NATIONAL QUALITY FORUM
Driving measurable health improvements together

Memo

November 17, 2020

To: Consensus Standards Approval Committee (CSAC)

From: Primary Care and Chronic Illness Project Team

Re: Primary Care and Chronic Illness Spring 2020[a]

CSAC Action Required

The CSAC will review recommendations from the Primary Care and Chronic Illness project at its November 17, 2020 meeting and vote on whether to uphold the recommendations from the Committee.

This memo includes a summary of the project, measure recommendations, themes identified and responses to the public and member comments and the results from the NQF member expression of support. The following documents accompany this memo:

1. **Primary Care and Chronic Illness Spring 2020 Draft Report**. The draft report has been updated to reflect the changes made following the Standing Committee's discussion of public and member comments. The complete draft report and supplemental materials are available on the project webpage.
2. **Comment Table**. Staff has identified themes within the comments received. This table lists 34 comments received during the post-meeting comment period and the NQF/Standing Committee responses.

Background

Primary care has a central role in improving the health of people and populations. Primary care practitioners manage the uniqueness and complexities of each patient. In this setting, the diagnosis and treatment of the patient is focused on the health of the entire patient and not a single disease. Chronic illnesses are long-lasting or persistent health conditions or diseases that patients and providers must manage on an ongoing basis. The Primary Care and Chronic Illness portfolio includes endocrine conditions; nonsurgical eyes, ears, nose, and throat conditions; infectious disease; musculoskeletal disorders; and pulmonary disease.

The 24-person Primary Care and Chronic Illness Standing Committee reviewed three measures: all three were recommended for endorsement, and one was not recommended for endorsement.

Draft Report

The Primary Care and Chronic Illness Spring 2020 draft report presents the results of the evaluation of three measures considered under the Consensus Development Process (CDP). All three were not

[a] This memo is funded by the Centers for Medicare and Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001.

recommended.

The measures were evaluated against the 2019 version of the measure evaluation criteria.

	Maintenance	New	Total
Measures under consideration	0	3	3
Measures not recommended for endorsement or trial use	0	3	3
Reasons for not recommending	Importance - 0 Scientific Acceptability - 0 Use - 0 Overall - 0 Competing Measure - 0	Importance - 1 Scientific Acceptability - 1 Use - 0 Overall - 1 Competing Measure - 0	

CSAC Action Required

Pursuant to the CDP, the CSAC is asked to discuss the three measures not recommended for endorsement.

Measures Not Recommended for Endorsement

(See Appendix B for the Committee's votes and rationale)

- NQF 3569e Prediabetes: Screening for Abnormal Blood Glucose (AMA)
- NQF 3570e Intervention for Prediabetes (American Medical Association (AMA))
- NQF 3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes (AMA)

Comments and Their Disposition

NQF received 34 comments from 12 organizations (including 1 member organizations) and individuals pertaining to the draft report and to the measures under consideration.

A table of comments submitted during the comment period, with the responses to each comment and the actions taken by the Standing Committee and measure developers, is posted to the Primary Care and Chronic Illness project webpage.

Theme 1 - Alignment of measure exclusions

Several commenters noted that the three measures which were evaluated have different exclusions. Commenters also suggested adding an exclusion such as exclusion of patients who are older and/or have multiple comorbidities and limited life expectancy.

> **Committee Response:**
> Thank you for your comment. The Committee carefully reviewed the exclusions with the developer during the post comment call. The Committee agrees that for measure NQF 3569e there should be more constraints around the age criteria to align with USPSTF recommendations.

Theme 2 - Concerns with data capture

Several commenters raised concerns over the measures' feasibility noting that currently there is no easy way to capture some of the interventions and they are likely not well-documented in EHRs.

Committee Response:

Thank you for your comments. The Committee reviewed these comments as well as the developer's response and discussed this theme at length during the post comment meeting. The Committee has heard the measure developer's argument that the measures' Feasibility Scorecard issues within the accuracy domain were offset by the validity testing using the parallel forms methodology. In general, the Committee has indicated that the validity testing has adequately demonstrated that the data elements necessary to calculate the measure may be represented inside of the EHRs where the measure was tested. The Committee expressed other concerns associated with the validity of NQF 3569e and 3570e. The Committee did not address validity concerns with 3571e because the Committee did not pass the measure due to weaknesses on the evidence criterion.

Member Expression of Support

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support ('support' or 'do not support') for each measure submitted for endorsement consideration to inform the Committee's recommendations. No NQF members provided their expression of support or non-support.

Removal of NQF Endorsement

No measures previously endorsed by NQF have not been re-submitted, nor endorsement has been removed.

Appendix A: CSAC Checklist

The table below lists the key considerations to inform the CSAC's review of the measures submitted for endorsement consideration.

Key Consideration	Yes/No	Notes
Were there any process concerns raised during the CDP project? If so, briefly explain.	No	
Did the Standing Committee receive requests for reconsideration? If so, briefly explain.	Yes	The developer submitted a request for reconsideration of the Committee's decision on #3570e. The developer believes that the Committee conflated the feasibility and reliability criteria, were not clear on how to apply the feasibility criteria and provided inconsistent recommendations for the three measures under consideration.
Did the Standing Committee overturn any of the Scientific Methods Panel's ratings of Scientific Acceptability? If so, state the measure and why the measure was overturned.	No	
If a recommended measure is a related and/or competing measure, was a rationale provided for the Standing Committee's recommendation? If not, briefly explain.	No	
Were any measurement gap areas addressed? If so, identify the areas.	Yes	The committee noted that there are no measures for prediabetes in the NQF portfolio.
Are there additional concerns that require CSAC discussion? If so, briefly explain.	No	

Appendix B: Measures Not Recommended for Endorsement

The table below lists the Committee's vote and rationale for measures not recommended for endorsement.

Legend: H = High; M = Moderate; L = Low; I = Insufficient

Measure	Voting Results	Standing Committee Rationale
3569e Prediabetes: Screening for Abnormal Blood Glucose (AMA)	**Evidence** H-4; M-17; L-1; I-0 **Gap** H-5; M-17; L-0; I-0 **Reliability** H-1; M-16; L-5; I-0 **Validity** H-0; M-11; L-8; I-3 **Feasibility** H-0; M-5; L-14; I-1 **Usability and Use** *Use* Pass-17; No Pass-3 *Usability* H-0; M-18; L-1; I-1 **Post Comment Call Vote:** Validity: M-8; L-9; I-1	The Committee expressed concerns that this measure has no upper age limit, which represents a threat to the measure's clinical validity. The Committee noted that USPSTF guidelines specify an age range of 40-70 years associated with the focus of the measure. The Committee recommends that the developer align with the guidelines
3570e Intervention for Prediabetes (AMA)	**Evidence** H-0; M-16; L-2; I-2 **Gap** H-2; M-16; L-1; I-1 **Reliability** H-0; M-16; L-3; I-0 **Validity** H-0; M-13; L-3; I-3 **Feasibility** H-0; M-5; L-15; I-1 **Usability and Use** *Use* Pass-18; No Pass-0 *Usability* H-0; M-10; L-6; I-2 **Post Comment Call Vote:** Request for Reconsideration: Yes-2; No-14	The Committee raised concerns that the measure has too many feasibility challenges to warrant an endorsement recommendation. Specifically, the Committee noted that the measure requires clinicians to either prescribe metformin or refer the patient out, which was noted to be especially burdensome and to not represent the range of options that clinicians in primary care settings have at their disposal to address this challenge.

Measure	Voting Results	Standing Committee Rationale
3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes (AMA)	**Evidence** H-0; M-4; L-6; I-7 **Insufficient Evidence with Exception** Yes-10; No-7 **Gap** H-1; M-10; L-3; I-3 **Reliability** H-0; M-12; L-5; I-0 **Validity** H-0; M-9; L-7; I-1 **Feasibility** H-0; M-7; L-9; I-0 **Usability and Use** *Use* Pass-15; No Pass-2 *Usability* H-0; M-7; L-6; I-3 **Post Comment Call Vote:** Evidence: M- 7; L- 7; I-4 Exception to Evidence: Yes- 9; No-8	The Committee noted that there is a lack of evidence supporting the frequency of retesting.

Appendix C: NQF Member Expression of Support Results

No NQF members provided their expression of support for the measures under consideration.

Appendix D: Details of Measure Evaluation

Rating Scale: H=High; **M**=Moderate; **L**=Low; **I**=Insufficient; **NA**=Not Applicable

Measures Not Recommended

3569e Prediabetes: Screening for Abnormal Blood Glucose

Submission

Description: Percentage of patients aged 40 years and older with a BMI greater than or equal to 25 who are seen for at least two office visits or at least one preventive visit during the 12-month period who were screened for abnormal blood glucose at least once in the last 3 years

Numerator Statement: *Screening for abnormal blood glucose may include using a fasting plasma glucose, 2-h plasma glucose during a 75g oral glucose tolerance test, or A1C.

Denominator Statement: All patients aged 43 years and older with a BMI greater than or equal to 25 seen for at least two office visits or at least one preventive visit during the 12-month measurement period

Exclusions: Denominator Exclusions

"Patient is Pregnant at Encounter"

or "Patient Has Active Diabetes Diagnosis at Encounter"

or "Hospice During Measurement Period"

or "Palliative Care During Measurement Period"

or "Comfort Measures During Measurement Period"

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Electronic Health Records

Measure Steward: American Medical Association

STANDING COMMITTEE MEETING 06/25/2020

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-4; M-17; L-1; I-0**; 1b. Performance Gap: **H-5; M-17; L-0; I-0**

Rationale:

- Developer cites evidence found in guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA).
    - The focus of the recommendations is lifestyle change.
    - USPSTF incorporated this evidence into the updated recommendation regarding screening for abnormal glucose and type 2 diabetes.
    - The grade B recommendation states that physicians should screen individuals for abnormal glucose if they are between the ages of 40 and 70 and are overweight or obese, or younger if they have additional risk factors.
    - The ADA recommends screening patients who are overweight or obese with one risk factor, regardless of age. Additionally, those who have no risk factors should start screening at age 45.
    - Developer notes that the risk factors included in this measure bring together both the USPSTF and ADA risk factors.
    - Testing for prediabetes and risk for future diabetes in asymptomatic people should be considered in adults of any age who are overweight or obese (BMI ≥25kg/mg or ≥23kg/m2 in Asian Americans) and who have one or more additional risk factors for diabetes. (ADA, 2018) (B Recommendation)
    - Grade B recommendation means: “The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.”

