NATIONAL QUALITY FORUM

Meeting Summary

Primary Care and Chronic Illness Standing Committee— Measure Evaluation In-Person and Post-Evaluation Web Meetings

The National Quality Forum (NQF) convened the Primary Care and Chronic Illness Standing Committee for an in-person meeting on June 26, 2019 at the NQF offices in Washington, DC and two post-measure evaluation web meetings on July 1 and July 8. The Committee reviewed a total of 10 measures during these three evaluation sessions.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the in-person meeting. NQF staff reviewed the meeting objectives. Committee members each introduced themselves and disclosed any conflicts of interest.

Topic Area Introduction and Overview of Evaluation Process

NQF staff provided an overview of the topic area and the current NQF portfolio of endorsed measures. There are currently 50 in the Primary Care and Chronic Illness portfolio. Additionally, NQF reviewed the Consensus Development Process (CDP) and the measure evaluation criteria.

Measure Evaluation

During the meeting, the Primary Care and Chronic Illness Standing Committee evaluated 10 measures for endorsement consideration. A summary of the Committee deliberations will be compiled and provided in the draft technical report. NQF will post the draft technical report on August 1, 2019 for public comment on the NQF website. The draft technical report will be posted for 30 calendar days.

Measure Evaluation Criteria Rating Key: H – High; M – Medium; L – Low; I – Insufficient

2522 Rheumatoid Arthritis: Tuberculosis Screening (American College of Rheumatology)

Measure Steward/Developer Representatives at the Meeting

Lisa Suter, Rachel Myslinski, and Tracy Johansson

Standing Committee Votes

- Evidence: H-0; M-18; L-0; I-0
- Performance Gap: H-7; M-12; L-0; I-0
- Reliability: H-3; M-15; L-2; I-0
- Validity: H-1; M-18; L-0; I-0
- Feasibility: H-8; M-11; L-1; I-0
- Use: Pass-20; No Pass-0
- Usability: H-6; M-14; L-0; I-0

Standing Committee Recommendation for Endorsement: Yes-19; No-1

Committee members discussed the role of registries and registry-based data in quality measurement. The Committee noted there is evidence that screening prevents and results in treatment of tuberculosis, and after some clarifying discussion on the NQF evidence algorithm, the measure passed the evidence criteria. Committee members noted that while performance is improving, there remains a gap of about 15 percent. This led Committee members to question whether there was an actual gap in care or just problems with capturing the data out of EHRs. The developer explained that they have done rigorous validation of the data elements, and after confirming there are actual gaps in screening, the Committee passed the measure on gap. The Committee discussed the types of testing included in the measure specifications: It noted challenges with reading skin tests and requested that the developer provide more guidance to ensure consistency, flagging these challenges as potential causes of both over- and under-treatment. The developer noted that they anticipate tuberculosis skin testing rates will continue to decline in favor of blood tests.

The Committee noted that a particular medication should not be included in the measure (Rituximab) because it does not cause the same problems, and the developer agreed to remove it. The developer provided additional data on testing for the individual provider level after the original submission deadline. The Committee requested, and the developer agreed, that the measure require a minimum threshold of 10 cases for accountability purposes to ensure the measure is fully reliable. The Committee did not consider the measure to have strong reliability below 10 patients, but there will be no minimum threshold for quality improvement purposes. With the two changes specified, and in light of the additional information submitted, the Committee agreed that the measure met NQF's reliability and validity criteria. Committee members noted that the measure's data elements are pulled from structured fields. This fact and the trend toward assay testing (and away from skin testing) further increase the feasibility. Since the measure is currently in use, the Committee had no major concerns on the use or usability. In response to questions, the developer explained that patients had been included in the development team for the measure. The Standing Committee recommended the measure for continued NQF endorsement.

