NATIONAL
QUALITY FORUM
Driving measurable health
improvements together

Rural-Relevant Quality Measures for Testing of Statistical Approaches to Address Low Case-Volume

SUMMARY OF RECOMMENDATIONS

September 25, 2020

This report is funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I – 75FCMC19F0007.

Table of Contents

Rural-Relevant Quality Measures for Testing of Statistical Approaches to Address Low Case-Volume 1
Table of Contents 2
Background 3
Process 3
Measure Recommendations 6
Measure List Characteristics 11
Reporting Challenges 12
Gaps and Future Considerations for Rural Health Measurement 13
Conclusion 14
Appendix A: Rural Health Workgroup Members and NQF Staff 16
Appendix B: Public Comments 18

Background

Low case-volume poses a measurement challenge for many healthcare providers in rural areas. Low population density, in combination with limited access to care, can reduce the number of patients eligible for inclusion in healthcare quality measures in Medicare public reporting and value-based purchasing programs. These low sample sizes affect the reliability and validity of measure scores, making it difficult to compare performance between providers or track changes in quality over time.

NQF convened the multistakeholder Measure Applications Partnership (MAP) Rural Health Workgroup (Workgroup), which included clinicians and healthcare providers, state and local agency staff, healthcare consumers, representatives of private nonprofit organizations, and other experts with background in rural or tribal areas to help identify performance measures that are high impact and meaningful to rural Americans, feasible for providers to report to Medicare programs, and resistant to low case-volume challenges. The resulting Core Set of Rural-Relevant Measures (Core Set) was released in the report of this work published in August 2018. The Core Set included both cross-cutting measures and condition-specific measures pertinent to rural populations, including measures on mental health, substance abuse, medication reconciliation, diabetes, hypertension, chronic obstructive pulmonary disease (COPD), hospital readmissions, perinatal care, and pediatric care. During public commenting, stakeholders commented that six of the measures in the Core Set may face low case-volume challenges in some areas (read full report here).

To further advance measurement science related to case-volume, in 2018, Centers for Medicare & Medicaid Services (CMS) tasked NQF with eliciting expert input on promising statistical approaches that could be used to address the low case-volume challenge. NQF convened a Technical Expert Panel (TEP) which made four recommendations: "borrow strength" for low case-volume rural providers by incorporating additional data (e.g., from past performance, other providers, other measures); recognize the need for robust statistical expertise and computational power to implement "borrowing strength"; report exceedance probabilities, which reflect the uncertainty of measure scores; and actively anticipate the potential for unintended consequences of measurement. The TEP also made recommendations for future activities, including testing the "borrowing strength" approach through activities such as simulation studies or challenge grants (read full report here).

In fall 2019, building upon previous efforts, NQF was tasked with identifying a list of high-priority, rural-relevant measures susceptible to low case-volume challenges for future testing of the TEP's recommended statistical approaches. To accomplish this objective, NQF performed an environmental scan and convened several web meetings of the Workgroup to develop a priority measure list and discuss reporting challenges specific to measurement in rural areas.

Process

Building on recommendations from the Workgroup and the TEP, NQF completed an environmental scan of approximately 250 rural-relevant quality measures included in Medicare quality reporting and value-based purchasing programs that are advised upon by MAP. NQF also included measures used in select Center for Medicare & Medicaid Innovation (CMMI) Alternative Payment Models (APMs). These included the Oncology Care Model, Bundled Payments for Care Improvement Advanced, Next Generation ACO Model, and Comprehensive Primary Care Plus. Quality measures used in these models

were considered for inclusion based on the models' high profile and experience in using quality measurement to incentivize delivery of high-quality care and efficient use of healthcare resources.

Measures were deemed rural relevant if they addressed topics previously identified as rural relevant by the Workgroup or defined as rural relevant in published literature. NQF extracted measures using the Centers for Medicare & Medicaid Services Measures Inventory Tool (CMIT). CMIT filters such as conditions, sub-conditions, Meaningful Measure area, current status, as well as key word searches, were used to identify measures that relate to rural-relevant topics and are implemented or finalized in federal programs. The environmental scan included measure titles, reference numbers, NQF endorsement, measure types, measure specifications, risk adjustment data, minimum case requirements, and rural-relevant topics addressed for each included measure. Data not included in CMIT, like program-specific minimum case numbers and risk adjustment factors, were collected manually from program resources, technical manuals, NQF measure repository notes, and other sources.

After developing the initial list of measures, NQF discussed the environmental scan methodology and initial results with the Workgroup during a web meeting on May 6, 2020. The Workgroup provided initial input on important measure attributes and topics to consider during measure prioritization.

NQF then fielded a brief survey with the Workgroup asking members to rate the importance of different measure attributes and to select high priority topic areas that build on the current Core Set. The Workgroup recommended the following topics and attributes be prioritized to identify measures that would be suitable candidates for the statistical testing.

Topics prioritized:

1. Access to care
2. Vaccinations
3. Cancer screening
4. Stroke
5. Healthcare-associated infections (HAIs)
6. Emergency department use

The Workgroup prioritized access to care, noting that it is the most relevant issue for rural health and healthcare and remains an important measurement gap area. Vaccinations and cancer screening were considered important aspects of preventive care that may not be received by rural residents in a timely manner due to access issues. Stroke was emphasized as an important issue for rural residents due to comparatively higher mortality rates. Infections such as catheter-associated urinary tract infections (CAUTI) and hospital-onset Clostridium difficile infection (CDI) were noted as important threats to patient safety that are addressed by existing quality measures and programs. But it was also noted that rural hospitals are not subject to these programs and can have challenges reporting on the measures due to low case-volume. Emergency department use was considered an important topic, and, in particular, communication around patient transfers; measures on admit-to-discharge time were considered not as relevant in rural contexts. Also identified as important were the topics of end-of-life/advance directives, pneumonia, heart failure, surgical care, heart attack, asthma, and obesity.