3569e Prediabetes: Screening for Abnormal Blood Glucose

- The Committee questioned the fact that the measure does not have an age upper limit, noting the USPSTF guidance related to screening for diabetes for patients with high BMI ages between 40-70.
- Developer provided a summary of the literature related to gaps in care. Developer states that their review of the literature suggests that the uninsured are less likely to be screened; Hispanics and black people are also more likely to be screened than white people.
- The Committee agreed that gap that exists based on the literature despite the lack actual data of patient care.

2. Scientific Acceptability of Measure Properties: <u>The measure did not pass the Scientific Acceptability criteria</u>

(2a. Reliability precise specifications, testing; 2b. Validity testing, threats to validity
2a. Reliability: **H-1; M-16; L-5; I-0** 2b. Validity: **H-0; M-8; L-9; I-1**

<u>Rationale</u>:

- Developer used same testing for both data element reliability and validity.
- Developer performed data element reliability/validity testing at two facilities on two common EHR systems, Epic and Cerner.
    - Test Site #1: An ambulatory facility in South Carolina, part of a larger health system comprised of 8 inpatient hospitals and more than 100 outpatient facilities. This facility uses Epic EHR in their facility.
    - Test Site #2: An ambulatory facility in South Carolina, part of a larger system comprised of a 1,600+ bed comprehensive integrated health system, serving 1 million patients. This facility uses Cerner EHR in their facility.
- The feasibility assessment indicated the following data elements had issues in the accuracy domain indicating that these data elements may not be correct:
    - "Laboratory Test, Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (in Cerner and Epic) (measure developer noted that Fasting status of glucose testing is not captured in discrete fields in either EHR, however capturing A1C testing is feasible. To test for prediabetes, fasting plasma glucose, 2-h plasma glucose during 75-g oral glucose tolerance test, and A1C are equally appropriate)
    - "Intervention Order: Comfort Measures" (in Cerner) (measure developer noted that Comfort Care as an exclusion is standard in in NQF endocrine registry measures and it is expected that EMR developers to create a distinct field to collect this data in the future)
    - "Laboratory Test, Not Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Glucose [Mass/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner)
    - "Laboratory Test, Not Performed: Glucose [Moles/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner)
    - "Laboratory Test, Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
- Data element reliability/validity testing was conducted utilizing Parallel Forms Reliability Testing methodology to determine if data elements found through electronic data pulls could be confirmed by manual abstraction of the same data elements.
    - Verification of the data elements was obtained through automated data search strategies against a reference strategy (considered the gold standard) for obtaining the data elements.
    - Manual review of the data elements was used as the reference strategy against which automated data search and extraction strategies were evaluated.
    - Interrater reliability (crude agreement and Cohen's Kappa) was used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record.
- Committee members raised concerns that this measure was only tested in two electronic health record (EHR) systems and was not tested with an EHR system less robust than Epic or Cerner.
- The Committee noted that the accuracy results were not clear and that there may be poorer results in smaller EHR systems.

3569e Prediabetes: Screening for Abnormal Blood Glucose

- The Committee noted that several of the data elements had accuracy issues and could present challenges with acquiring data across different providers.
- The Committee expressed particular concern that there were no upper limits for age on this measure.

3. Feasibility: H-0; M-5; L-14; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- During the discussion of feasibility, the Committee raised concerns about the lack of fasting glucose being listed as such in the EMR and the fact that that comfort measures are not necessarily standard.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-17; No Pass-3** *4b. Usability:* **H-0; M-18; L-1; I-1**

Rationale:

- The Committee did not express any concerns with use and usability.

5. Related and Competing Measures

- No related or competing measures noted.

6. Standing Committee Recommendation for Endorsement: N/A

The Standing Committee did not vote on an endorsement recommendation because the measure did not pass the scientific acceptability criteria.

7. Public and Member Comment

- Preferred “abnormal blood glucose” over “prediabetes” which is suggested to be a risk factor rather than a disease.
- Some commenters opposed the missing upper age limit (40-70 years) included in AAFP and USPSTF guidelines.
- One commenter suggested that confirmation of results should be included in this measure.
- Several commenters had concerns with data capture, such as fasting glucose or exclusions not in EHR distinct field, and that the measure was only tested in EPIC and Cerner.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-X; No-X (November 17, 2020: [Endorsed or Not Endorsed])

The CSAC upheld [or did not uphold] the Standing Committee’s decision to recommend the measure for endorsement.

9. Appeals

3570e Intervention for Prediabetes

Submission | Specifications

Description: Percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention*

Numerator Statement: Patients who were provided an intervention*

*Intervention must include one of the following: referral to a CDC-recognized diabetes prevention program; referral to medical nutrition therapy with a registered dietician; prescription of metformin.

Denominator Statement: All patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period

**Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol).

3570e Intervention for Prediabetes

Exclusions: Denominator Exclusions:

Exclude patients who are pregnant.

Exclude patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY])

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Electronic Health Records

Measure Steward: American Medical Association

STANDING COMMITTEE MEETING 06/25/2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-0; M-16; L-2; I-2**; 1b. Performance Gap: **H-2; M-16; L-1; I-1**

<u>Rationale</u>:

- Developer cites evidence primarily found in guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA).
    - USPSTF grade B recommendation states that adults aged 40 to 70 years of age who are overweight or obese should be screened for abnormal blood glucose as part of cardiovascular risk assessment.
    - Referral of patients with abnormal blood glucose to intensive behavioral counseling interventions is also recommended by the USPSTF (grade B recommendation).
    - The ADA recommends an intensive behavioral lifestyle intervention program modeled on the Diabetes Prevention Program for prediabetes patients (grade A recommendation).
    - The developer cites ADA's grade A recommendation on Metformin therapy for preventing type 2 diabetes in individuals with prediabetes (<60 years, BMI ≥35 kg/m2 and women with prior gestational diabetes mellitus).
    - An individualized medical nutrition therapy is recommended by ADA for all with type 1 or type 2 diabetes or gestational diabetes mellitus (grade A recommendation).
- The Committee noted that this measure could be specified as an outcome measure but acknowledged that providers may not yet have the processes in place to achieve outcomes.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability precise specifications, testing; 2b. Validity testing, threats to validity

2a. Reliability: **H-0; M-16; L-3; I-0**; 2b. Validity: **H-0; M-13; L-3; I-3**

<u>Rationale</u>:

- Developer used same testing for both data element reliability and validity.
- Developer performed data element reliability/validity testing at two facilities on two common EHR systems.
    - Test Site #1: An ambulatory facility in South Carolina, part of a larger health system comprised of 8 inpatient hospitals and more than 100 outpatient facilities. This facility uses Epic EHR in their facility.
    - Test Site #2: An ambulatory facility in South Carolina, part of a larger system comprised of a 1,600+ bed comprehensive integrated health system, serving 1 million patients. This facility uses Cerner EHR in their facility.
- Submission includes simulated data set results demonstrating unit testing covering 100% of the measure logic.
- The feasibility assessment indicated the following data elements had issues in the accuracy domain indicating that these data elements may not be correct:
    - "Laboratory Test, Performed: Glucose [Mass/volume] in Serum or Plasma --2 hours post 75 g glucose PO" Measure developer noted that fasting status of glucose testing is not captured in discrete fields in either EHR, however capturing A1C testing is feasible. To test for pre-

3570e Intervention for Prediabetes

diabetes, fasting plasma glucose, 2-h plasma glucose during 75-g oral glucose tolerance test, and A1C are equally appropriate (in Cerner and Epic)
    - "Laboratory Test, Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Intervention, Performed: Referral to Dietitian" (measure developer noted that It is expected that EMR developers to create a distinct field to collect this data in the future.) (in Cerner and Epic)
    - "Intervention, Performed: Referral to Diabetes Prevention Program" (measure developer noted that It is expected that EMR developers to create a distinct field to collect this data in the future) (in Cerner and Epic)
    - "Intervention, Not Performed: Referral to Diabetes Prevention Program" (in Cerner and Epic)
    - "Intervention, Not Performed: Referral to Dietitian" (in Cerner and Epic)
    - "Medication, Not Ordered: Metformin" (in Cerner)
    - "Diagnosis: Limited Life Expectancy" (in Cerner)
    - "Encounter, Performed: Nursing Facility Visit" (in Cerner)
- Data element reliability/validity testing was conducted utilizing Parallel Forms Reliability Testing methodology to determine if data elements found through electronic data pulls could be confirmed by manual abstraction of the same data elements.
    - Verification of the data elements was obtained through automated data search strategies against a reference strategy (considered the gold standard) for obtaining the data elements.
    - Manual review of the data elements was used as the reference strategy against which automated data search and extraction strategies were evaluated.
    - Interrater reliability (crude agreement and Cohen's Kappa) was used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record.
- In terms of reliability, the Committee raised concerns around the sampling methodology. The Committee noted that convenience sampling did not necessarily indicate systematic bias.
- The Committee expressed concern that the measure may not have had all data elements tested and that the eCQM feasibility scorecard assessment indicated the many data elements had issues in the accuracy domain, indicating that these data elements may not be accurately captured.

3. Feasibility: H-0; M-5; L-15; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- The Committee raised concerns that the fields needed to collect this measure are not present in the EHR.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-18; No Pass-0** *4b. Usability:* **H-0; M-10; L-6; I-2**

Rationale:

- For usability, the Committee noted that there are potential issues with lack of discrete fields to document the referral and patient lacking access to a diabetes prevention program because their insurance doesn't cover the services.

5. Related and Competing Measures

- No related or competing measures noted.

6. Standing Committee Recommendation for Endorsement: Yes-5; No-13

7. Public and Member Comment

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-X; No-X (November 17, 2020: [Endorsed or Not Endorsed])

The CSAC upheld [or did not uphold] the Standing Committee's decision to recommend the measure for

3570e Intervention for Prediabetes
endorsement.
9. Appeals

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes
Submission
Description: Percentage of patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the previous year who have a blood glucose test performed in the one-year measurement period **Numerator Statement**: Patients who had a blood glucose test performed *Retesting for abnormal blood glucose may include using a fasting plasma glucose, 2-h plasma glucose during a 75g oral glucose tolerance test, or A1C. **Denominator Statement**: All patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the year prior to the one-year measurement period **Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol). **Exclusions**: Denominator Exclusions: Exclude patients who are pregnant. Exclude patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY]). Exclude patients in palliative care/hospice **Adjustment/Stratification**: No risk adjustment or risk stratification **Level of Analysis:** Clinician : Group/Practice, Clinician : Individual **Setting of Care:** Outpatient Services **Type of Measure**: Process **Data Source**: Electronic Health Records **Measure Steward**: American Medical Association

STANDING COMMITTEE MEETING 06/25/2020
1. Importance to Measure and Report: <u>The measure did not pass the Importance criteria</u>
(1a. Evidence, 1b. Performance Gap)
1a. Evidence: **H-0; M-7; L-7; I-4;** 1b. Performance Gap: **H-1; M-10; L-3; I-3;** Evidence Exception: **Yes-9; No-8**
<u>Rationale</u>:
• Developer cites evidence found in guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA).
  ○ At least annual monitoring for the development of diabetes in those with prediabetes is suggested. (ADA, 2018) (E Recommendation)
  ○ Developer provides evidence of disease prevalence and systematic misses of opportunities to intervene by clinicians.
  ○ Developer does not provide studies that offer evidence that annual monitoring is associated with positive outcomes.
• The Committee noted that there is a lack of evidence to support this measure.
• The Committee raised the concern that the quality measurement enterprise generally has sufficient process measures and not enough outcome measures.
• The Committee observed the developer's review of the literature that suggests a gap in care, noting that the United States has 84 million adults with prediabetes, that 9 out of 10 patients who have prediabetes are not aware, and that missed opportunities among primary care providers in diagnosing and managing patients with prediabetes represent a gap in care.