2523 Rheumatoid Arthritis: Assessment of Disease Activity (American College of Rheumatology)

Measure Steward/Developer Representatives at the Meeting

Lisa Suter, Rachel Myslinski, and Tracy Johansson

Standing Committee Votes

- Evidence: H-5; M-15; L-0; I-0
- Performance Gap: H-5; M-14; L-1; I-0
- Reliability: H-8; M-11; L-1; I-0
- Validity: H-2; M-15; L-2; I-1
- Feasibility: H-0; M-10; L-10; I-0
- Use: Pass-15; No Pass-5

- Usability: H-1; M-14; L-5; I-0

Standing Committee Recommendation for Endorsement: Yes-15; No-5

Committee members requested clarification on how visits are counted, noting that patients could see their general practitioner and discuss their rheumatoid arthritis (therefore coding it as discussed) but that provider wouldn't be screening for disease activity. The developer explained that only providers in the registry are participating in the measure, participation is voluntary, and that they have set a lower bar for capturing disease activity (at 50 percent of visits) because there are encounters when a provider would appropriately not be capturing disease activity. Committee members noted, and the developer agreed, that there are potential scalability issues to implementing the measure outside the registry, but that not all patients with rheumatoid arthritis are being treated by rheumatologists. Committee members suggested minor adjustments to the coding to assist with this. The developer agreed to consider these comments as the measure is expanded. The measure is based on the guidelines, which are themselves based on systematic reviews, so the Committee agreed that the measure met the evidence criteria. The Committee agreed there is a gap in care, and the measure passed performance gap.

Similar to the previous measure (2522), the developer provided additional testing information for the individual provider level of analysis, and the Committee noted that this measure achieved better reliability scores than 2522. The measure passed reliability. After some discussion of the process of calculating the measure and what counts as a disease activity measure, the Committee agreed the measure is valid. Committee members noted feasibility challenges, stating that in practice, providers are doing this with paper and check boxes and waiting for the test results to come back, and later inputting the data, and that EHRs have not yet caught up with practice. Committee members also noted that having six different tools is meant to make the measure more feasible, but since only some of the tools require lab work and some do not, there may be differing results. The developer noted there is not a best-in-class disease activity assessment tool and that different providers prefer different tools. They further noted it is burdensome for providers to collect needed data but that it is very important to treat the disease properly, and that the ACR is continuing to work to improve the feasibility across more EHRs.

Committee members noted that measures can help drive the field as well, as a measure may make it more likely that EHR vendors will include the appropriate structured data fields. Committee members noted that the assessment of disease activity itself is incredibly important and is feasible, but that the challenges are with getting the data into the EHR properly, and that could lead to potential negative impacts for providers whose EHRs can't manage, therefore potentially leading to these providers refusing to take patients. The developer noted they have just started working with Epic, which greatly increased the number of providers who can easily use the measure. The Committee agreed that the measure was feasible for providers using the RISE database, which only includes about 30 percent of practicing rheumatologists, but that 95 percent of rheumatologists are ACR members and eligible to use the RISE registry. Ultimately, the Committee did not reach consensus on whether the measure is feasible (50 percent rated moderate and 50 percent rated low), but feasibility is not a must-pass criterion, so consideration of the measure continued. The measure is currently in use in the RISE registry and will be reported on in MIPS in 2020, and feedback is given to participating providers. The Committee

agreed that the measure met both the use and usability criteria. The Standing Committee recommended the measure for NQF endorsement.

2525 Rheumatoid Arthritis: Disease Modifying Anti-Rheumatic Drug (DMARD) Therapy (American College of Rheumatology)

Measure Steward/Developer Representatives at the Meeting

Lisa Suter, Rachel Myslinski, and Tracy Johansson

Standing Committee Votes

- Evidence: H-0; M-20; L-0; I-0
- Performance Gap: H-0; M-20; L-0; I-0
- Reliability: H-7; M-12; L-1; I-0
- Validity: H-3; M-16; L-1; I-0
- Feasibility: H-2; M-18; L-0; I-0
- Use: Pass-20; No Pass-0
- Usability: H-2; M-17; L-1; I-0

Standing Committee Recommendation for Endorsement: Yes-20; No-0

The measure is based on guidelines, which were developed based on evidence from systematic reviews; the Committee had no concerns and agreed it met the evidence criterion. There is a limited gap, with over 90 percent adherence; the Committee raised the question of whether the measure might be topped out or nearly topped out. The developer noted that new practices are using the measure all the time, and that it is useful to help them understand their performance; they see rapid improvement when the measure is implemented. They also noted the need to understand the role of disparities in the measure performance. The Committee noted that the measure looks at providers' actions of prescribing, but that does not necessarily follow through to whether a prescription was filled and used, so the gap in care received is likely larger. The Committee discussed various exclusion criteria; the developer clarified patient refusal is not included due to concerns about gaming and the role of shared decision making which should ensure patients are selecting drugs that work for them. Ultimately the Committee agreed there was likely a larger gap in care than current performance suggests and the measure passed gap.