Measure attributes used for prioritization:

1. NQF endorsement

2. Outcome measures, especially patient-reported outcome-based performance measures (PRO-PMs)
3. Cross-cutting measures
4. Measures used in multiple federal programs

Rationale for each measure attribute were that NQF endorsement indicates scientific acceptability of measure properties, feasibility, usability, and evidence of a performance gap. Outcome measures and PRO-PMs assess the impact of a healthcare service or intervention on health status or experience of a patient and emphasize patient-centeredness. Cross-cutting measures reflect broad applicability to patient populations by not limiting measurement to a specific diagnosis or process, and therefore could reach many patients. The Workgroup recognized the need to balance the inclusion of cross-cutting measures as well as condition-specific measures in the candidate measures for testing. This approach would acknowledge the importance of both measure types and recognize that singular focus on cross-cutting measures may neglect focus on quality measurement in rural-relevant, specialty areas. The criterion “use in multiple federal programs” could mean greater ability of rural providers to participate in federal programs if the statistical approaches were found to be successful.

NQF used the Workgroup’s importance ratings for each attribute to develop a composite score that was assigned to each measure in the environmental scan. The Workgroup’s importance rating was based on averaged Likert scale responses to each attribute (0=not important; 1=slightly important; 2=moderately important; 3=important; 4=very important). NQF staff then tagged each measure with a “1” or “0” to indicate whether or not the measure was NQF-endorsed, an outcome or PRO-PM, cross-cutting, or used in multiple federal programs. Staff then multiplied each attribute 1 or 0 by a distributed attribute “weight,” based on the averaged importance rating, and summed each component to obtain the composite score for each measure. NQF then grouped measures into high-priority rural-relevant topic areas and selected high-scoring measures within each group for further consideration.

Measure Attributes	**Averaged Importance Likert Rating**	**Distributed Attribute “Weight”**
NQF Endorsement	3.28	0.28
Outcome or PRO-PM	3	0.26
Cross-Cutting	2.91	0.25
Use in Multiple Programs	2.36	0.20
Example composite score calculation:		
Measure	Follow-Up After Hospitalization for Mental Illness (NQF# 0576)	
Attributes	NQF Endorsed (Yes): 1; Outcome Measure (No): 0; Cross-Cutting (Yes): 1; Use in Multiple Programs (Yes): 1	
Score calculation	(1*0.28) + (1*0.25) + (1*0.20) = 0.73	

In selecting measures for consideration, NQF attempted to ensure an adequate mix of measure type, risk adjustment, use in programs, care settings, and reporting levels. This resulted in a short list of approximately 40 measures. The shortlist also included six measures from the Core Set; stakeholders shared during previous public comment that these measures may pose measurement challenges due to low case-volume. The shortlist was then shared with the Workgroup to offer an opportunity for members to recommend removal or addition of specific measures. The Workgroup recommended removal of four measures addressing coronary bypass artery graft procedures since they are not performed very often in small rural hospitals and one measure addressing overuse of bone scan for

staging low-risk prostate cancer patients as the group was more concerned about underuse of this imaging vs overuse in rural settings. The Workgroup decided to add measures #0500 *Severe Sepsis and Septic Shock* and #0277 *PQI-08 Heart Failure Admission Rate* to the shortlist of measures for consideration.

During extended web meetings on May 27 and May 29, 2020, the Workgroup had in-depth discussions on 34 measures. Individual Workgroup members were randomly assigned as lead discussants and were asked to provide initial reactions to five questions:

1. Is the measure problematic due to low case-volume and why?
2. Is the measure pertinent to the rural population and does it have a significant impact on patient care?
3. Does the hospital/clinician have influence over measure performance?
4. What is the opportunity for performance improvement?
5. Is the measure feasible to report for rural providers?

NQF staff and Workgroup co-chairs facilitated group discussion on each measure and, following discussion, Workgroup members voted to recommend or not recommend measures for statistical testing. Measures that received a “yes” vote by 60% or more made it to the final recommendations list.

A draft of this report along with an accompanying spreadsheet (XLSX) was posted on the NQF website from July 10 to July 30 for public comment. Comments received were discussed and addressed in a web meeting with the Workgroup on August 26. This web meeting was announced in advance and made available to all NQF members and members of the public. Comments and feedback received have been incorporated into this document.

Measure Recommendations

The Workgroup selected 15 measures susceptible to low case-volume and recommended they be prioritized for future testing of statistical approaches to overcome this challenge. This measure list puts forth recommendations for prioritizing which measures should be tested. It is not intended to represent the Workgroup’s opinion of the measures’ appropriateness for use in specific federal programs, nor are these measures being considered for addition to the Core Set at this time. The 15 measures are listed below, along with their rationale for inclusion and reporting challenges and are described in further detail in an Excel spreadsheet available online on the MAP Rural Health project page.