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes

2. Scientific Acceptability of Measure Properties: <u>The measure did not achieve consensus on the Scientific Acceptability criteria</u>

(2a. Reliability precise specifications, testing; 2b. Validity testing, threats to validity

2a. Reliability: **H-0; M-12; L-5; I-0**; 2b. Validity: **H-0; M-9; L-7; I-1**

<u>Rationale</u>:

- Developer used same testing for both data element reliability and validity.
- Developer performed data element reliability/validity testing at two facilities on two common EHR systems.
    - Test Site #1: An ambulatory facility in South Carolina, part of a larger health system comprised of 8 inpatient hospitals and more than 100 outpatient facilities. This facility uses Epic EHR in their facility.
    - Test Site #2: An ambulatory facility in South Carolina, part of a larger system comprised of a 1,600+ bed comprehensive integrated health system, serving 1 million patients. This facility uses Cerner EHR in their facility.
- Submission includes simulated data set results demonstrating unit testing covering 100% of the measure logic.
- The feasibility assessment indicated the following data elements had issues in the accuracy domain indicating that these data elements may not be correct:
    - Laboratory Test, Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (measure developer noted that fasting status of glucose testing is not captured in discrete fields in either EHR, however capturing A1C testing is feasible. To test for prediabetes, fasting plasma glucose, 2-h plasma glucose during 75-g oral glucose tolerance test, and A1C are equally appropriate. (in Cerner and Epic)
    - "Laboratory Test, Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Glucose [Moles/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner)
    - "Laboratory Test, Not Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Glucose [Mass/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner)
    - "Intervention, Order: Comfort Measures" using "Comfort Measures (2.16.840.1.113883.17.4077.3.2030)" (measure developer noted that Comfort Care as an exclusion is standard in NQF endocrine registry measures and it is expected that EMR developers to create a distinct field to collect this data in the future) (in Cerner)
- Data element reliability/validity testing was conducted utilizing Parallel Forms Reliability Testing methodology to determine if data elements found through electronic data pulls could be confirmed by manual abstraction of the same data elements.
    - Verification of the data elements was obtained through automated data search strategies against a reference strategy (considered the gold standard) for obtaining the data elements.
    - Manual review of the data elements was used as the reference strategy against which automated data search and extraction strategies were evaluated.
    - Interrater reliability (crude agreement and Cohen's Kappa) was used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record.
- The Committee did not reach consensus on validity, but noted that the measure had concerns associated with the feasibility scorecard in that the accuracy of the data elements was questionable.

3. Feasibility: H-0; M-7; L-9; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

<u>Rationale</u>:

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes
• In the review of the measure's feasibility, the Committee was concerned that reporting the measure may be challenging since the accuracy of the data elements was not clear.
4. Use and Usability *4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)* 4a. Use: **Pass-15; No Pass-2** *4b. Usability:* **H-0; M-7; L-6; I-3** Rationale: • The Committee noted that the measure has not been implemented, but the developer has the intention of submitting the measure to CMS for the MIPS program
5. Related and Competing Measures • No related or competing measures noted.
6. Standing Committee Recommendation for Endorsement: **N/A** The Standing Committee did not vote on an endorsement recommendation because the measure did not pass the scientific acceptability criteria.
7. Public and Member Comment • One commentator noted, "there is limited evidence on the best rescreening intervals for adults with normal results; however, screening every 3 years is a reasonable option." In contrast, this measure requires re-testing at least annually. • In addition, the exclusions for this measure are different from the others. Comfort care is not included in this measure. • Other comment agreed that retesting is needed but that the testing should include a variety of tests, a specific timeframe, coverage by insurance, and ease of access to tests.
8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-X; No-X (November 17, 2020: [Endorsed or Not Endorsed]) The CSAC upheld [or did not uphold] the Standing Committee's decision to recommend the measure for endorsement.
9. Appeals

NATIONAL QUALITY FORUM
Driving measurable health improvements together

http://www.qualityforum.org

Primary Care and Chronic Illness Spring 2020 Review Cycle

CSAC Review and Endorsement

November 17, 2020

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

NATIONAL QUALITY FORUM

Standing Committee Recommendations

- Three measures reviewed for Spring 2020
  - No measures reviewed by the Scientific Methods Panel
- Three measures not recommended for endorsement
  - **NQF 3570e** Intervention for Prediabetes (New Measure)
  - **NQF 3569e** Prediabetes: Screening for Abnormal Blood Glucose (AMA) (New Measure)
  - **NQF 3571e** Retesting of Abnormal Blood Glucose in Patients with Prediabetes (AMA) (New Measure)

NATIONAL QUALITY FORUM

Overarching Issues

- Upper age limits
  - NQF 3569e did not have an upper age limit, although this was included in some of the evidence cited by the developer. The Committee found this particular item especially concerning
- EHRs used in testing
  - The three measures were tested in two of the strongest EHRs—Epic and Cerner. The concern was that this measure would not be feasible in smaller EHRs
- Data element accuracy and data capture
  - Developer noted in submission that several data elements, especially for measure numerators, may have accuracy challenges as expressed in the eCQM Feasibility Scorecard
  - This was counterbalanced by the developer's parallel forms validity results

Public and Member Comment and Member Expressions of Support

- 34 total comments received
- No NQF member of expressions of support or not support received

Questions?

- Project team:
  - Sam Stolpe, PharmD, MPH, Senior Director
  - Yemi Kidane, PMP, Project Manager
  - Erin Buchanan, MPH, Manager
  - Isaac Sakyi, MSGH, Analyst

- Project webpage: https://www.qualityforum.org/Primary_Care_and_Chronic_Illness.aspx

- Project email address: primarycare@qualityforum.org

THANK YOU.

NATIONAL QUALITY FORUM

http://www.qualityforum.org

NATIONAL QUALITY FORUM
Driving measurable health improvements together

Primary Care and Chronic Illness, Spring 2020 Cycle: CDP Report

DRAFT REPORT FOR CSAC REVIEW
NOVEMBER 17, 2020

This report is funded by the Centers for Medicare and Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001

http://www.qualityforum.org

Contents

Executive Summary

Primary care providers serve as the most common contact point for many people within the U.S. healthcare system. As such, primary care has a central role in improving the health of people and populations. Primary care practitioners work with each patient to manage the health of that individual. In the primary care setting, the diagnosis and treatment of the patient focuses on the health of the entire patient and not a single disease.

Chronic illnesses are long-lasting or persistent health conditions or diseases that patients and providers must manage on an ongoing basis. The incidence, impact, and cost of chronic disease is increasing in the United States. For example, more than 30 million Americans (9.4 percent) are living with diabetes, and in 2017, the U.S. spent $237 billion on diabetes care, making it one of the most expensive health conditions.[1,2] In addition, studies have estimated the yearly costs for glaucoma, rheumatoid arthritis and hepatitis C at $5.8 billion, $19.3 billion, and $6.5 billion, respectively.[3–5] The net economic burden for medication nonadherence—a common issue with primary care patients—has been estimated at nearly $300 billion per year.[6]

The review and evaluation of measures impacting primary care and dealing with chronic illness has long been a priority of National Quality Forum (NQF), with endorsement for such measures going back to NQF's inception. At present, there are 47 NQF-endorsed primary care and chronic illness measures. The background and description of NQF's most recent Primary Care and Chronic Illness (PCCI) Standing Committee meeting as well as previous meetings are available on NQF's project webpage. This Committee oversees the measurement portfolio used to advance accountability and quality in the delivery of primary care services.

For this project, the Standing Committee evaluated three newly submitted measures against NQF's standard evaluation criteria. During the measure evaluation meeting, the Committee did not recommend one measure for endorsement (NQF 3570e) and did not reach consensus on two measures during the measure evaluation meeting (3569e and 3571e).

Following the post-comment meeting, the Committee did not recommend the following measures:

- **NQF 3569e** Prediabetes: Screening for Abnormal Blood Glucose (American Medical Association)
- **NQF 3570e** Intervention for Prediabetes (American Medical Association)
- **NQF 3571e** Retesting of Abnormal Blood Glucose in Patients with Prediabetes (American Medical Association)

Brief summaries of the measures currently under review are included in the body of the report; detailed summaries of the Committee's discussion and ratings of the criteria for each measure are in Appendix A.

Introduction

Primary care providers serve as the most common contact point for many people within the U.S. healthcare system. As such, primary care has a central role in improving the health of people and populations.

Over the last 15 years, NQF has endorsed dozens of measures addressing improvements in primary care and chronic illnesses. These measures are used in many national and state-level public reporting and accountability programs, as well as for quality improvement. With the formation of the Primary Care and Chronic Illness Standing Committee in 2017, NQF was able to consolidate and streamline the measure maintenance and endorsement process for a broad set of measures related to primary care and chronic illness.

High-quality performance measurement that captures the complexity of primary care and chronic illnesses is essential to improve diagnosis, treatment, and management of conditions. NQF will review measures in these important healthcare areas under a consolidated measure portfolio that reflects the importance of caring for chronic illness in primary care settings. Measures may focus on nonsurgical eyes or ears, nose, and throat conditions; diabetes care, osteoporosis; HIV; rheumatoid arthritis; gout; back pain; asthma; chronic obstructive pulmonary disease (COPD); and acute bronchitis.

NQF Portfolio of Performance Measures for Primary Care and Chronic Illness Conditions

The Primary Care and Chronic Illness Standing Committee (Appendix C) oversees NQF's portfolio of Primary Care and Chronic Illness measures (Appendix B) that includes 47 measures: 40 process measures, five outcome and resource use measures, one intermediate outcome measure, and one composite measure (see table below).