The Committee discussed the scalability question again, similar to measure 2523. The Committee agreed that the measure performed well on reliability testing and met the reliability criteria. During the validity discussion, the developer clarified that the list of drugs is updated annually, and the Committee agreed the measure is valid. The Committee noted data for this measure are available in discrete data fields and had no concerns about feasibility. The measure is currently only in use in the RISE registry, and it is similar to the previous two measures; the Committee voted to pass on both use and usability. The Committee briefly discussed a public comment received on the measure during the pre-meeting commenting period. The Committee then voted to recommend the measure for NQF endorsement.

0541 Proportion of Days Covered (Pharmacy Quality Alliance)

Measure Steward/Developer Representatives at the Meeting

Lynn Pezzullo, Irene Nsiah, Lisa Hines

Standing Committee Votes

- Evidence: H-0; M-14; L-6; I-0
- Performance Gap: H-3; M-16; L-0; I-0
- Reliability: Yes-19, No-0
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel. The measure was ranked as Moderate.
- Validity: Yes-18, No-2
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel. The measure was ranked as Moderate.
- Feasibility: H-2; M-16; L-2; I-0
- Use: Pass-14; No Pass-6
- Usability: H-3; M-9; L-7; I-0

Standing Committee Recommendation for Endorsement: Yes-16; No-4

The Committee noted that these are known measures with broad national adoption. Committee discussion was prefaced with the note that the data source for this measure is electronic pharmacy claims, a source with significantly higher precision than conventional medical claims. Nonetheless, pharmacy data do not contain the breadth of information that is found either in the EHR, or what may be present in traditional medical claims. Committee members questioned the measure developer on the logic model that connects pharmacy claims with positive patient outcomes, specifically voicing the concern that pharmacy claims might not be an adequate proxy for patient medication taking behavior. The lead discussant pointed to evidence provided by the developer that adherence measures using the proportion of days covered (PDC) methodology have been repeatedly demonstrated to serve as a strong proxy for medication adherence, with clear connections to positive patient medical outcomes and decreased cost of care at the population level.

The Committee asked the developer what occurs when patients experience side effects or significant adverse drug events (ADE) associated with medication use. The developer responded that the measures demonstrate a robust resilience to these effects, for two reasons. First, the measure specifications stipulate that a patient must have two fills of a medication in order to appear in the denominator, with most patients discontinuing therapy because of side effects or ADEs on the first fill of a given medication. Second, assuming an equal distribution of these types of events across populations, health plans would theoretically be affected by such discontinuations at the same rate, and hence still have accurate comparability using these three PDC rates. The Committee was satisfied with the evidence and performance gap for the measure. This measure was deemed complex due to risk adjustment and was evaluated by the NQF

Scientific Methods Panel. The measure developer submitted a first-of-its-kind risk-adjustment model for a process measure.

The Committee had limited discussion on the reliability of the measure and elected to uphold the Methods Panel reliability ranking. The validity discussion centered on risk adjustment, stratification, and correlation with other measures. The developer noted that the thresholds for performance indicate that validity correlations were moderate by conventional evaluation standards for Pearson correlation coefficients between quality measures. The Committee upheld Methods Panel validity ranking.

During the discussion of feasibility, the Committee introduced concerns that prescriptions that are not caught by claims will not be captured in the data. This could result in consequences for health plans as well as downstream consequences for providers and pharmacists accountable for patients who appear to be nonadherent to their medications, but simply have not been captured by claims data. The developer noted that they are currently in the process of specifying measures that draw exclusively on pharmacy dispensing data, which would alleviate this concern. In the discussion on use and usability, it was noted that these measures are currently in use. The Committee noted hospice and ESRD exclusions, but after some discussion determined these exclusions to be appropriate. When the Committee asked how plans can improve performance, the developer noted how research has demonstrated that interventions such as medication therapy management, performance reports, dashboards, outreach to patients, among other approaches, return positive improvements in population level adherence rates. The Committee also noted that rates in Medicare PDC performance have continually improved year-over-year, and that Medicare has acknowledged significant financial benefits associated with increased medication adherence across Medicare beneficiaries. The Standing Committee recommended the measure for continued endorsement.