CMIT #	NQF #	Measure Title	Rationale for Inclusion	Reporting Challenges
2517	0005	Consumer Assessment of Healthcare Providers & Systems (CAHPS) Clinician/ Group Survey	This measure is currently included in the Core Set but public comments suggest that the similar Hospital CAHPS (HCAHPS) measure is challenging to report on due to low case-volumes in rural areas. The Workgroup agreed that the clinician has influence over measure performance and that this measure is pertinent to rural populations and impacts care.	The Workgroup noted that feasibility of data collection is a problem for Critical Access Hospitals (CAHs) due to cost and reporting rules that are difficult for rural providers to meet. Limiting allowable data collection to either mail-in surveys or via telephone creates undue administrative burden and is one reason this measure is challenging to

CMIT #	NQF #	Measure Title	Rationale for Inclusion	Reporting Challenges
				report on in rural settings. The Workgroup recommends that CMS consider allowing electronic data collection. Despite these challenges, this survey is used widely in different programs and it would be helpful to apply the statistical testing approaches to this measure to assess reliability.
113	0166	HCAHPS	This measure is currently included in the Core Set, but public comments suggest that it is challenging to report on due to low case-volumes in rural areas. Public comments suggest that nearly 60% of CAHs submitting HCAHPS data do not meet the CMS Star Rating threshold of 100 completed surveys over four quarters, and 12% of reporting CAHs had fewer than 25 surveys returned. The Workgroup noted that this measure is rural relevant, impacts care, and is influenced by clinicians.	The Workgroup cited the same reporting challenges for this measure that are outlined for #0005 (above). It was noted that CAHPS and HCAHPS have similar data collection processes, and if resources are limited, HCAHPS should be prioritized for statistical testing.
2046	2079	HIV Medical Visit Frequency	The Workgroup agreed that this measure faces reporting challenges due to low case-volume, is pertinent to a rural population, and has a significant impact on patient care. The measure was noted as important from a health equity perspective, as African American patients are disproportionately represented among rural HIV cases. The Workgroup also noted that this measure addresses access to care—a critically important issue for rural health—and that the current Core Set does not include any HIV measures.	Measure performance may be impacted by factors outside of a clinician's control, such as lack of transportation options for rural patients with HIV. The Workgroup recommends that "medical visit" include a telehealth option. The Workgroup noted that this measure is endorsed by NQF at the facility level, but that it is analyzed in the Merit-Based Incentive Payment System (MIPS) at the clinician level.
2519	0108	Follow-Up Care for Children Prescribed Attention-Deficit/Hyperactivity Disorder (ADHD) Medication (ADD-CH)	The Workgroup noted that this measure does not face low case-volume challenges at the health plan level, as endorsed by NQF, but that it does at the group practice/clinician level, as it is used in MIPS reporting. The measure was considered pertinent to rural populations and impactful, especially given the implications for mental health	It was noted that clinicians have some influence on measure performance by initiating follow-up, but that the actual number of visits might depend on patient-level factors.

CMIT #	NQF #	Measure Title	Rationale for Inclusion	Reporting Challenges
			and substance use later in life. The Workgroup noted that MIPS data demonstrate an opportunity for performance improvement. While the Workgroup voted to recommend this measure for statistical testing, there was uncertainty around including it in the Core Set in the future, as there may be more broadly applicable behavioral health measures that could be prioritized.	
745	0576	Follow-Up After Hospitalization for Mental Illness	The Workgroup discussed that the measure would not have low case-volume problems at the health plan level, as endorsed by NQF, but likely faces low case-volume reporting challenges at the clinician level, as used in MIPS. The measure was considered rural relevant and impactful; feasible for clinicians to report from existing claims data; and has opportunity for improvement. Because this measure includes patients starting at the age of six years, it does address pediatric health. Measures related to transitions of care, such as this one, are a priority, though not many measures on this topic were available for potential inclusion.	A shortage of behavioral health specialists in rural areas creates a challenge in ensuring timely follow-up for behavioral health appointments. Measures related to transitions of care need to be feasible for rural providers. Exchanging data can be difficult in some rural facilities as they may not have integrated data systems.
2818	0275	Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate (PQI05-AD)	The Workgroup noted that the measure was rural relevant and impactful; the clinician would have influence over the measure performance (especially in team-based care); there was opportunity for improvement; and the measure would likely be feasible to report as it is claims-based. There was some uncertainty around whether this measure truly faces low case-volume reporting challenges, but the Workgroup consensus was that it may face these challenges at the group/practice level.	The Workgroup recommends that this measure be tested at the group/practice level, rather than at the population level.
1364	0138	National Healthcare Safety Network (NHSN) Catheter-Associated Urinary Tract Infection	This measure is currently included in the Core Set, but public comments suggest that it is challenging to report on due to low case-volume in rural areas. Public comments suggest that it is vital for CAHs to be reporting healthcare-associated infection data to the NHSN, but that few CAHs have enough cases	The Workgroup recommends that in analysis and testing, the final product should provide guidance on whether differences in infections between individual facilities can be determined given low

CMIT #	NQF #	Measure Title	Rationale for Inclusion	Reporting Challenges
		Outcome Measure	for a quality metric of a standardized infection ratio (SIR) to be calculated on a quarterly or even annual basis. The Workgroup noted that this measure is a high priority for rural populations, feasible to report, has opportunity for performance improvement, and that measure performance is under a clinician's influence as there are clear guidelines for using catheters appropriately.	case-volumes in the rural setting.
831	1717	National Healthcare Safety Network (NHSN) Facility-wide Inpatient Hospital-onset *Clostridium difficile* Infection (CDI) Outcome Measure	This measure is currently included in the Core Set, but public comments suggest that it is challenging to report on due to low case-volumes in rural areas. The Workgroup agreed that this measure encompassed important topics, including environmental hygiene, infection and prevention control policies, and antibiotic stewardship.	Previous public comments suggest that it is vitally important for CAHs to be reporting healthcare-associated infection data to the NHSN but that few CAHs have enough cases for a quality metric of a standardized infection ratio to be calculated on a quarterly or annual basis.
2831	0471	PC-02 Cesarean Birth	This measure is currently included in the Core Set, but public comments suggest that it is challenging to report on due to low case-volume in rural areas. The Workgroup discussed that this measure was rural relevant, demonstrated an opportunity for improvement due to uneven performance, could be influenced by the clinician, and was feasible to report because of the option to pull data from electronic health records. The group also noted that the measure was risk-adjusted but did not include adjustment based on the type of provider performing the C-section, and also had a number of exclusions (e.g., it is only for first-time mothers who are not transferred to another facility for care, medical exclusions also apply).	N/A
182	0173	Emergency Department Use without Hospitalization During the First 60 days of Home Health	The Workgroup agreed that this measure was important for care, could be influenced by the clinician, demonstrated room for improvement, and was feasible to report. It is also related to home health—a setting not captured by other measures on the list.	N/A

