For SMP discussion:
 - Additional clarifying information from the developers.
 - Are there any concerns about the reliability or validity testing methodology, or the results?
 - Are there concerns regarding the use of patient-months?



#3695 Percentage of Prevalent Patients Waitlisted (PPPW)

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-4, M-4, L-0, I-2 Pass
 - Validity: H-2, M-4, L-3, I-1 Consensus Not Reached
- Lead Discussant: Eugene Nuccio; Secondary Discussant: Zhenqiu Lin
- Measure Developer: University of Michigan Kidney Epidemiology and Cost Center
- Measure Steward: Centers for Medicare & Medicaid Services
- Discussion Guide page 21
- For SMP discussion:
 - Additional clarifying information from the developers.
 - Are there any concerns about the reliability or validity testing methodology, or the results?
 - Are there concerns regarding the use of patient-months?

Break

Will resume at 1:20 PM EST



#3679 Home Dialysis Rate

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-6, M-0, L-1, I-3 Consensus Not Reached
 - Validity: H-2, M-2, L-3, I-3 Consensus Not Reached
- Lead Discussant: Patrick Romano; Secondary Discussant: Daniel Deutscher
- Measure Developer: Kidney Care Quality Alliance
- Measure Steward: Kidney Care Quality Alliance
- Discussion Guide page 22
- For SMP discussion:
 - Additional clarifying items from the developer
 - Are the methods appropriate for testing reliability for this measure?
 - Is the measure calculation by individual facility or HRR as structured appropriate?
 - Is the high rate of identification of outliers in the reported measure scores a cause for concern?
 - Are the face validity results sufficient to meet the validity requirements?



#3697 Home Dialysis Retention

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-0, M-2, L-6, I-2 No Pass
 - Validity: H-0, M-5, L-3, I-2 Consensus Not Reached
- Lead Discussant: Patrick Romano; Secondary Discussant: Daniel Deutscher
- Measure Developer: Kidney Care Quality Alliance
- Measure Steward: Kidney Care Quality Alliance
- Discussion Guide page 25
- For SMP discussion:
 - Additional clarifying information from the developer
 - Are the methods appropriate for testing reliability for this measure?
 - Is the measure calculation by individual facility or HRR as structured appropriate?
 - Is the high rate of identification of outliers in the reported measure scores a cause for concern?
 - Are the face validity results sufficient to meet the validity requirements?
 - Is the developer's rationale for not assessing risk adjustment independently (of the paired measure NQF 3679) for this measure sufficient or should independent risk adjustment analyses be conducted?



#1460 Bloodstream Infection in Hemodialysis Outpatients

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-0, M-3, L-4, I-2 No Pass
 - Validity: H-0, M-4, L-4, I-2 Consensus Not Reached
- Lead Discussant Reliability: Terri Warholak
- Lead Discussant Validity: Jeff Geppert
- Measure Developer: Centers for Disease Control and Prevention
- Measure Steward: Centers for Disease Control and Prevention
- Discussion Guide page 5
- For SMP discussion:
 - Are there any concerns about the overall underreporting of bloodstream infections?
 - Are the testing results sufficient to demonstrate reliability and validity?



#0471e ePC-02 Cesarean Birth

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-0, M-3, L-4, I-4 No Pass
 - Validity: H-0, M-3, L-4, I-4 No Pass
- Lead Discussant: Sam Simon; Secondary Discussant: Paul Kurlansky
- Measure Developer: The Joint Commission
- Measure Steward: The Joint Commission
- Discussion Guide page 6
- For SMP discussion:
 - Additional clarifying information from the developer, including whether the correct data elements were assessed for each measure
 - The SMP did not pass this measure for reliability and validity with concerns about testing results. Is there further argument that the testing results are in fact sufficient to demonstrate encounter-level reliability/validity.



#0716e ePC-06 Unexpected Newborn Complications in Term Newborns

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-1, M-2, L-4, I-4 No Pass
 - Validity: H-0, M-5, L-3, I-3 Consensus Not Reached
- Lead Discussant: Paul Kurlansky; Secondary Discussant: Sam Simon
- Measure Developer: The Joint Commission
- Measure Steward: The Joint Commission
- Discussion Guide page 9
- For SMP discussion:
 - Additional clarifying information from the developer, including whether the correct data elements were assessed for each measure
 - The SMP did not pass this measure on reliability and was consensus not reached on validity. Is there further discussion that would provide support to demonstrate reliability and validity of the measure?