Table 1. NQF Primary Care and Chronic Illness Portfolio of Measures

	Process	Outcome	Intermediate Outcome	Composite
Ears, Nose, Throat (ENT), Eye Care	14	0	0	0
Endocrine	6	3	0	1
Infectious Disease	8	2	1	0
Musculoskeletal	6	0	0	0
Pulmonary	5	0	0	0
Other	1	0	0	0
Total	40	5	1	1

Other measures related to primary care and chronic illness have been assigned to other portfolios. These include functional status measures (Patient Experience and Function), opioid use measures (Patient Safety and Behavioral Health and Substance Abuse), diabetes-related admission rate measures

(Prevention and Population Health), and a variety of condition- or population-specific measures (Cardiovascular, Pediatric, Geriatrics and Palliative Care, etc.).

Primary Care and Chronic Illness Measure Evaluation

On June 25, June 26, and July 7, the Primary Care and Chronic Illness Standing Committee evaluated 3 new measures against NQF's standard measure evaluation criteria.

Table 2. Primary Care and Chronic Illness Measure Evaluation Summary

	Maintenance	New	Total
Measures under consideration	0	3	3
Measures recommended for endorsement	0	0	0
Measures where consensus is not yet reached	0	2	2
Measures not recommended for endorsement	0	1	1
Reasons for not recommending	Importance – 0 Scientific Acceptability – 0 Use – 0 Overall Suitability – 0 Competing Measure – 0	Importance – 1 Scientific Acceptability – 1 Use – 0 Overall Suitability – 1 Competing Measure – 0	

Comments Received Prior to Committee Evaluation

NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). In addition, NQF solicits comments for a continuous period during each evaluation cycle via an online tool located on the project webpage. Pre-meeting commenting closed on September 3, 2020. As of that date, no comments were submitted.

Comments Received After Committee Evaluation

The continuous public commenting period with NQF member support closed on August 5, 2020. Following the Committee's evaluation of the measures under consideration, NQF received 34 comments from 12 organizations (including 1 member organizations) and individuals pertaining to the draft report and to the measures under consideration. All comments for each measure under consideration have been summarized in Appendix A.

Throughout the continuous public commenting period, NQF members had the opportunity to express their support ('support' or 'do not support') for each measure submitted for endorsement consideration to inform the Committee's recommendations. no NQF members provided their expression of support.

Overarching Issue

During the Standing Committee's discussion of the measures, an overarching issue emerged that was factored into the Committee's ratings and recommendations for multiple measures and is not repeated in detail with each individual measure.

Accuracy Issues in Data Capture for eCQMs

Each of the measures evaluated this cycle was an electronic clinical quality measure (eCQM). NQF requires measure developers to complete a feasibility scorecard for eCQMs that explores the ability of a given electronic health record (EHR) to capture the data fields necessary to calculate the measure according to the measure specifications. NQF has emphasized the need to move toward reducing the burden associated with quality measurement, with electronic measures that use data that occurs as part of documentation of normal care delivery within structured EHR fields as an important alternative to the more cumbersome measures that draw from medical chart abstraction. Nonetheless, many EHRs were not originally designed to serve as data sources for quality measurement and this can be problematic in calculating eCQMs. Moreover, the structured fields that would be useful to populate a measure are often not present even in more advanced EHRs. This creates tension in the measure evaluation process when eCQMs do not exhibit high accuracy during feasibility scorecard testing. There has been a concern that providers could be prospectively held accountable for eCQMs that do not display reliable calculation based on accuracy issues during the feasibility scorecard testing. Developers often address those concerns by citing EHR vendor commitment to the implementation of structured fields for capture of data critical to eCQM calculation if and when those measures are required, for example as part of reporting eCQMs within federal quality programs.

The NQF Primary Care and Chronic Illness Committee noted that the measure developer tested the three eCQMs evaluated this cycle within Epic and Cerner. These are the two largest EHR vendors and widely regarded as among the most advanced. The Committee expressed concerns that strong accuracy was not reflected in the feasibility scorecard testing. This was especially true of data elements related to the focus of the measure, for example the capture of fasting blood glucose testing for NQF 3569e Prediabetes: Screening for Abnormal Blood Glucose. The Committee expressed concerns that the clinician could order such a test for patients who would fall in the denominator of the measure, such a test could be performed and documented, but the test is not accurately captured by the measure. The Committee considered accuracy issues in feasibility scorecard testing for the eCQMs to be threats to both the validity and feasibility of the measures.

Summary of Measure Evaluation

The following brief summaries of the measure evaluation highlight the major issues that the Committee considered. Details of the Committee's discussion and ratings of the criteria for each measure are included in Appendix A.

3569e Prediabetes: Screening for Abnormal Blood Glucose (American Medical Association): Not Recommended

Description: Percentage of patients aged 40 years and older with a BMI greater than or equal to 25 who are seen for at least two office visits or at least one preventive visit during the 12-month period who were screened for abnormal blood glucose at least once in the last 3 years; **Measure Type**: Process;

Level of Analysis: Clinician: Group/Practice, Clinician: Individual; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

The Standing Committee did not recommend this measure for endorsement.

The Committee noted that this is a new process measure which assesses the percentage of patients aged 40 years and older with a BMI greater than or equal to 25 who are seen for at least two office visits or at least one preventive visit during the 12-month period who were screened for abnormal blood glucose at least once in the last 3 years. The Committee indicated support of measures that address prediabetes, acknowledging a gap in NQF-endorsed measures that specifically address prediabetes. Concerning the evidence criterion, Committee members agreed this is an important area of measurement and determined that the evidence submitted generally supports the measure. The Committee noted that the developer cited guidelines from the American Diabetes Association (ADA) as well as from the United States Preventative Services Task Force (USPSTF). The Committee questioned the reason that the measure does not have an age upper limit, noting the USPSTF guidance related to screening for diabetes for patients with high BMI between 40-70. The Committee agreed a performance gap exists based on the literature despite the lack actual data on patient care. During the discussion around reliability, Committee members raised concerns that while it does conform to the NQF evaluation criteria requiring it to be tested in more than one EHR, this measure was not tested in an EHR system less robust than Epic or Cerner. The Committee was concerned with the eCQM feasibility scorecard for Epic and Cerner, noting that the accuracy results were not clear and that there may be poorer results in smaller EHR systems. The Committee passed the measure on reliability. Regarding validity testing, the Committee raised several concerns. The Committee noted that several of the data elements had accuracy issues and could present challenges with acquiring data across different providers. In particular, the Committee reviewed accuracy issues in the feasibility scorecard for eCQMs for fields related to fasting plasma glucose lab testing, glucose in serum plasma lab testing, and exclusions related to intervention orders for comfort care. The Committee expressed a concern that since the focus of the measure is determining whether an appropriate test has been conducted, the measure should be especially accurate in detecting when such a test has occurred for patients in the denominator of the measure. Consensus was not reached on the validity of this measure. The measure was not regarded as feasible by Committee members citing the fact that fasting plasma glucose is not routinely captured during care and the fact that that the exclusion of comfort measures is not easily captured in EHR software. The Committee did not express any concerns with use and usability.

Measure NQF 3569e did not achieve consensus on validity during the initial measure evaluation meeting. In the discussion of comments received related to NQF 3569e, the Committee first turned to the measure developer, the American Medical Association (AMA) to ask for a summary of their responses to the comments. In response to the comment that suggested that the term “prediabetes” was inappropriate because it confers the suggestion of a disease and expressed a preference for the term “abnormal blood glucose”, the developer first provided an acknowledgement that the measure title itself includes reference to abnormal glucose and that they have noted the input point. The developer also responded to comments related to the fact that this measure does not include an upper age limit exclusion, noting that this point was debated within their own technical expert panel (TEP) resulting in a consensus not reached vote which lead directly in not including an upper age limit. The developer stated that not including an upper age limit is aligned with the recommendation from

American Diabetes Association (ADA) guidelines. The developer also alluded to evidence that suggested that an upper age limit for the measure is not appropriate given that older patients have been shown to benefit from screening as well. The requirement of confirmation of results was noted by AMA's TEP to not be pragmatic nor aligned with United States Preventative Services Task Force (USPSTF) guidelines. The comments also reflected previous Committee discussion around validity concerns expressed during the initial July measure evaluation meeting. At that time, several Committee members had expressed reservations associated with the accuracy domain of the Feasibility Scorecard for data elements related to fasting glucose tests. The developer noted that the parallel forms validity testing that they performed resolved initial concerns related to the calculation of the measure as suggested in the Feasibility Scorecard, noting moderate to excellent crude agreement and kappa statistics between abstractors and calculations from the eCQM. The developer emphasized that the fasting glucose data element that had accuracy concerns were directly addressed through the feasibility testing. The developer was questioned on their assertion that the overwhelming majority of records were HbA1c data elements and was asked what percentage of data elements pulled were fasting blood glucose. The developer indicated that the fasting blood glucose accounted for less than 10% of the data.

Several Committee members disagreed with the developer on not including an upper age limit, viewing the lack of the upper age limit as a threat to the validity of the measure. The Committee asked the developer to highlight the evidence that older patients benefit from such interventions. The developer reviewed their references included in responses to comments with the Committee, including ADA guideline screening recommendations, a smaller study by Kramer, et al., and evidence from the National Diabetes Prevention Program. AMA also noted that their TEP felt that having measure exclusions for patients with limited functional status or limited life expectancy were sufficient to identify those who should not be screened. One member agreed with the developer that not having an upper age limit was appropriate based on experiences managing lifestyle change programs. Another member noted this but added that they were concerned that during an appointment with especially older patients, a clinician may be required to perform a screening that they did not consider appropriate or face be penalized on their measure performance. Other members pointed out that the developer could simply adopt the 40-to-70-year age group suggested by the comments from American Academy of Family Physicians (AAFP) and American Geriatrics Society (AGS). This solution would not prohibit clinicians from still screening patients who were older but rather provides a known and supported age span for the purposes of accountability, allowing for a consistent measure denominator. The developer responded that the expectation for performance on the measure is not to achieve perfection and that performance cut points can be used to account for instances where clinicians may determine that it may not be appropriate for certain patients to be screened.

The Committee asked if age-range concerns were appropriate to consider within a validity discussion, noting reservations around supporting the measure with the current age limits. NQF staff reaffirmed that if the Committee felt that the definitions that were used to capture the patient population within the measure do not align with clinical recommendations, it has direct bearing on the question of whether the metric does in fact measure what it purports to measure, which is the central question of validity.

The Committee did not support the measure on validity, a must-pass criterion. The Committee did not recommend the measure for endorsement.