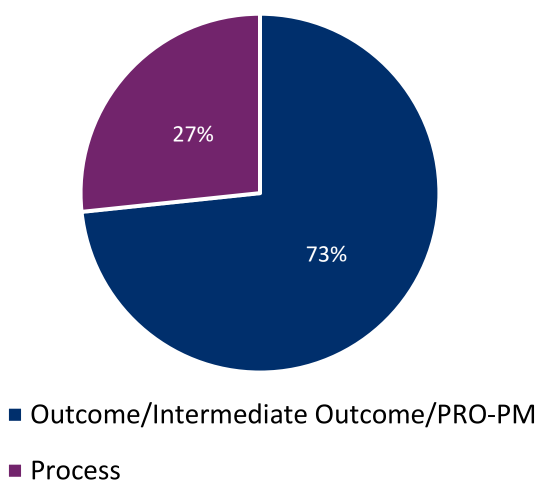
CMIT #	NQF #	Measure Title	Rationale for Inclusion	Reporting Challenges
6040	1789	Risk-Standardized, All Condition Readmission	This measure is currently included in the Core Set, but public comments suggest that it is challenging to report on due to low case-volumes in rural areas. The Workgroup agreed that there is room for performance improvement and that this measure is feasible to report on as it is based on claims data. It was noted that this measure is endorsed at both the Accountable Care Organization (ACO) and facility levels, and that at the facility level, it is likely challenging to report due to low case-volume. It was also noted that CMS has indicated they will be shifting to a hybrid version of this measure and will no longer be utilizing the claims-only measure in the future. It will be important to take this into consideration in deciding which version to use for statistical testing.	The Workgroup recommends this measure be tested at the facility level.
2432	2510	Skilled Nursing Facility 30-Day All-Cause Readmission Measure (SNFRM)	The Workgroup discussed that this measure is subject to low case-volume reporting challenges and is feasible to report on as it is claims-based and reported at the nursing facility level.	The Workgroup expressed uncertainty that a clinician would have significant influence over measure performance, which may be explored during testing.
899	1551	Hospital-Level 30 Day, All-Cause, Risk-Standardized Re admission Rate (RSRR) Followin g Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplas ty (TKA)	The Workgroup expressed that this measure is susceptible to low case-volume reporting challenges, is pertinent and impactful to an aging rural population, and is feasible to report as it is already used for reporting through Hospital Compare and can be influenced by clinicians.	This is a useful benchmark that has been used without adjustment, but one or two additional readmissions can greatly impact performance for some facilities. It was also noted that CMS publicly reports this information, but it is very hard to use it to distinguish between facilities as the numerator is so low and most hospitals report a rate of zero for this measure.
2086	2539	Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colo noscopy (OP32)	The Workgroup stated that clinicians have some influence over this measure and that it is feasible to report. It was noted that colonoscopies are procedures that bring patients into the healthcare system and serve as an access point for care, and patients express that they do not want to travel to receive colonoscopies, rendering this measure impactful and rural relevant.	N/A

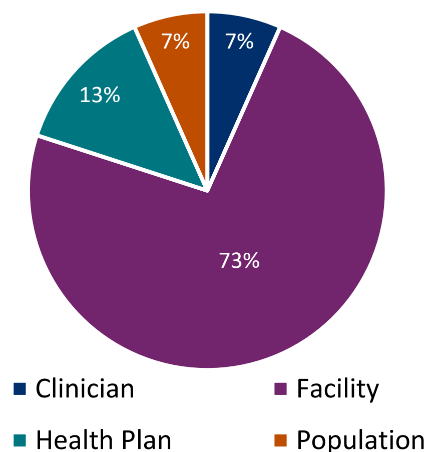
CMIT #	NQF #	Measure Title	Rationale for Inclusion	Reporting Challenges
1017	0500	Severe Sepsis and Septic Shock: Management Bundle	The Workgroup agreed that this measure was subject to low case-volume reporting issues in the rural context. The measure would be a high-value inclusion for improving care for a mix of provider types and noted that it is under exploration for potential inclusion in the Medicare Beneficiary Quality Improvement Project (MBQIP) as a measure for CAHs.	Small rural facilities may transfer patients with sepsis to larger facilities to finish treatment. The question was raised as to whether a smaller part of the composite might be appropriate to measure for small rural hospitals. Some rural hospitals do treat sepsis in full. For those that do not, the measure could be used to address whether care was managed correctly up to the point of transfer.