3059e One-Time Screening for Hepatitis C Virus (HCV) for Patients at Risk (PCPI Foundation)

This is an electronic clinical quality measure (eCQM)

Measure Steward/Developer Representatives at the Meeting

Jamie Lehner, Eduardo Segovia, Nadene Chambers, Samantha Tierney, and John Wong

Standing Committee Votes

- Evidence: H-10; M-9; L-0; I-0
- Performance Gap: H-11; M-8; L-0; I-0
- Reliability: H-1; M-13; L-1; I-5
- Validity: H-0; M-15; L-5; I-0
- Feasibility: H-3; M-14; L-3; I-0
- Use: Pass-17; No Pass-2
- Usability: H-1; M-16; L-1; I-1

Standing Committee Recommendation for Endorsement: Yes-16; No-3

This is a new eMeasure going through NQF full initial endorsement; the measure was previously in Approved for Trial Use status. The Committee reviewed the evidence and performance gap and commented that there are very few measures in the portfolio of NQF endorsed measures that address hepatitis C screening and treatment, an important area of clinical concern. The Committee was satisfied with the developer's demonstration of evidence and performance gap. In the reliability discussion, the Committee expressed some concern around the lack of clarity for the care settings contained in the developer's testing sample. The specifications for the measure outlined care settings where the measure could be deployed, with no indication in the testing if those settings were indeed present in the data. The developer explained that their data were provided to them by CMS but afforded them limited ability to identify provider types. The Committee requested that the developer secure data that allow them to test measures to specifications for future submissions. In the discussion related to validity, the Standing Committee noted that as this is a new measure, the developer was only required to submit face validity testing. However, the Committee had fairly extensive discussion surrounding the exceptions, including the concern that the measure doesn't address the stigmas associated with intravenous drug use and the potential penalization of providers for things that are outside of the provider's control, such as whether or not patients receive a blood test recommended by the provider.

The feasibility discussion also connected with some themes in the exclusion criteria, namely that patients potentially may have a strong disinclination to having intravenous drug use documented within a structured data field, and many providers do not include coding to that effect due to the stigmas associated with intravenous drug use. It was noted during the discussion of use that the measure was on the Measures Under Consideration List for potential inclusion in the Merit-based Incentive Payment System. As this is a new measure, use is not a must-pass criterion. The conversation about usability revealed a concern by the Committee for potential over-screening if the documentation isn't there and noted the difficulty in obtaining certain data elements, such as blood transfusion before 1992 and history of injection drug use. Potential harms of stigma or anxiety waiting for results were considered to not outweigh the benefits of the measure. The Standing Committee recommended the measure for NQF endorsement.

3060e Annual Hepatitis C Virus (HCV) Screening for Patients who are Active Injection Drug Users (PCPI Foundation)

This is an electronic clinical quality measure (eCQM)

Measure Steward/Developer Representatives at the Meeting

Jamie Lehner, Eduardo Segovia, Nadene Chambers, Samantha Tierney, John Wong, Beth Foster

Standing Committee Votes

- Evidence: H-4; M-14; L-0; I-1
- Performance Gap: H-11; M-7; L-0; I-1
- Reliability: H-0; M-8; L-9; I-2
- Validity: H-0; M-12; L-7; I-0
- Feasibility: H-0; M-4; L-15; I-0

- Use: Pass-12; No Pass-6
- Usability: H-0; M-8; L-10; I-0

Standing Committee Recommendation for Endorsement: Yes-N/A; No-N/A

This is a new eMeasure going through NQF full initial endorsement; the measure was previously in Approved for Trial Use status. The Committee noted that the evidence for this measure was similar to that for the previous measure 3059e discussed in that it is supported by guidelines, but they noted concern about the grade of the evidence. The Committee was also concerned that there is a proliferation of measures, and not a clear need for a metric on every desirable outcome. While the developer didn't present formalized gap analysis using primary data, they did summarize articles that noted an independent disparity gap, with Caucasians and women being less likely to be tested. The Committee noted a gap based on the number of people that probably should be tested, according to the data presented by the developer.