3570e Intervention for Prediabetes (American Medical Association): Not Recommended

Description: Percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention; **Measure Type**: Process; **Level of Analysis**: Clinician : Group/Practice, Clinician : Individual; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

The Standing Committee did not recommend the measure for endorsement. This is a new process measure which assesses the percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention. The Committee noted that overall there was good evidence for this measure and passed on this criterion. The Committee also noted that this measure could be an outcome measure but recognized that providers may not have the processes in place to achieve those outcomes and therefore a process measure is still useful. The Committee had no concerns about performance gap. In terms of reliability, the Committee raised concerns about sampling methodology. The Committee noted that convenience sampling did not necessarily indicate systematic bias. The Committee passed this measure on reliability. The Committee passed the measure on validity, but noted that the measure had concerns associated with the feasibility scorecard in that the accuracy of the data elements was questionable. The Committee did not pass the measure on feasibility raising concerns that the fields needed to collect this measure are not present in the EHR. The Committee acknowledged that the missing data will most likely be able to be captured in the future but note that currently this measure presents too great of a burden for implementation as manual review would be needed to confirm accuracy. The Committee did not have any concerns on use. For usability, the Committee noted that there are potential issues with lack of discrete fields to document the referral and patient lacking access to a diabetes prevention program because their insurance doesn't cover it. The Committee passed this measure on usability. The Committee observed that there are no related and competing measures to discuss for this measure.

During the Spring 2020 post-comment meeting, NQF staff summarized the comments received, noting that some commenters called into question the interventions contained within the measure specifications, namely either referral to CDC-recognized diabetes prevention program (DPP), referral to medical nutrition therapy with a registered dietician, or prescription of metformin. One commenter noted that intensive behavioral counseling or other interventions are not adequately represented in this measure, making the measure poorly aligned with current guidelines and best practices. Commenters noted that the options of prescribing metformin or referring patients out will either be burdensome and drive up cost, or result in a narrow, specific pharmacotherapeutic option. It was also noted that DPPs are not widely available through the entire country. One Committee member noted that programs based on DPP protocols are fairly well available throughout the country. Another member added that poor bandwidth is now the primary barrier, but telehealth and virtual dashboards are beginning to address access challenges for rural areas, also noting that many health plans are covering the service. Another Committee member noted that they had a challenge in accessing this service himself under his insurance carrier unless he was coded as diabetic. Another Committee member noted that there are provisions for Medicare beneficiaries that makes DPP widely accessible. Other Committee members expressed concern that the measure equates the three interventions when evidence suggests that behavioral interventions are stronger than metformin. Another Committee member expressed support for this remark and added that from a feasibility perspective this fact alone creates a lot of challenges.

Committee members expressed concerns associated with the unintended consequences of driving a greater utilization of metformin.

The Committee then invited the measure developer to outline their rationale for their reconsideration request. The AMA noted that the measures were submitted according to NQF measure evaluation criteria and with significant effort put into the development and testing of the measures. The AMA stated that they are concerned that the criteria for feasibility, scientific acceptability (particularly validity), and usability were not applied appropriately. Related to the feasibility and validity concerns, the developer stated that the results of the validity testing demonstrate that the results of the parallel forms tested resulted in kappa statistics indicating moderate to near-perfect agreement. AMA further acknowledged that not every data element was captured in the two EHR systems tested but noted that the validity testing showed that the results produced were acceptable. Moreover, AMA noted that EHR systems will improve to better capture the data elements needed as organizations begin working to implement and track these measures. AMA referred to previous dialogue related to fasting blood glucose, which was relevant for all three measures. The concern that this element was not captured in structured data fields was not found to be problematic within AMA's data element validity testing, with what AMA characterized as a nearly 0% occurrence. A representative endocrinologist from AMA's TEP added additional commentary that the measure does not say that the three interventions are equivalent, but that there are different options. Further, she noted that the Committee emphasized that within the DPP study, intensive lifestyle interventions were the most efficacious but there are other studies concluding that metformin is equivalent for certain populations and certain conditions. It was also emphasized that comparative effectiveness studies of virtually delivered DPP interventions have shown similar weight loss outcomes.

One Committee member noted that the measure is not doing enough to improve patient outcomes, adding that while there are options, the measure does treat the three interventions as equivalent and that there may be unintended consequences associated with that. The member further suggested that if it were framed as all of these options being offered to the patient, then that would be different, but as the measure is constructed there is only one box that can be checked, and this will not necessarily lead to the same results.

During the measure evaluation meeting, the Committee did not pass the measure on feasibility and did not recommend the measure for endorsement. During the post-comment meeting, the Committee did not approve this measure for reconsideration

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes (American Medical Association): Not Recommended

Description: Percentage of patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the previous year who have a blood glucose test performed in the one-year measurement period; **Measure Type**: Process; **Level of Analysis**: Clinician : Group/Practice, Clinician : Individual; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

The Standing Committee did not recommend the measure for endorsement.

This is a new process measure which assesses the percentage of patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the previous year who have a blood glucose test performed in the one-year measurement period. The Committee began the discussion with a review of the evidence. The Committee questioned whether there was evidence to suggest that testing within one year is the correct time frame. The developer noted that the ADA recommended at least an annual retesting. Nonetheless, the Committee noted that there may be unintended consequences associated with testing frequently, namely false positives in testing for diabetes which will increase along with testing frequency. One Committee member noted that this is a process measure with less evidence to support it and expressed concern that the quality measurement enterprise generally has sufficient process measures and not enough outcome measures. When voting on evidence, the Committee did not pass the measure. Since the evidence, this measure is based on is expert opinion rather than randomized control trials, the committee took a vote to grant an exception to evidence. Consensus was not reached on the vote to grant an exception to the evidence provided. The Committee observed the developer's review of the literature that suggests a gap in care, noting that the United States has 84 million adults with prediabetes, that 9 out of 10 patients who have prediabetes are not aware, and that missed opportunities among primary care providers in diagnosing and managing patients with prediabetes represent a gap in care. In the discussion on validity, the Committee expressed some concern that the measure may not have had all data elements tested and that the eCQM feasibility scorecard assessment suggested that many data elements had issues in the accuracy domain, indicating that these data elements may not be accurately captured. The Committee did not achieve consensus on validity. In the review of the measure's feasibility, the Committee was also concerned that reporting the measure may be challenging since the accuracy of the data elements was not clear. The Committee did not reach consensus for the measure on feasibility. In the discussion on use, the Committee noted that the measure has not been implemented, but the developer has the intention of submitting the measure to CMS for the MIPS program. During the discussion on usability, the Committee noted that diabetes testing is not completely harmless since going into a primary care provider for regular screening can be burdensome for patients due to peripheral costs and inconvenience. The Committee did not achieve consensus on usability.

NQF staff noted that the measure did not pass during the initial measure evaluation meeting and that the developer had since provided a reconsideration request. The developer suggested within that request that the Committee had been inconsistent in the application of NQF criteria and that the Committee had conflated validity and feasibility. Moreover, the developer suggested that it was not clear why it was that the measure passed on validity where the other two did not achieve consensus on validity and did not pass on feasibility. The developer also noted that the measure passed all must-pass criteria but did not receive overall endorsement.

During the spring 2020 post-comment meeting, NQF staff noted that consensus was not reached during the measure evaluation meeting for NQF 3571e on evidence and validity, both must-pass criteria. The comments received reflected concerns associated with evidence on the screening interval of one year, that exclusions for this measure differed from the other eCQMs submitted by AMA (e.g. comfort care not included in this measure), and that testing should include a variety of tests, a specific time frame, and include considerations associated with access. The developer responded to those concerns by noting that public comments were generally supportive of an exception to evidence and emphasized

that their validity testing performed addressed many of the issues raised related to the accuracy of data elements.

The Committee began the discussion by reviewing a comment and response from AAFP focused on the screening interval. AAFP asserted that a three-year interval would be more appropriate. In their response, the developer noted that a three-year interval is appropriate for normal glucose readings (USPSTF), but that annual testing is appropriate if an abnormal glucose result is obtained (ADA). A Committee member suggested that there is not sufficient evidence that supports the one-year rescreening interval because it has not been directly correlated with better outcomes but is based on expert opinion. Another Committee member countered that the test itself is not overly burdensome and seems appropriate.

NQF staff then reviewed the evidence discussion in the July measure evaluation meeting, noting that the developer cited the USPSTF and ADA guidelines as evidence for the measure. The developer noted that the annual testing recommendation came directly from the ADA guideline where it was given an "E" grade, meaning that it is based on expert opinion. Staff then reviewed the NQF criteria for evidence submissions, including a detailed walkthrough of the evidence algorithm found in NQF's 2019 measure evaluation criteria, highlighting the pathway of exception to evidence for measures rated as "insufficient" because they are based on expert opinion. The Committee asked the developer if there was a systematic review associated with benefits and risks of the intervention as part of the expert opinion recommendations. The developer referred to the ADA guidelines and their own TEP review of the measure to indicate that a careful review of existing evidence was conducted prior to providing that expert opinion.

The developer was asked if patients who were prediabetic and found to be stable for a lengthy period of time would be excluded, but the developer noted that there is not a ready approach to guide such an exclusion.

The Committee did not pass the measure on evidence, a must-pass criterion. The Committee did not recommend the measure for endorsement.

Measures Withdrawn from Consideration

There were no measures withdrawn from consideration this cycle.

References

1 Centers for Disease Control and Prevention. *National Diabetes Statistics Report, 2017*. Atlanta, GA: Centers for Disease Control and Prevention; 2017. https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf. Last accessed July 2019.

2 American Diabetes Association. Economic costs of diabetes in the U.S. in 2017. *Diabetes Care*. 2018;41(5):917-928.

3 Birnbaum H, Pike C, Kaufman R, et al. Societal cost of rheumatoid arthritis patients in the US. *Curr Med Res Opin*. 2010;26(1):77-90.

4 Prevent Blindness. Glaucoma Costs Reach $5.8 Billion Annually. Prevent Blindness. https://www.preventblindness.org/glaucoma-costs-reach-5-point-8-billion-annually. Last accessed July 2019.

5 Razavi H, ElKhoury AC, Elbasha E, et al. Chronic hepatitis C virus (HCV) disease burden and cost in the United States. *Hepatology*. 2013;57(6):2164-2170.

6 Network for Excellence in Health Innovation (NEHI). *Improving Patient Medication Adherence: A $290 Billion Opportunity*. Boston, MA: NEHI; 2011. https://www.nehi.net/bendthecurve/sup/documents/Medication_Adherence_Brief.pdf. Last accessed July 2019.

Appendix A: Details of Measure Evaluation

Rating Scale: H=High; **M**=Moderate; **L**=Low; **I**=Insufficient; **NA**=Not Applicable

Measures Not Recommended

3569e Prediabetes: Screening for Abnormal Blood Glucose

Submission | Specifications

Description: Percentage of patients aged 40 years and older with a BMI greater than or equal to 25 who are seen for at least two office visits or at least one preventive visit during the 12-month period who were screened for abnormal blood glucose at least once in the last 3 years

Numerator Statement: *Screening for abnormal blood glucose may include using a fasting plasma glucose, 2-h plasma glucose during a 75g oral glucose tolerance test, or A1C.