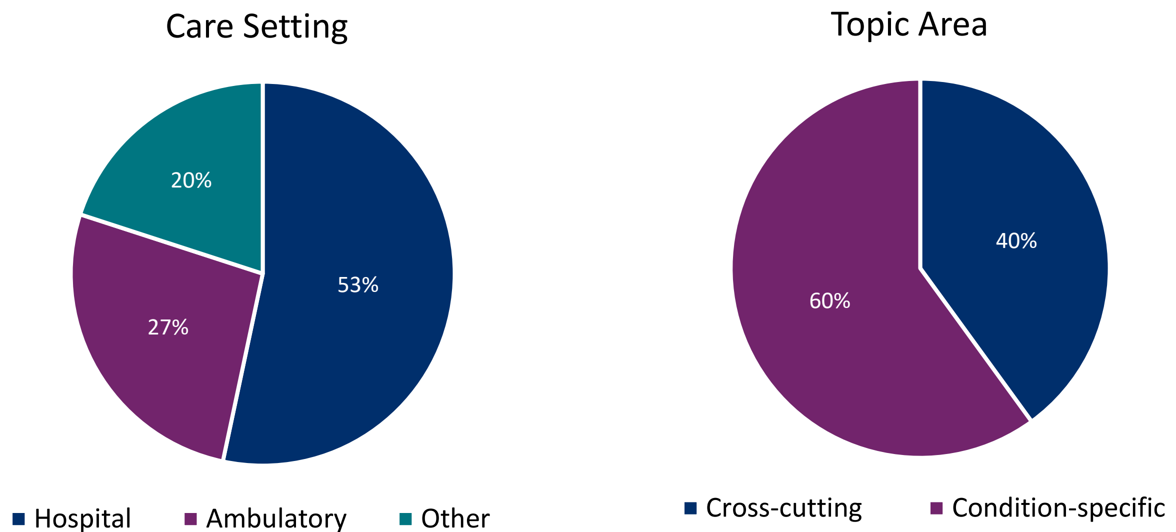
To differentiate this measure list from the Core Set of Rural-Relevant Measures created in 2018, measures included in the Core Set are intended to be widely implementable and resistant to low case-volume challenges. Measures on this list are required to be susceptible to low case-volume challenges that may limit their usefulness for making performance comparisons or driving quality improvement for rural providers. Note this measure list includes several measures in the Core Set due to public comments suggesting they may face low-case volume challenges.

Measure List Characteristics

The Workgroup emphasized the importance of achieving an adequate mix of measure attributes and topic areas in the final list of recommended measures for testing of the statistical approaches. Several aspects of the measures that were deemed important to vary included measure type, level of analysis, care setting, and cross-cutting versus condition-specific topic area. These attributes were considered when making decisions about which measures to include in the final list. Characteristics of these 15 measures are highlighted below.







Note that some measures included on this list are analyzed at varying levels and likely do not face low case-volume challenges at the health plan or population level. For these measures, the Workgroup recommends applying statistical testing at a more granular level to assess appropriateness of use at, for example, the clinician or facility level of analysis. If the statistical testing approaches are successful and appropriate contextual factors are considered during testing, measures that were once only appropriate for reporting at higher levels of analysis (e.g., health plan or population) may be suitable to assess performance at other levels of analysis. The Workgroup encouraged this to be carefully investigated during statistical testing.

Reporting Challenges

Several themes emerged from the Workgroup's discussion on reporting challenges.

Data Challenges

Some measures discussed did not seem to have case-volume challenges based on the denominator population. However, certain rural providers can still face difficulties obtaining the data needed to meet measure requirements or to inform care decisions. Data challenges include lack of sufficient information flow from specialists to primary care providers, between providers in rural and urban areas, and/or between providers and other entities (e.g., payers). The Workgroup also discussed the potential impact of expanding the universe of claims included in the calculation of claims-based measures to incorporate Medicare Advantage or all-payer data. To ensure comparability between and across providers, if this were to be considered, it should apply to all hospitals/providers in all settings, not just those in rural areas.

Measure Reporting Options

The Workgroup appreciated the movement towards greater use of electronic clinical quality measures (eCQMs) to reduce burden but highlighted several considerations related to their use by rural providers.

- Differing availability of certain data sources (e.g., access to electronic health record (EHR) data) in rural care settings
- Lack of clarity regarding how many CAHs, as well as other rural hospitals, are reporting (and using) eCQMs
- Rural providers are less likely to be using one of the major EHR companies and are usually using smaller, less expensive, and less advanced EHR systems

- Rural providers are less likely to have in-house expertise to perform data extraction and analysis
- Rural providers are more likely to be independent and not part of a larger system, which may negatively impact their performance on measures relying on inter-provider data communication
- eCQMs in some CMS programs (e.g., Hospital Inpatient Quality Reporting Program and Promoting Interoperability Programs) are of limited relevance to rural providers
- Some rural providers are changing workflows and documentation processes in order to participate in eCQM reporting, but there is opportunity for the measure results to be more actionable for rural providers in order to drive improvement
- The Workgroup recommended that the CAHPS measures should have electronic data collection options

Infrastructure Requirements

In order to effectively implement the TEP-recommended borrowing strength statistical approaches, infrastructure is an important consideration. This includes the robust statistical expertise and computational power that would be needed to establish benchmarks or thresholds, observe statistical correlations or persistence, and estimate correlated signal variances.

Measure Alignment

Measures related to hospital or emergency department visits after certain procedures should be aligned to the extent possible. For example, measures used to address the quality of surgical procedures should be aligned across ambulatory surgical centers and outpatient facilities.