The Committee cited a number of concerns related to reliability. First, the occurrence rate is very small, with only 30 events in the first data set, and 22,000 events from 4.8 million visits in the second. This implies that there may be an issue with who is self-reporting as an active intravenous drug user, compounded by the potential for self-reporters to be the same population that would be willing to get tested. The Committee also noted that injection drug users do not typically schedule care, so the exclusion of emergency departments as a care setting is also a potential confounder. The developer noted that the larger data set excluded all providers who had fewer than 10 events due to potential reidentification issues in the deidentified data. This indicates that the measure was not tested to specifications due to misalignment of exclusion criteria in the testing and specifications. Due to these concerns, the Committee was not able to achieve consensus on the vote for reliability.

Similar to the previous measure 3059e, the developer used face validity testing to fulfill the validity requirement. It was noted that there were a high number of exclusions in this measure, which was viewed as a threat to validity. In the discussion of the feasibility of the measure, Committee members noted that the measure should be a byproduct of routine patient care. There was some concern that the distinction between active and inactive drug use may not lend itself to good measurement. The developer noted the importance of this distinction, and also added that this is a yearly evaluation for patients who remain at continued risk, which is different from the one-time screening in the previous measure 3059e. The measure did not pass feasibility, but it is not a must-pass criterion. The Committee noted that because this is a new measure with potential for inclusion in accountability programs, it would still be appropriate to pass for use even though it is yet to be adopted. In the discussion of usability, the Committee appreciated that there were no harms identified in the measure, but added that the identification of the population that needs screening remains a challenge.

The Standing Committee did not vote on the recommendation for endorsement because the Committee did not reach consensus on reliability—a must-pass criterion. The Committee will re-vote on the measure on the post-comment web meeting on September 24, 2019.

0086 Primary Open-Angle Glaucoma (POAG): Optic Nerve Evaluation (PCPI Foundation)

Measure Steward/Developer Representatives at the Meeting

Dr. Scott Perek, Jamie Lehner, Eduardo Segovia, Elvia Chavarria, Samantha Tierney, and Nadine Chambers

Standing Committee Votes

- Evidence: H-9; M-8; L-0; I-0
- Performance Gap: H-4; M-14; L-0; I-0
- Reliability: H-2; M-15; L-1; I-0
- Validity: H-0; M-11; L-7; I-0
- Feasibility: H-0; M-17; L-0; I-0
- Use: Pass-18; No Pass-0
- Usability: H-2; M-15; L-0; I-0

Standing Committee Recommendation for Endorsement: Yes-17; No-1

The Committee agreed that this process measure is important to measure the percentage of patients aged 18 years and older with a diagnosis of primary open-angle glaucoma (POAG) who have an optic nerve head evaluation. This measure is reported through claims and registry, whereas 0086e is reported through the electronic health records. The Committee agreed that the evidence continues to remain strong and a performance gap continues to exist and did not have further discussion.

The Committee had some discussion on the reliability and validity testing of the measure. Since testing on the measure was not at the clinician: individual level of analysis, this measure would be evaluated by the Committee at the clinician: group/practice level of analysis only. The developer noted that they were unable to parse out their data at the clinician: individual level of analysis for this measure. One Committee member noted that ICD 10 coding of this measure included normal-tension and low-tension glaucoma in the definition of primary open-angle glaucoma. The developer noted they will share that coding feedback with their technical expert panel during their annual update. The Committee noted that the empirical validity results using Pearson's correlation coefficients to compare performance of 0086 with PQRS #117 *Diabetes: Eye Exam* were moderate at the registry level (0.57), but weak at the claims level (0.22).

The Committee had no further discussion or concerns on the feasibility and use. In regard to the usability criterion, a few Committee members expressed support that this measure will encourage optic nerve evaluations being performed and hopefully in the future encourage measures that address optic nerve evaluation. The Committee noted that there is one related measure, 0563 *Primary Open-Angle Glaucoma: Reduction of Intraocular Pressure* by 15 percent or *Documentation of a Plan of Care*; however, 0563 has a different measure focus than 0086. One Committee member noted that 0563 and 0086 differ with respect to including patients who have normal or low-tension glaucoma and would like to see harmonization in the target populations of the two measures. A few Committee members suggested that the developer consider whether the appropriate measure title and target population is primary open-angle glaucoma or the

general glaucoma population. The developer will share that feedback with their technical expert panel during their annual update. The Standing Committee recommended the measure for continued endorsement.