Denominator Statement: All patients aged 43 years and older with a BMI greater than or equal to 25 seen for at least two office visits or at least one preventive visit during the 12-month measurement period

Exclusions: Denominator Exclusions

"Patient is Pregnant at Encounter"

or "Patient Has Active Diabetes Diagnosis at Encounter"

or "Hospice During Measurement Period"

or "Palliative Care During Measurement Period"

or "Comfort Measures During Measurement Period"

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Electronic Health Records

Measure Steward: American Medical Association

STANDING COMMITTEE MEETING 06/25/2020

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-4; M-17; L-1; I-0**; 1b. Performance Gap: **H-5; M-17; L-0; I-0**

Rationale:

- Developer cites evidence found in guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA).
    - The focus of the recommendations is lifestyle change.
    - USPSTF incorporated this evidence into the updated recommendation regarding screening for abnormal glucose and type 2 diabetes.
    - The grade B recommendation states that physicians should screen individuals for abnormal glucose if they are between the ages of 40 and 70 and are overweight or obese, or younger if they have additional risk factors.
    - The ADA recommends screening patients who are overweight or obese with one risk factor, regardless of age. Additionally, those who have no risk factors should start screening at age 45.
    - Developer notes that the risk factors included in this measure bring together both the USPSTF and ADA risk factors.
    - Testing for prediabetes and risk for future diabetes in asymptomatic people should be considered in adults of any age who are overweight or obese (BMI ≥25kg/mg or ≥23kg/m2 in

Asian Americans) and who have one or more additional risk factors for diabetes. (ADA, 2018) (B Recommendation)
    - Grade B recommendation means: "The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial."
- The Committee questioned the fact that the measure does not have an age upper limit, noting the USPSTF guidance related to screening for diabetes for patients with high BMI ages between 40-70.
- Developer provided a summary of the literature related to gaps in care. Developer states that their review of the literature suggests that the uninsured are less likely to be screened; Hispanics and black people are also more likely to be screened than white people.
- The Committee agreed that gap that exists based on the literature despite the lack actual data of patient care.

2. Scientific Acceptability of Measure Properties: <u>The measure did not pass the Scientific Acceptability criteria</u>

(2a. Reliability precise specifications, testing; 2b. Validity testing, threats to validity

2a. Reliability: **H-1; M-16; L-5; I-0** 2b. Validity: **H-0; M-8; L-9; I-1**

<u>Rationale</u>:

- Developer used same testing for both data element reliability and validity.
- Developer performed data element reliability/validity testing at two facilities on two common EHR systems, Epic and Cerner.
    - Test Site #1: An ambulatory facility in South Carolina, part of a larger health system comprised of 8 inpatient hospitals and more than 100 outpatient facilities. This facility uses Epic EHR in their facility.
    - Test Site #2: An ambulatory facility in South Carolina, part of a larger system comprised of a 1,600+ bed comprehensive integrated health system, serving 1 million patients. This facility uses Cerner EHR in their facility.
- The feasibility assessment indicated the following data elements had issues in the accuracy domain indicating that these data elements may not be correct:
    - "Laboratory Test, Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (in Cerner and Epic) (measure developer noted that Fasting status of glucose testing is not captured in discrete fields in either EHR, however capturing A1C testing is feasible. To test for prediabetes, fasting plasma glucose, 2-h plasma glucose during 75-g oral glucose tolerance test, and A1C are equally appropriate)
    - "Intervention Order: Comfort Measures" (in Cerner) (measure developer noted that Comfort Care as an exclusion is standard in in NQF endocrine registry measures and it is expected that EMR developers to create a distinct field to collect this data in the future)
    - "Laboratory Test, Not Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Glucose [Mass/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner)
    - "Laboratory Test, Not Performed: Glucose [Moles/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner)
    - "Laboratory Test, Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
- Data element reliability/validity testing was conducted utilizing Parallel Forms Reliability Testing methodology to determine if data elements found through electronic data pulls could be confirmed by manual abstraction of the same data elements.
    - Verification of the data elements was obtained through automated data search strategies against a reference strategy (considered the gold standard) for obtaining the data elements.

3569e Prediabetes: Screening for Abnormal Blood Glucose

- Manual review of the data elements was used as the reference strategy against which automated data search and extraction strategies were evaluated.
- Interrater reliability (crude agreement and Cohen's Kappa) was used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record.

- Committee members raised concerns that this measure was only tested in two electronic health record (EHR) systems and was not tested with an EHR system less robust than Epic or Cerner.
- The Committee noted that the accuracy results were not clear and that there may be poorer results in smaller EHR systems.
- The Committee noted that several of the data elements had accuracy issues and could present challenges with acquiring data across different providers.
- The Committee expressed particular concern that there were no upper limits for age on this measure.

3. Feasibility: H-0; M-5; L-14; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- During the discussion of feasibility, the Committee raised concerns about the lack of fasting glucose being listed as such in the EMR and the fact that that comfort measures are not necessarily standard.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-17; No Pass-3** *4b. Usability:* **H-0; M-18; L-1; I-1**

Rationale:

- The Committee did not express any concerns with use and usability.

5. Related and Competing Measures

- No related or competing measures noted.

6. Standing Committee Recommendation for Endorsement: N/A

7. Public and Member Comment

- Preferred "abnormal blood glucose" over "prediabetes" which is suggested to be a risk factor rather than a disease.
- Some commenters opposed the missing upper age limit (40-70 years) included in AAFP and USPSTF guidelines.
- One commenter suggested that confirmation of results should be included in this measure.
- Several commenters had concerns with data capture, such as fasting glucose or exclusions not in EHR distinct field, and that the measure was only tested in EPIC and Cerner.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-X; No-X (November 17, 2020: [Endorsed or Not Endorsed])

The CSAC upheld [or did not uphold] the Standing Committee's decision to recommend the measure for endorsement.

9. Appeals

3570e Intervention for Prediabetes

Submission | Specifications

3570e Intervention for Prediabetes

Description: Percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention*

Numerator Statement: Patients who were provided an intervention*

*Intervention must include one of the following: referral to a CDC-recognized diabetes prevention program; referral to medical nutrition therapy with a registered dietician; prescription of metformin.

Denominator Statement: All patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period

**Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol).

Exclusions: Denominator Exclusions:

Exclude patients who are pregnant.

Exclude patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY])

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Electronic Health Records

Measure Steward: American Medical Association

STANDING COMMITTEE MEETING 06/25/2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-0; M-16; L-2; I-2**; 1b. Performance Gap: **H-2; M-16; L-1; I-1**

<u>Rationale</u>:

- Developer cites evidence primarily found in guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA).
    - USPSTF grade B recommendation states that adults aged 40 to 70 years of age who are overweight or obese should be screened for abnormal blood glucose as part of cardiovascular risk assessment.
    - Referral of patients with abnormal blood glucose to intensive behavioral counseling interventions is also recommended by the USPSTF (grade B recommendation).
    - The ADA recommends an intensive behavioral lifestyle intervention program modeled on the Diabetes Prevention Program for prediabetes patients (grade A recommendation).
    - The developer cites ADA's grade A recommendation on Metformin therapy for preventing type 2 diabetes in individuals with prediabetes (<60 years, BMI ≥35 kg/m2 and women with prior gestational diabetes mellitus).
    - An individualized medical nutrition therapy is recommended by ADA for all with type 1 or type 2 diabetes or gestational diabetes mellitus (grade A recommendation).
- The Committee noted that this measure could be specified as an outcome measure but acknowledged that providers may not yet have the processes in place to achieve outcomes.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

(2a. Reliability precise specifications, testing; 2b. Validity testing, threats to validity

2a. Reliability: **H-0; M-16; L-3; I-0**; 2b. Validity: **H-0; M-13; L-3; I-3**

<u>Rationale</u>:

- Developer used same testing for both data element reliability and validity.

3570e Intervention for Prediabetes

- Developer performed data element reliability/validity testing at two facilities on two common EHR systems.
    - Test Site #1: An ambulatory facility in South Carolina, part of a larger health system comprised of 8 inpatient hospitals and more than 100 outpatient facilities. This facility uses Epic EHR in their facility.
    - Test Site #2: An ambulatory facility in South Carolina, part of a larger system comprised of a 1,600+ bed comprehensive integrated health system, serving 1 million patients. This facility uses Cerner EHR in their facility.
- Submission includes simulated data set results demonstrating unit testing covering 100% of the measure logic.
- The feasibility assessment indicated the following data elements had issues in the accuracy domain indicating that these data elements may not be correct:
    - "Laboratory Test, Performed: Glucose [Mass/volume] in Serum or Plasma --2 hours post 75 g glucose PO" Measure developer noted that fasting status of glucose testing is not captured in discrete fields in either EHR, however capturing A1C testing is feasible. To test for pre-diabetes, fasting plasma glucose, 2-h plasma glucose during 75-g oral glucose tolerance test, and A1C are equally appropriate (in Cerner and Epic)
    - "Laboratory Test, Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Intervention, Performed: Referral to Dietitian" (measure developer noted that It is expected that EMR developers to create a distinct field to collect this data in the future.) (in Cerner and Epic)
    - "Intervention, Performed: Referral to Diabetes Prevention Program" (measure developer noted that It is expected that EMR developers to create a distinct field to collect this data in the future) (in Cerner and Epic)
    - "Intervention, Not Performed: Referral to Diabetes Prevention Program" (in Cerner and Epic)
    - "Intervention, Not Performed: Referral to Dietitian" (in Cerner and Epic)
    - "Medication, Not Ordered: Metformin" (in Cerner)
    - "Diagnosis: Limited Life Expectancy" (in Cerner)
    - "Encounter, Performed: Nursing Facility Visit" (in Cerner)
- Data element reliability/validity testing was conducted utilizing Parallel Forms Reliability Testing methodology to determine if data elements found through electronic data pulls could be confirmed by manual abstraction of the same data elements.
    - Verification of the data elements was obtained through automated data search strategies against a reference strategy (considered the gold standard) for obtaining the data elements.
    - Manual review of the data elements was used as the reference strategy against which automated data search and extraction strategies were evaluated.
    - Interrater reliability (crude agreement and Cohen's Kappa) was used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record.
- In terms of reliability, the Committee raised concerns around the sampling methodology. The Committee noted that convenience sampling did not necessarily indicate systematic bias.
- The Committee expressed concern that the measure may not have had all data elements tested and that the eCQM feasibility scorecard assessment indicated the many data elements had issues in the accuracy domain, indicating that these data elements may not be accurately captured.