Unintended Consequences Related to Statistical Testing

One approach to overcoming low case-volume is to pool data over several years for one provider, but the Workgroup noted that this would affect the ability to track improvement over time due to lag, which might pose a challenge for pay-for-performance programs intended to serve this purpose. Implementing borrowing strength approaches that leverage persistent statistical relationships over time, when and where possible, may mitigate this problem. Measure attribution should be carefully considered during testing. For example, physician assistants and nurse practitioners may be the actual providers of care in many cases; however, services are required to be submitted under the supervising physician. Post testing, implementation of measures that rely on statistical methods to address case-volume challenges should carefully consider program characteristics and intent. Additionally, it is crucial to monitor unintended consequences to ensure that measures used to assess care provided in rural areas do not reduce access to care, disincentivize providers from offering certain types of care in rural or underserved areas based on risk of reduced payment, or encourage providers to avoid providing procedures like caesarean sections—even when in the best interest of individual patients.

Gaps and Future Considerations for Rural Health Measurement

The Workgroup identified the following gap areas related to quality measurement in rural areas.

Person-Centered Measurement

The Workgroup encouraged the balance of using quantitative measures that are easier to capture with measures that use qualitative methodologies to represent patient and caregiver voices and experiences (e.g., patient-reported information). A core aspect of person-centered care is ensuring that patients

actively participate in their own healthcare, including decision making around advance care directives and end-of-life care. These continue to be important topics that require measure adaptation and development for rural providers and patients.

Measures Related to COVID-19 and Telehealth

The healthcare system is continually evolving, even more so recently with major changes in the delivery of care due to the COVID-19 pandemic. The measurement enterprise should consider the impact of these changes and of the consolidation and regionalization of healthcare. COVID-19 has also exacerbated barriers to health equity and the role of social determinants of health. It is important to recognize these disparities for rural communities and other underserved populations. Infection prevention, health system preparedness, patient resilience, and health system resilience are areas in need of greater data and opportunities for advancing measurement. In addition, due to a rapid increase in the use of telehealth to provide more services, measure specifications should consider and include this technology when appropriate. However, there is also a need to better understand limitations of services that are delivered virtually and if there are differences in the quality of care delivered virtually versus in-person, especially for chronic illness care. There may also be concerns about telehealth access issues for rural communities (e.g., lack of access to high-speed internet).

Community and Population Health

Community-based measures (e.g., those that assess systems of care across a community), keeping populations healthy, and correlating access to care with population health outcomes are areas that are currently not adequately addressed by quality measurement. There are challenges for rural providers in communicating quality information across care settings. The Workgroup suggested that better information flow and communication between providers, payers, and community-based human service agencies could drive greater coordination of patient care and improved patient outcomes. There is also an opportunity for greater coordination among health plans, health systems, and community-based organizations to drive sustainable improvements in care. To fully capture the healthcare areas most important to patients in rural communities, it was suggested that it may be helpful to supplement the Core Set with population-based measures that could assess characteristics that may be difficult to evaluate at the provider or facility level (e.g., social, economic, and environmental determinants of health, community-level indicators of health and disease, prevention programs).

Conclusion

The Workgroup used a multistakeholder, consensus-based process to select 15 rural-relevant measures that should be prioritized for testing statistical approaches to address low case-volume. These measures cover a range of topics relevant to healthcare quality for rural populations—patient experience, access to care, behavioral health, COPD, HAIs, perinatal care, readmissions, transitions of care, and sepsis. They represent a mix of measure types, analysis levels, and care settings.

Although CMS is the primary audience for the recommendations in this report, other healthcare measurement stakeholders can benefit from understanding the opportunities to advance quality measurement in rural settings. If future testing to overcome case-volume challenges proves successful, this measure list may represent a key source of rural-relevant measures that can be considered for use in measurement programs. The creation of this prioritized list is an important step towards achieving

high-quality and high-value outcomes for all Americans, regardless if their area of residence is rural or geographically remote.

This work serves as the basis for advancing approaches that can make performance measurement more useful for providers and patients in rural areas as well as other stakeholders that have rural members. Future related work of the Workgroup may include reviewing the statistical testing results for these measures and determining if the measures are appropriate for inclusion in the Core Set of Rural-Relevant Measures.

Appendix A: Rural Health Workgroup Members and NQF Staff

Workgroup

Ira Moscovice, PhD, Co-Chair
University of Minnesota School of Public Health

Aaron Garman, MD, Co-Chair
Coal Country Community Health Center

Kimberly Rask, MD, PhD, FACP
Alliant Health Solutions

David Schmitz, MD, FAAFP
American Academy of Family Physicians (AAFP)

Daniel Coll, MHS, PA-C, DFAAPA
American Academy of PAs

Margaret Greenwood-Ericksen, MD
American College of Emergency Physicians

Stephen Tahta, MD
American Hospital Association

Erika Thomas, M.B.A., B.S.Pharm
American Society of Health-System Pharmacists

Jennifer Greene
Cardinal Innovations

Karen Murphy, PhD, RN
Geisinger Health

Jesse Spencer, MD
Intermountain Healthcare

Crystal Barter, MS
Michigan Center for Rural Health

Julie Sonier, MPA
Minnesota Community Measurement

Bill Finerfrock
National Association of Rural Health Clinics

Brock Slabach, MPH, FACHE
National Rural Health Association

Cameron Deml
National Rural Letter Carriers' Association (NRLCA)

Keith Mueller, PhD
RUPRI Center for Rural Health Policy Analysis

Tim Size, MBA
Rural Wisconsin Health Cooperative

Heather Brown-Palsgrove
IBM Watson Health Company

Michael Fadden, MD
Cerner

John Gale, MS
Maine Rural Health Research Center, University of Southern Maine

Curtis Lowery, MD
University of Arkansas for Medical Sciences

Melinda Murphy, RN, MS
Melinda Lew Murphy

Jessica Schumacher, PhD
University of Wisconsin - Madison, Surgical Collaborative of Wisconsin