0086e Primary Open-Angle Glaucoma (POAG): Optic Nerve Evaluation (PCPI Foundation)

This is an electronic clinical quality measure (eCQM)

Measure Steward/Developer Representatives at the Meeting

John Thompson, Jamie Lehner, and Samantha Tierney

Standing Committee Votes

- Evidence: H-9; M-7; L-0; I-0
- Performance Gap: H-0; M-15; L-1; I-0
- Reliability: H-0; M-12; L-4; I-0
- Validity: H-0; M-7; L-8; I-1
- Feasibility: H-1; M-15; L-0; I-0
- Use: Pass-16; No Pass-0
- Usability: H-1; M-15; L-0; I-0

Standing Committee Recommendation for Endorsement: Yes-N/A; No-N/A

This process measure is the eMeasure version of 0086 which measures the percentage of patients aged 18 years and older with a diagnosis of primary open-angle glaucoma (POAG) who have an optic nerve head evaluation. The Committee agreed to pull the votes on evidence from 0086 as it is identical information. The Committee agreed that a performance gap continues to exist and did not have further discussion on the criterion.

The Committee did not reach consensus on the validity of the measure. In regard to validity of the specification, the Committee members raised again with the developer for consideration the appropriate coding of this measure which includes normal-tension and low-tension glaucoma; and also if the appropriate measure title and target population is primary open-angle glaucoma or the general glaucoma population. The developer noted again their plan to share that feedback with their technical expert panel during their annual update process. One Committee member discussed if the appropriate sample of specialists is reporting on the measure. The developer noted that specialists are able to choose which measure they report on so would generally report on measures for which they have expertise. The Committee noted that the empirical validity result using Pearson's correlation coefficients to compare performance of 0086e with PQRS #117 *Diabetes: Eye Exam* was weak at the EHR level (0.36); however, one Committee member felt the correlation coefficients would be stronger except that the providers reporting the two measures may be taking care of different types of patients. One Committee member had a concern that the measure is not risk adjusted for potential social determinants of health and/or age. However, other Committee members did not feel this measure needs risk adjustment. The Committee had no further discussions or concerns on the feasibility, use, and usability of the measure.

The Standing Committee did not vote on the recommendation for endorsement at the meeting because the Committee did not reach consensus on validity—a must-pass criterion. The Committee will re-vote on the measure on the post-comment web meeting on September 24, 2019.

0089 Diabetic Retinopathy: Communication with the Physician Managing Ongoing Diabetes Care (PCPI Foundation)

Measure Steward/Developer Representatives at the Meeting

John Thompson, Jamie Lehner, Elvia Chavarria, Eduardo Segovia, and Samantha Tierney

Standing Committee Votes

- Evidence: H-0; M-1; L-2; I-13
- Performance Gap: H-0; M-15; L-0; I-0
- Evidence Exception: Yes-7; No-8
- Reliability: H-1; M-7; L-6; I-1
- Validity: H-0; M-5; L-11; I-0 (Did not pass at validity criterion. Voting stopped.)
- Feasibility: H-N/A; M-N/A; L-N/A; I-N/A
- Use: Pass- N/A; No Pass- N/A
- Usability: H- N/A; M- N/A; L- N/A; I- N/A

Standing Committee Recommendation for Endorsement: Yes-N/A; No-N/A

The Standing Committee did not vote on the recommendation for endorsement because the measure did not pass the validity criterion—a must-pass criterion. In addition, the measure did not reach consensus on the evidence and reliability criteria. This process measure measures the percentage of patients aged 18 years and older with a diagnosis of diabetic retinopathy who had a dilated macular or fundus exam performed with documented communication to the physician who manages the ongoing care of the patient with diabetes mellitus regarding the findings of the macular or fundus exam. This measure is reported through claims and registry, whereas 0089e is reported through the electronic health records.