3. Feasibility: H-0; M-5; L-15; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- The Committee raised concerns that the fields needed to collect this measure are not present in the EHR.

3570e Intervention for Prediabetes

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-18; No Pass-0** *4b. Usability:* **H-0; M-10; L-6; I-2**

Rationale:

- For usability, the Committee noted that there are potential issues with lack of discrete fields to document the referral and patient lacking access to a diabetes prevention program because their insurance doesn't cover the services.

5. Related and Competing Measures

- No related or competing measures noted.

6. Standing Committee Recommendation for Endorsement: **Yes-5; No-13**

The developer submitted a reconsideration request for this measure. The Standing Committee voted to not reconsider this measure.

7. Public and Member Comment

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-X; No-X (November 17, 2020: [Endorsed or Not Endorsed])

The CSAC upheld [or did not uphold] the Standing Committee's decision to recommend the measure for endorsement.

9. Appeals

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes

Submission | Specifications

Description: Percentage of patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the previous year who have a blood glucose test performed in the one-year measurement period

Numerator Statement: Patients who had a blood glucose test performed

*Retesting for abnormal blood glucose may include using a fasting plasma glucose, 2-h plasma glucose during a 75g oral glucose tolerance test, or A1C.

Denominator Statement: All patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the year prior to the one-year measurement period

**Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol).

Exclusions: Denominator Exclusions:

Exclude patients who are pregnant.

Exclude patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY]).

Exclude patients in palliative care/hospice

Adjustment/Stratification: No risk adjustment or risk stratification

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Outpatient Services

Type of Measure: Process

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes

Data Source: Electronic Health Records

Measure Steward: American Medical Association

STANDING COMMITTEE MEETING 06/25/2020

1. Importance to Measure and Report: <u>The measure did not pass the Importance criteria</u>

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **H-0; M-7; L-7; I-4;** 1b. Performance Gap: **H-1; M-10; L-3; I-3;** Evidence Exception: **Yes-9; No-8**

<u>Rationale</u>:

- Developer cites evidence found in guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA).
    - At least annual monitoring for the development of diabetes in those with prediabetes is suggested. (ADA, 2018) (E Recommendation)
    - Developer provides evidence of disease prevalence and systematic misses of opportunities to intervene by clinicians.
    - Developer does not provide studies that offer evidence that annual monitoring is associated with positive outcomes.
- The Committee noted that there is a lack of evidence to support this measure.
- The Committee raised the concern that the quality measurement enterprise generally has sufficient process measures and not enough outcome measures.
- The Committee observed the developer's review of the literature that suggests a gap in care, noting that the United States has 84 million adults with prediabetes, that 9 out of 10 patients who have prediabetes are not aware, and that missed opportunities among primary care providers in diagnosing and managing patients with prediabetes represent a gap in care.

2. Scientific Acceptability of Measure Properties: <u>The measure did not achieve consensus on the Scientific Acceptability criteria</u>

(2a. Reliability precise specifications, testing; 2b. Validity testing, threats to validity

2a. Reliability: **H-0; M-12; L-5; I-0**; 2b. Validity: **H-0; M-9; L-7; I-1**

<u>Rationale</u>:

- Developer used same testing for both data element reliability and validity.
- Developer performed data element reliability/validity testing at two facilities on two common EHR systems.
    - Test Site #1: An ambulatory facility in South Carolina, part of a larger health system comprised of 8 inpatient hospitals and more than 100 outpatient facilities. This facility uses Epic EHR in their facility.
    - Test Site #2: An ambulatory facility in South Carolina, part of a larger system comprised of a 1,600+ bed comprehensive integrated health system, serving 1 million patients. This facility uses Cerner EHR in their facility.
- Submission includes simulated data set results demonstrating unit testing covering 100% of the measure logic.
- The feasibility assessment indicated the following data elements had issues in the accuracy domain indicating that these data elements may not be correct:
    - Laboratory Test, Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (measure developer noted that fasting status of glucose testing is not captured in discrete fields in either EHR, however capturing A1C testing is feasible. To test for prediabetes, fasting plasma glucose, 2-h plasma glucose during 75-g oral glucose tolerance test, and A1C are equally appropriate. (in Cerner and Epic)
    - "Laboratory Test, Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)
    - "Laboratory Test, Not Performed: Fasting glucose [Moles/volume] in Serum or Plasma" (in Cerner and Epic)

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes
○ "Laboratory Test, Not Performed: Glucose [Moles/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner) ○ "Laboratory Test, Not Performed: Fasting Plasma Glucose Lab Test Mass Per Volume" (in Cerner and Epic) ○ "Laboratory Test, Not Performed: Glucose [Mass/volume] in Serum or Plasma --2 hours post 75 g glucose PO" (in Cerner) ○ "Intervention, Order: Comfort Measures" using "Comfort Measures (2.16.840.1.113883.17.4077.3.2030)" (measure developer noted that Comfort Care as an exclusion is standard in NQF endocrine registry measures and it is expected that EMR developers to create a distinct field to collect this data in the future) (in Cerner) • Data element reliability/validity testing was conducted utilizing Parallel Forms Reliability Testing methodology to determine if data elements found through electronic data pulls could be confirmed by manual abstraction of the same data elements. ○ Verification of the data elements was obtained through automated data search strategies against a reference strategy (considered the gold standard) for obtaining the data elements. ○ Manual review of the data elements was used as the reference strategy against which automated data search and extraction strategies were evaluated. ○ Interrater reliability (crude agreement and Cohen's Kappa) was used to assess the reliability of the measure based on results from two independent reviewers trained in the same way reviewing the same patient record. • The Committee did not reach consensus on validity, but noted that the measure had concerns associated with the feasibility scorecard in that the accuracy of the data elements was questionable.
3. Feasibility: H-0; M-7; L-9; I-0 *(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)* Rationale: • In the review of the measure's feasibility, the Committee was concerned that reporting the measure may be challenging since the accuracy of the data elements was not clear.
4. Use and Usability *4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)* 4a. Use: **Pass-15; No Pass-2** *4b. Usability:* **H-0; M-7; L-6; I-3** Rationale: • The Committee noted that the measure has not been implemented, but the developer has the intention of submitting the measure to CMS for the MIPS program
5. Related and Competing Measures • No related or competing measures noted.
6. Standing Committee Recommendation for Endorsement: **N/A** The Standing Committee did not vote on an endorsement recommendation because the measure did not pass the scientific acceptability criteria.
7. Public and Member Comment • One commentator noted, "there is limited evidence on the best rescreening intervals for adults with normal results; however, screening every 3 years is a reasonable option." In contrast, this measure requires re-testing at least annually. • In addition, the exclusions for this measure are different from the others. Comfort care is not included in this measure.

3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes
• Other comment agreed that retesting is needed but that the testing should include a variety of tests, a specific timeframe, coverage by insurance, and ease of access to tests.
8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Yes-X; No-X (November 17, 2020: [Endorsed or Not Endorsed]) The CSAC upheld [or did not uphold] the Standing Committee's decision to recommend the measure for endorsement.
9. Appeals

Appendix B: Primary Care and Chronic Illness Portfolio—Use in Federal Programs[1]

NQF #	Title	Federal Programs: Implemented or Finalized as of February 22, 2019
0046	Screening for Osteoporosis for Women 65-85 Years of Age	Merit-Based Incentive Payment System (MIPS) Program (Finalized)
0047	Asthma: Pharmacologic Therapy for Persistent Asthma	None
0053	Osteoporosis Management in Women Who Had a Fracture	MIPS Program (Finalized), Medicare Part C Star Rating (Implemented)
0054	Disease-Modifying Anti-Rheumatic Drug Therapy for Rheumatoid Arthritis (ART)	None
0055	Comprehensive Diabetes Care: Eye Exam (retinal) performed	MIPS Program (Finalized), Qualified Health Plan (QHP) Quality Rating System (QRS) (Implemented)
0056	Comprehensive Diabetes Care: Foot Exam	MIPS Program (Finalized)
0057	Comprehensive Diabetes Care: Hemoglobin A1c (HbA1c) Testing	Medicaid (Implemented), Qualified Health Plan (QHP) Quality Rating System (QRS) (Implemented)
0058	Avoidance of Antibiotic Treatment in Adults With Acute Bronchitis (AAB)	Medicare Physician Quality Reporting System, MIPS Program (Finalized), Qualified Health Plan (QHP) Quality Rating System (QRS) (Implemented)
0059	Comprehensive Diabetes Care: Hemoglobin A1c (HbA1c) Poor Control (>9.0%)	Medicaid (Implemented), Medicare Shared Savings Program (Implemented), MIPS Program (Finalized)
0061	Comprehensive Diabetes Care: Blood Pressure Control (<140/90 mm Hg)	None
0062	Comprehensive Diabetes Care: Medical Attention for Nephropathy	MIPS Program (Finalized), Qualified Health Plan (QHP) Quality Rating System (QRS) (Implemented)
0086	Primary Open-Angle Glaucoma (POAG): Optic Nerve Evaluation	MIPS Program (Finalized)
0087	Age-Related Macular Degeneration: Dilated Macular Examination	MIPS Program (Finalized)
0088	Diabetic Retinopathy: Documentation of Presence or Absence of Macular Edema and Level of Severity of Retinopathy	None

[1] Per CMS Measures Inventory Tool as of 07/09/2020

NQF #	Title	Federal Programs: Implemented or Finalized as of February 22, 2019
0089	Diabetic Retinopathy: Communication with the Physician Managing Ongoing Diabetes Care	Merit-Based Incentive Payment System (MIPS) Program (Finalized)
0091	COPD: Spirometry Evaluation	MIPS Program (Finalized)
0405	HIV/AIDS: Pneumocystis jiroveci pneumonia (PCP) Prophylaxis	MIPS Program (Finalized)
0409	HIV/AIDS: Sexually Transmitted Diseases – Screening for Chlamydia, Gonorrhea, and Syphilis	MIPS Program (Finalized)
0416	Diabetic Foot & Ankle Care, Ulcer Prevention – Evaluation of Footwear	MIPS Program (Finalized)
0417	Diabetic Foot & Ankle Care, Peripheral Neuropathy – Neurological Evaluation	MIPS Program (Finalized)
0541	Proportion of Days Covered (PDC): 3 Rates by Therapeutic Category	Qualified Health Plan (QHP) Quality Rating System (QRS) (Implemented)
0563	Primary Open-Angle Glaucoma: Reduction of Intraocular Pressure by 15% or Documentation of a Plan of Care	MIPS Program (Finalized)
0566	Age-Related Macular Degeneration (AMD): Counseling on Antioxidant Supplement	None
0575	Comprehensive Diabetes Care: Hemoglobin A1c (HbA1c) Control (<8.0%)	Qualified Health Plan (QHP) Quality Rating System (QRS) (Implemented)
0577	Use of Spirometry Testing in the Assessment and Diagnosis of COPD	None
0653	Acute Otitis Externa: Topical Therapy	MIPS Program (Finalized)
0654	Acute Otitis Externa: Systemic Antimicrobial Therapy – Avoidance of Inappropriate Use	MIPS Program (Finalized)
0655	Otitis Media with Effusion: Antihistamines or decongestants – Avoidance of inappropriate use	None
0657	Otitis Media with Effusion: Systemic antimicrobials – Avoidance of inappropriate use	MIPS Program (Implemented)
0729	Optimal Diabetes Care	None
1800	Asthma Medication Ratio	Medicaid (Implemented)
2079	HIV medical visit frequency	MIPS Program (Finalized)