Ana Verzone, MS, APRN, FNP, CNM
Alaska Native Medical Center

Holly Wolff, MHA
Ashley Medical Center

Craig Caplan (non-voting)
Federal Office of Rural Health Policy, DHHS/HRSA

Bruce Finke (non-voting)
Indian Health Services, DHHS

Emily Moore (non-voting)
Center for Medicare & Medicaid Innovation, Centers for Medicare & Medicaid Services (CMS)

NQF Staff

Sheri Winsper, RN, MSN, MSHA
Maha Taylor, MHA, PMP
Nicolette Mehas, PharmD
Andre Weldy, MPH
Mike DiVecchia, MBA, PMP
Amy Guo, MS
Kabir Suri

Appendix B: Public Comments

Battelle Memorial Institute (commenter: Jeffrey Geppert)

Measure Recommendations: Do you agree that the 15 measures recommended for statistical testing are relevant for rural populations and are susceptible to low case-volume reporting challenges? If not, please provide feedback.

Among the measure attributes used for prioritization item #3 (Cross-cutting measures reflect broad applicability to patient populations by not limiting measurement to a specific diagnosis or process) may be somewhat contrary to the TEPs recommendation. The use of cross-cutting measure was one of the measurement recommendations the TEP considered. Specially (p. 9)

The TEP also noted some potential drawbacks of this approach. In particular, limiting the selection of measures to those that are applicable for most rural providers places artificial constraints on the available measures. This could result in the neglect of other measures that are important for rural populations. For example, a focus on screening or immunizations might jeopardize quality improvement efforts in rural areas for other important conditions or healthcare activities such as specialty care or surgical services. TEP members also suggested that such a focus might, in some cases, tilt selection away from use of outcome measures. Finally, there may not be an objective way to determine which measures meet the criterion of “broadly applicable” (or a way to otherwise reach consensus on what it means to be broadly applicable).

Rather than using a cross-cutting measures criterion, a more relevant criterion to the "borrowing strength" approach would have been groups of measures that have a common causal pathway. What makes borrowing strength "work" is the existence of underlying structural similarities in that causal pathway across time, peer providers, or related process and outcome measures. Those structural similarities are also what makes the borrowing strength approach either actionable (if those elements of structure are loosely under the provider’s or system’s control) or illuminating of an unintended consequence (if not).

Reporting Challenges: Are there additional reporting challenges that should be considered in future rural health measurement work? If so, please describe.

The comment "pooling data over several years for one provider would affect the ability to track improvement over time due to lag, which might pose a challenge for pay-for-performance programs intended to serve this purpose" (p. 11) again seems to miss the utility of the "borrowing strength" approach, which does not in fact require combining data over multiple years (or across peer providers) but rather leverages the persistent statistical relationship across years.

Gaps and Future Considerations: Are there additional gaps that should be considered in future rural health measurement work? If so, please describe.

The report does not really address the infrastructure requirements for implementing the "borrowing strength" approach which may have informed the selection of measures for testing. For example, the ability to establish benchmarks or thresholds, observed statistical correlations or persistence, and the ability to estimate correlated signal variances.

Stratis Health (commenter: Karla Weng)

What general comments do you have on the recommendations report?

Stratis Health is a non-profit organization whose mission is to lead collaboration and innovation in health care quality and safety. We have a long history of working closely with Critical Access Hospitals (CAHs) and other rural health care organizations and clinicians, with a focus on supporting quality reporting and improvement.

We applaud the ongoing work by NQF to address rural-relevant and low-case volume measurement, but strongly encourage additional support for development of rural sensitive measures to allow CAHs and other small rural hospitals to demonstrate the quality of care they provide, and to continue to participate in improvement and payment programs which lead to higher quality and lower cost care for Americans living in rural places.

Thank you for the opportunity to submit comments, and in particular, to help assure that patients living in rural places continue to receive the highest quality care possible in our nation's rural hospitals.

Measure Recommendations: Do you agree that the 15 measures recommended for statistical testing are relevant for rural populations and are susceptible to low case-volume reporting challenges? If not, please provide feedback.

We encourage consideration of availability of services in rural hospitals before applying testing of statistical methods for low-volume. For example, a limited number of CAH have labor and delivery services available. For those that do offer that service, is low case volume still an issue for the PC-02 measure, or is there just a limited number of rural hospitals providing that service? The same question would apply for the THA/TKA readmission measure and OP-32.
The potential inclusion of the Sepsis measure in the MBQIP program is exploratory at this time, and no decision has been made by the Federal Office of Rural Health Policy to include it as a core measure for that program.
Per the Risk Standardized, All Condition Readmission measure. The majority of CAHs currently meet the minimum threshold for calculation of Hospital-Wide Readmission measure, and we encourage review of that information prior to inclusion of the measure as a priority for statistical testing. CMS has indicated that they will be shifting to utilization of a hybrid measure for calculation of Hospital-Wide Readmissions, and will no longer be utilizing the claims-only measure starting in 2023. If testing is pursued on this measure, we'd encourage it be done on the hybrid version (voluntary reporting for the hybrid version begins in 2021).
We found the legends for the charts on page 10 hard to read and interpret, we'd encourage you to use larger color boxes in the legend or use a different format for that information.

Gaps and Future Considerations: Are there additional gaps that should be considered in future rural health measurement work? If so, please describe.