Opportunity for Public Comment

Next Steps



Next Steps

- Tomorrow's (3/23) meeting will be from 1:00PM-3:00PM EST
- Agenda
 - Measure Methodology Discussion
 - » 2820
 - » 3687e

Adjourn

Day 2: Welcome, Disclosures of Interest, Review Agenda

Welcome



Housekeeping Reminders – Day 2

This is a Webex meeting with audio and video capabilities:

Meeting link: https://nqf.webex.com/nqf/j.php?MTID=mdd23063d22e2c46233463788664b88f9

Meeting number: 2335 654 2657

Password: MScAEvent

Optional: Dial 1-844-621-3956 and enter passcode

- Please place yourself on mute when you are not speaking
- We encourage you to use the following features
 - Chat box: to message NQF staff or the group
 - Raise hand: to be called upon to speak
- We will conduct Scientific Methods Panel roll call once the meeting begins

If you are experiencing technical issues, please contact the NQF project team at methodspanel@qualityforum.org



Meeting Ground Rules

- No rank in the room
- Remain engaged and actively participate
- Be prepared, having reviewed the measures beforehand
- Base evaluation and recommendations on the measure evaluation criteria and guidance
- Keep comments concise and focused
- Be respectful and allow others to contribute
- Share your experiences
- Learn from others



Meeting Agenda: Day 2

- Welcome and Recap of Day 1
- Measure Methodology Discussion
- Opportunity for Public Comments
- Next Steps
- Adjourn



Scientific Methods Panel Members (Cont.)

- David Nerenz, PhD, Co-chair
- Christie Teigland, PhD, Co-chair
- J. Matt Austin, PhD
- John Bott, MBA, MSSW
- Daniel Deutscher, PT, PhD
- Marybeth Farquhar, PhD, MSN, RN
- •Jeffrey Geppert, EdM, JD
- Laurent Glance, MD
- Joseph Hyder, MD
- Sherrie Kaplan, PhD, MPH
- Joseph Kunisch, PhD, RN-BC, CPHQ
- Paul Kurlansky, MD

- Zhenqiu Lin, PhD
- Jack Needleman, PhD
- Eugene Nuccio, PhD
- Sean O'Brien, PhD
- Jennifer Perloff, PhD
- ■Patrick Romano, MD, MPH
- Sam Simon, PhD
- •Alex Sox-Harris, PhD, MS
- Ronald Walters, MD, MBA, MHA, MS
- ■Terri Warholak, PhD, RPh, CPHQ, FAPhA
- Eric Weinhandl, PhD, MS
- Susan White, PhD, RHIA, CHDA

Recap of Day 1

Measure Methodology Discussion



#3687e ePC-07 Severe Obstetric Complications

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-4, M-5, L-1, I-0 Pass
 - Validity: H-2, M-6, L-0, I-2 Pass
- Lead Discussant: Jack Needleman
- Measure Developer: The Joint Commission
- Measure Steward: The Joint Commission
- Discussion Guide page 14
- For SMP discussion:
 - Correlation of the hospital level rates of transfusions and the non-transfusion components of the measure
 - How is the social risk factor variable in the risk adjustment model (economic/housing instability) collected or recorded?
 - Clarification on testing results on non-transfusion cases in the risk adjustment model



#2820 Pediatric Computed Tomography (CT) Radiation Dose

- Subgroup 1
- Preliminary Voting Result:
 - Reliability: H-5, M-4, L-0, I-1 Pass
 - Validity: H-1, M-7, L-1, I-1 Pass
- Lead Discussant: Alex Sox-Harris
- Measure Developer: University of California, San Francisco
- Measure Steward: University of California, San Francisco
- Discussion Guide page 12
- For SMP discussion:
 - Additional clarifying information from the developer
 - Additional clarification on the method for scoring this measure and how it identifies outliers

Opportunity for Public Comment

Next Steps



Next Steps and Reminders

- Full measure submission deadlines: April 4 and 11
- NQF staff will summarize the relevant measure information and discussions of the SMP and provide to the various Standing Committees
 - » These Standing Committees will evaluate measures in June/July
 - » CSAC will review Spring measures in the November/December timeframe
- Next Intent to Submit deadline (Fall 2022): August 1



2022 SMP Meetings

- April 27, 2022
 - 10:00 AM 12:00 PM EST
- May 24, 2022
 - 12:00 PM 2:00 PM EST
- **July 14, 2022**
 - 12:00 PM 2:00 PM EST



Project Contact Info

■ Email: <u>MethodsPanel@qualityforum.org</u>

■ NQF phone: 202-783-1300

Project page:
 http://www.qualityforum.org/Measuring Performance/Scientific Methods Panel.aspx

SharePoint site:
https://share.qualityforum.org/portfolio/ScientificMethodsPanel/SitePages/Home.aspx

Adjourn

THANK YOU.

NATIONAL QUALITY FORUM

https://www.qualityforum.org