NQF #	Title	Federal Programs: Implemented or Finalized as of February 22, 2019
2080	Gap in HIV medical visits	None
2082	HIV viral load suppression	Medicaid (Implemented), MIPS Program (Finalized)
2083	Prescription of HIV Antiretroviral Therapy	None
2522e	Rheumatoid Arthritis: Tuberculosis Screening	None
2523e	Rheumatoid Arthritis: Assessment of Disease Activity	None
2524e	Rheumatoid Arthritis: Functional Status Assessment	None
2525e	Rheumatoid Arthritis: Disease Modifying Anti-Rheumatic Drug (DMARD) Therapy	None
2549e	Gout: Serum Urate Target	None
2550e	Gout: ULT Therapy (Recommended for eMeasure Trial Approval)	None
2811e	Acute Otitis Media - Appropriate First-Line Antibiotics	None
2856	Pharmacotherapy Management of COPD Exacerbation	None
3086	Population Level HIV Viral Load Suppression	None
3209e	HIV medical visit frequency	None
3210e	HIV viral load suppression	None
3211e	Prescription of HIV Antiretroviral Therapy	None
3569e	Prediabetes: Screening for Abnormal Blood Glucose	None
3570e	Intervention for Prediabetes	None
3571e	Retesting of Abnormal Blood Glucose in Patients with Prediabetes	None

Appendix C: Primary Care and Chronic Illness Standing Committee and NQF Staff

STANDING COMMITTEE

Dale Bratzler, DO, MPH (Co-Chair)
University of Oklahoma Health Sciences Center-College of Public Health
Oklahoma City, OK

Adam Thompson, BA (Co-Chair)
Kennedy Health Alliance
Berlin, NJ

Robert A. Bailey, MD
Janssen Scientific Affairs, LLC
Titusville, NJ

Lindsay Botsford, MD, MBA, MBA/FAAFP
Physicians at Sugar Creek
Sugar Land, TX

Kathleen Brady, MD, MSCE
Philadelphia Department of Public Health
Philadelphia, PA

Roger Chou, MD
Oregon Health & Science University
Cochrane Back and Neck Group
Portland, OR

James M. Daniels, MD, MPH, RMSK, FAAFP, FACOEM, FACPM
Illinois University
Quincy, Illinois

Kim Elliott, PhD
Health Services Advisory Group, Inc.
Phoenix, AZ

Donald Goldmann, MD
Institute for Healthcare Improvement
Boston, Massachusetts

V. Katherine Gray, PhD
Sage Health Management Solutions, Inc.
Minneapolis, Minnesota

Faith Green, MSN, RN, CPHQ, CPC-A
Humana
Louisville, Kentucky

Daniel Greninger, MD
The Permanente Medical Group
Antioch, California

Stephen Grossbart, PhD
Health Catalyst
Salt Lake City, UT

James Michell Harris
Children's Hospital Association
Washington, DC

Starlin Haydon-Greatting, MS, BS, Pharm, FAPhA
Illinois Pharmacists Association
Springfield, Illinois

Ann Kearns, MD, PhD
Mayo Clinic
Rochester, MN

Grace Lee, MD
Virginia Mason Medical Center
Seattle, WA

Jason Matuszak, MD, FAAFP
The Sports Concussion Center
Excelsior Orthopaedics
Amherst, NY

Anna McCollister
Galileo Analytics
Washington, DC

Janice Miller, DNP, CRNP, AGPCNP-BC, CDE
Thomas Jefferson University
Philadelphia, PA

Crystal Riley, PharmD, MHA, MBA, CPHQ, CHPIT
Baxter Healthcare Corporation
Washington, DC

Rishi Singh, MD
Cleveland Clinic

Cleveland, Ohio

Steven Strode, M.D., M.Ed., M.P.H., FAAFP
eDocAmerica
Sherwood, Arkansas

NQF STAFF

Sheri Winsper, RN, MSN, MSHA
Senior Vice President, Quality Measurement

Apryl Clark, MHSA
Acting Vice President, Quality Measurement

Sai Ma, MPA, PhD
Managing Director/Senior Technical Expert, Quality Measurement

Samuel Stolpe, PharmD, MPH
Senior Director

Yemi Kidane, PMP
Project Manager

Erin Buchanan, MPH
Manager

Isaac Sakyi, MSGH
Analyst

Appendix D: Measure Specifications

	3569e Prediabetes: Screening for Abnormal Blood Glucose
Steward	American Medical Association
Description	Percentage of patients aged 40 years and older with a BMI greater than or equal to 25 who are seen for at least two office visits or at least one preventive visit during the 12-month period who were screened for abnormal blood glucose at least once in the last 3 years
Type	Process
Data Source	Electronic Health Records Measure data elements will be collected through health care organization electronic health record query, electronic health data queries
Level	Clinician : Group/Practice, Clinician : Individual
Setting	Outpatient Services
Numerator Statement	*Screening for abnormal blood glucose may include using a fasting plasma glucose, 2-h plasma glucose during a 75g oral glucose tolerance test, or A1C.
Numerator Details	exists "A1c Test Performed Within Past 3 Years" or exists "Fasting Plasma Glucose Test Performed Within Past 3 Years" or exists "Two Hour Plasma Glucose During 75 Gram Oral Glucose Tolerance Test Performed Within Past 3 Years" See additional code sets and materials in attachments
Denominator Statement	All patients aged 43 years and older with a BMI greater than or equal to 25 seen for at least two office visits or at least one preventive visit during the 12-month measurement period
Denominator Details	Denominator "Initial Population" and exists (["Patient Characteristic Birthdate": "Birth date"] BirthDate where Global."CalendarAgeInYearsAt" (BirthDate.birthDatetime, start of "Measurement Period") >= 43) and "Highest BMI Documented During Measurement Period is Greater Than or Equal to 25" See attachment in human readable file in S.2a
Exclusions	Denominator Exclusions "Patient is Pregnant at Encounter" or "Patient Has Active Diabetes Diagnosis at Encounter" or "Hospice During Measurement Period" or "Palliative Care During Measurement Period" or "Comfort Measures During Measurement Period"
Exclusion details	See attachment in human readable file in S.2a
Risk Adjustment	No risk adjustment or risk stratification
Stratification	N/A
Type Score	Rate/proportion better quality = higher score
Algorithm	See attachment in human readable file in S.2a 151659
Copyright / Disclaimer	© 2018 American Medical Association. All Rights Reserved.? Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of

	3569e Prediabetes: Screening for Abnormal Blood Glucose
	these code sets. The AMA disclaims all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications. CPT® contained in the Measures specifications is copyright 2004-2017 American Medical Association. LOINC® copyright 2004-2017 Regenstrief Institute, Inc. SNOMED CLINICAL TERMS (SNOMED CT®) copyright 2004-2017 The International Health Terminology Standards Development Organisation (IHTSDO). ICD-10 is copyright 2017 World Health Organization. All Rights Reserved.

	3570e Intervention for Prediabetes
Steward	American Medical Association
Description	Percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention*
Type	Process
Data Source	Electronic Health Records Measure data elements will be collected through health care organization electronic health record query, electronic health data queries
Level	Clinician : Group/Practice, Clinician : Individual
Setting	Outpatient Services
Numerator Statement	Patients who were provided an intervention* *Intervention must include one of the following: referral to a CDC-recognized diabetes prevention program; referral to medical nutrition therapy with a registered dietician; prescription of metformin.
Numerator Details	Please see attachment in S.2a for all information required to calculate numerator
Denominator Statement	All patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period **Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol).
Denominator Details	Please see attachment in S.2a for all information required to calculate denominator
Exclusions	Denominator Exclusions: Exclude patients who are pregnant. Exclude patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY])
Exclusion details	Please see attachment in S.2a for all information required to calculate denominator exclusions
Risk Adjustment	No risk adjustment or risk stratification
Stratification	N/A
Type Score	Rate/proportion better quality = higher score
Algorithm	Please see attachment in S.2a for all information required to calculate measure 151659

	3570e Intervention for Prediabetes
Copyright / Disclaimer	

	3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes
Steward	American Medical Association
Description	Percentage of patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the previous year who have a blood glucose test performed in the one-year measurement period
Type	Process
Data Source	Electronic Health Records Measure data elements will be collected through health care organization electronic health record query, electronic health data queries.
Level	Clinician : Group/Practice, Clinician : Individual
Setting	Outpatient Services
Numerator Statement	Patients who had a blood glucose test performed *Retesting for abnormal blood glucose may include using a fasting plasma glucose, 2-h plasma glucose during a 75g oral glucose tolerance test, or A1C.
Numerator Details	See attached file in S.2a and S.2b for information to calculate the numerator
Denominator Statement	All patients aged 18 years and older who had an abnormal fasting plasma glucose, oral glucose tolerance test, or hemoglobin A1c result in the range of prediabetes in the year prior to the one-year measurement period **Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol).
Denominator Details	See attached file in S.2a and S.2b for information to calculate the denominator
Exclusions	Denominator Exclusions: Exclude patients who are pregnant. Exclude patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY]). Exclude patients in palliative care/hospice
Exclusion details	See attached file in S.2a and S.2b for information to calculate the exclusions
Risk Adjustment	No risk adjustment or risk stratification
Stratification	n/a

	3571e Retesting of Abnormal Blood Glucose in Patients with Prediabetes
Type Score	Rate/proportion better quality = higher score
Algorithm	See attached file in S.2a for information to calculate the measure logic 151659
Copyright / Disclaimer	

Appendix E: Related and Competing Measures

No related or competing measures were identified.

Appendix F: Pre-Evaluation Comments

No comments were received.

National Quality Forum
1099 14th Street NW, Suite 500
Washington, DC 20005
http://www.qualityforum.org