There continues to be a significant need for measure adaptation and measure development to help address critical areas of quality and safety for rural health care such as priorities identified in previous NQF Rural reports including: access and timeliness of care, care transitions, substance use, cost, population health, advance care directives, end-of life care, and patient outcomes.

Reporting Challenges: Are there additional reporting challenges that should be considered in future rural health measurement work? If so, please describe.

We encourage exploration of the potential to expand the universe of claims available for calculation of the claims-based measures. A limitation of the current CMS measure reporting system is that measures such as hospital readmissions are only calculated using FFS Medicare claims. Expansion to inclusion of Medicare Advantage data, or ideally to an all-payer claims database (where available) would likely increase the utility of existing claims-based measures for rural and low-volume facilities.
We agree with the potential challenges identified regarding greater use of eCQMs, but would also highlight that an additional issue is that there is limited rural relevancy to the currently available eCQMs for hospital reporting through part of the CMS Inpatient Quality Reporting (IQR) and Promoting Interoperability Programs. Only 2 of the 8 eCQMs measures currently available (ED-2 and VTE-1) are relevant to most CAHs and hospitals are required to report on 4 measures (https://www.qualityreportingcenter.com/globalassets/iqr_resources/ecqm-resources-for-iqr/cy-2020/ecqm_cy-2020-available-ecqms-table_vfinal508.pdf). There are good opportunities for reduction of measurement burden through expanded use of eCQMs, but the challenges in availability of quality data for small rural hospitals will remain the same unless the eCQMs selected for inclusion in those programs are rural-relevant.

National Organization of State Offices of Rural Health (NOSORH) (commenter: Teryl Eisinger)

The National Quality Forum (NQF) recently released the latest draft of a report on ***Rural-Relevant Quality Measures for Testing of Statistical Approaches to Address Low Case-Volume***. In this communication National Organization of State Offices of Rural Health (NOSORH) is providing comments on that report. NOSORH had submitted comments on a previous iteration of NQF's rural-relevant quality measures – a copy of those comments is attached.

NOSORH was established in 1995 to assist State Offices of Rural Health (SORH)s in their efforts to improve access to, and the quality of, health care for nearly 57 million rural Americans. All 50 states have a SORH, and each SORH helps their state's rural communities to build effective health care delivery systems. SORHs work with the rural health care system nationwide, with a particular emphasis on rural hospitals, including smaller rural hospitals and Critical Access Hospitals (CAHs). NOSORH and its member SORHs have a long history of working with quality monitoring and quality improvement in these facilities, and are integrally involved in the operations of the Medicare Beneficiary Quality Improvement Project (MBQIP). MBQIP is the primary resource for quality monitoring for the more than 1,300 CAHs nationwide.

NOSORH is strongly supportive of several of the measure selection criteria used by NQF in the development of its report.

- **<u>Cross-cutting measures</u>**: applicable to the broadest range of patients and services
- **<u>Measures relevant to low volume service environments</u>**: applicable for measuring quality with relatively small numbers of reportable incidents.
- **<u>Measures that address transitions in care</u>**: assessing the broader context of care continuity across multiple environments, including a patient's home.

NOSORH believes that a quality measurement scheme developed with these selection criteria will be a major step forward in understanding quality trends in rural health services.

NOSORH believes that the final set of recommendations in the NQF would improve rural health system monitoring, but that additional changes would be beneficial. NOSORH has identified multiple measures in the final set included in the report that would be ***difficult to apply in CAH and other low-volume rural***

hospitals. The NQF report itself includes comments from the NQF Workgroup and the public that indicate that ***measures might not be feasible in these hospitals due to low case-volumes in rural areas***. The NQF report further notes Workgroup comments questioning the ***feasibility of data collection in CAHs*** for several measures, due to cost and reporting rules that are difficult for rural providers to meet.

Given these challenges, NOSORH recommends that additional changes be made to the report. NOSORH believes that ***no single measurement set should be created for all providers***. A 'one size fits all' approach has been taken by some CMS provider evaluation schemes - in particular the Hospital Star Rating system. Under this scheme hospitals are assessed on 57 separate reported measures grouped in 7 Domains. Few hospitals can acceptably report on all 57 measures. This has led to different hospitals being assessed on completely different numbers of measures and different mixes of measures. This severely limits the usefulness of the ultimate comparisons.

NOSORH has conducted a study of the Hospital Star Rating system and identified major problems with its treatment of rural providers. NOSORH analysis has indicated that, in the most recent iteration of the data reporting, ***fewer than half of all CAHs were able to report on enough measures to be rated***. In addition, among rated hospitals, fewer than 10% of all CAHs were rated on the important Patient Safety domain, compared to more than 90% of all acute care hospitals. This is very problematic. NOSORH believes that the current measurement recommendations included in the NQF report could lead to a quality measurement scheme that repeats these problems – excluding many CAHS from monitoring.

To address these problems, NOSORH recommends that NQF identify a core set of cross-cutting measures for all providers and also identify ***separate supplemental sets of measures that are specific to different provider categories***. Separate inpatient measure sets can be established for CAHs, general acute care facilities and for specialty care facilities. Separate outpatient category measure sets can be established for primary care providers and key categories of specialists/subspecialists.

NOSORH believes that ***process*** measures are more appropriate for lower volume facilities, such as CAHs. NOSORH feels that the MBQIP measure set, which includes several process measures, has shown its worth as quality index for low-volume hospitals. NOSORH recommends the development of a CAH-specific set of quality measures based upon MBQIP.

NOSORH recognizes the challenge faced by NQF in the development of quality measures relevant for low-volume rural providers. NOSORH commends NQF for its efforts and hopes that these comments can help support some additional improvements.