More than 60 percent of the Committee members voted Insufficient on evidence. Committee members noted that there is no evidence indicating communication between physicians performing the dilated macular or fundus exam and those treating the diabetes will lead to improved health outcomes for the patient. The Committee was able to vote on evidence with exception; however, the Committee did not reach consensus on evidence with exception. Some Committee members did not see value in a performance measure addressing this measure focus, in addition to their concern about the evidence. One Committee member also expressed that quality of care is mandatory; however, if a quality measure does not meet applicable standards, then the benefit of measurement may not justify the reporting burden. However, some Committee members had a different opinion: They did see value in the measure as a potential driver of improved outcomes. The developer noted that care coordination measures are an important gap in the measurement field.

The Committee agreed that a performance gap continues to exist and did not have further discussion on the criterion.

The Committee did not reach consensus on the reliability of the measure. Since testing on the measure was not at the clinician: individual level of analysis, this measure was evaluated at the clinician: group/practice level of analysis only. In addition, the developer specified the measure for outpatient, post-acute care and domiciliary settings, but these analyses were not conducted separately. A few Committee members with an ophthalmology background noted a very small percentage of ophthalmologists reporting on this measure would be from the domiciliary setting and would be predominantly reporting at the outpatient setting.

The Committee did not pass the measure on validity. The Committee noted that the empirical validity results using Pearson's correlation coefficients to compare performance of 0089 with PQRS #117 *Diabetes: Eye Exam* were weak at the claims and registry levels (0.11 and 0.16). However, one Committee member felt the correlation coefficients would be stronger except that the providers reporting the two measures may be taking care of different types of patients. Discussion and voting stopped at the validity criterion, as it is a must-pass criterion.

0089e Diabetic Retinopathy: Communication with the Physician Managing Ongoing Diabetes Care (PCPI Foundation)

This is an electronic clinical quality measure (eCQM)

Measure Steward/Developer Representatives at the Meeting

Samantha Tierney and Yvette Apura

Standing Committee Votes

- Evidence: H-0; M-3; L-3; I-8
- Performance Gap: H-3; M-10; L-1; I-0
- Evidence Exception: Yes-8; No-6
- Reliability: H-1; M-7; L-4; I-2
- Validity: H-0; M-4; L-9; I-1
- Feasibility: H-1; M-12; L-1; I-0
- Use: Pass-13; No Pass-1
- Usability: H-1; M-8; L-4; I-1

Standing Committee Recommendation for Endorsement: Yes-5; No-9

The Committee did not have quorum for voting on the measure at the July 8 post-meeting call and submitted their votes via SurveyMonkey afterwards. The measure did not pass the evidence and validity criteria—must-pass criteria. In addition, the measure did not reach consensus on the reliability criterion. This process measure measures the percentage of patients aged 18 years and older with a diagnosis of diabetic retinopathy who had a dilated macular or fundus exam performed with documented communication to the physician who manages the ongoing care of the patient with diabetes mellitus regarding the findings of the macular or fundus exam. This

measure is reported through electronic health records, whereas 0089 is reported through claims and registry.

The Committee had no additional discussion on evidence or gap of this measure. The evidence was thoroughly discussed previously on measure 0089, which has identical evidence information.

The Committee noted that the empirical validity result using Pearson's correlation coefficients to compare performance of 0089 with PQRS #117 *Diabetes: Eye Exam* was weak at the EHR level (0.08). There was a moderate correlation (0.59) with the measure, *Diabetic Retinopathy: Documentation of Presence or Absence of Macular Edema and Level of Severity of Retinopathy*. One Committee member asked the developer about the usability and feasibility of this eMeasure. The developer noted no issues thus far in the usability of the measure. The developer clarified for the Committee the type of communications qualifying for the measure. The Committee recapped previous Committee discussion on measure 0089 about the usability of the measure and whether the measure adds value and improves outcomes, which also applies to 0089e.

The Standing Committee did not recommend the measure for continued endorsement.

Public Comment

NQF held a pre-comment period that began on May 1 and ended on June 12, 2019. During that period, one comment from the public was received on measure 2525 related to the value set of the measure. No public or NQF member comments were provided during the measure evaluation meeting.

Next Steps

NQF will post the draft technical report on July 25, 2019 for public comment for 30 calendar days. The continuous public comment with member support will close on August 23, 2019. NQF will re-convene the Standing Committee for the post-comment web meeting on September 24, 2019.