NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Click to go to the link. ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 3455

Corresponding Measures:

De.2. Measure Title: Timely Follow-Up After Acute Exacerbations of Chronic Conditions

Co.1.1. Measure Steward: IMPAQ International

De.3. Brief Description of Measure: The percentage of issuer-product-level acute events requiring either an emergency department (ED) visit or hospitalization for one of the following 6 chronic conditions: hypertension, asthma, heart failure (HF), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), or diabetes mellitus (Type I or Type II), where follow-up was received within the timeframe recommended by clinical practice guidelines in a non-emergency outpatient setting.

1b.1. Developer Rationale: Follow-up care is an element of care coordination, which has recently been highlighted as a priority area within the CMS Quality Strategy goal of Eliminating Disparities. (1) Follow-up care allows providers to perform a number of important activities, including ensuring that patients understand and are adhering to their medication regimen, monitoring them for adverse events, and educating them to recognize warning signs. (2) To date, there are scant published studies or data, especially within the US, that provide aggregate rates of timely follow-up care after ED or hospitalization for the conditions considered in this measure, making it imperative to fill in the gap in the healthcare community.

The body of evidence described below suggests, broadly and for the specific conditions included in this measure, that failing to obtain timely follow-up care is empirically linked to higher readmission rates, which are known to be associated with increased costs and decreased patient satisfaction and are widely considered an indicator of sub-optimal patient care. Furthermore, in the attached evidence tables we provide clinical guidelines for each condition, recommending that patients receive follow-up care after inpatient or ED discharge within the respective timeframe set for each chronic condition in the measure.

The goal of this measure is to improve the quality of care provided and improve patient outcomes by incentivizing health plans (insurance products) to ensure patients receive appropriate follow-up care following acute exacerbations of chronic conditions. To achieve this goal, the measure will identify health plans that have significantly lower rates of appropriate follow-up visits for acute conditions, relative to other health plans with the same acute conditions for similar patient populations. In doing so, this measure will prompt health plans to carefully evaluate care processes and implement quality improvement strategies. (3) Ultimately, this measure will provide an opportunity for health plans to become aware of and to improve rates of appropriate follow-up visits following acute events leading to decreased morbidity and mortality for patients with any of the 6 conditions covered by this measure.

By incentivizing health plans (insurance products) to improve follow-up rates for these conditions, this measure will improve outcomes as demonstrated in the following logic model:

1. Patient presents to the hospital/ED with an acute exacerbation of 1 of the 6 conditions included in the measure.
2. Patient is treated (either admitted to the hospital or treated only in the emergency department) and discharged to the community.
3. Health plan (insurance products) encourages follow-up visit/care through strategies such as incentives to providers,* reminders to patients, providing data/reports and continuing education to providers, etc.
4. Patient receives follow-up visit/care based upon evidence-based clinical guidelines, conditions are appropriately managed, with improved patient health and function.
5. ED and hospital cost and utilization are reduced by preventing avoidable readmissions.

* Note: If a health plan provides bonuses or other financial incentives to providers based on quality performance, an unintended consequence of holding plans accountable could be that providers are financially penalized if the plan performs poorly. However, providers, in this case, have every incentive to help patients receive timely follow-up care, which could render the net effects being positive.

This logic model is made available in visual format as an appendix to the evidence form.

In general, the evidence demonstrates the presence of a link between timely follow-up care after discharge from the hospital/ED and lower rates of readmission. (4, 5) Specifically, the evidence also supports the link between follow-up care and improved outcomes for each of the conditions included in this measure, as described below.

• Asthma: A systematic review states that evidence from randomized controlled trials (RCTs) and non-RCTs suggests that timely follow-up with specialists reduces subsequent asthma exacerbations, as well as fewer symptoms and improved quality of life. (6) A population-based study on 7,829 patients with asthma or COPD found that follow-up visits within 30 days of discharge were associated with a significant reduction in the 90-day risk of an emergency readmission (RR=.79). (7) In a study of Medicare beneficiaries with COPD, 10% of whom also had asthma as a primary diagnosis code, the 30-day rate of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. Within the study population, patients who had a timely follow-up visit had a significantly reduced risk of ED visit (HR=.86) and readmission (HR=.91). (8)

• Heart Failure: Outpatient follow-up from a cardiology or general medicine provider within 7 days for a heart failure patient after hospitalization is associated with a lower chance of 30-day readmission (OR=.81). (9) Delayed outpatient follow-up after myocardial infarction has been associated with worse short-term and long-term patient medication adherence. Logically, this will add to morbidity and very likely both mortality and subsequent cardiovascular events/subsequent hospital admissions. (10) In the United States, the Hospital to Home (H2H) program, a national quality improvement initiative, also recommends that all heart failure and myocardial infarction patients have a follow-up appointment or cardiac rehab referral scheduled within 7 days of discharge. (11) Furthermore, the H2H program implemented a structured improvement project called the "See You in 7" initiative at 10 hospitals in Southeast Michigan, which included follow-up within a week of discharge as a core concept. After one year of participation, the adjusted 30-day readmission rates at collaborating hospitals decreased compared to non-participating hospitals (2.6% decrease vs. 0.6% decrease). (12) Heart failure and myocardial infarction patients in Taiwan were found to have a lower risk of 30-day readmission if they received outpatient visit with a physician within 7 days of discharge (HR=.54). (13) In a study of 3,136 patients, those who received cardiovascular follow-up saw fewer ED (38% vs. 80%) and hospital (13% vs. 94%) readmissions for cardiovascular reasons within the year, and lower unadjusted mortality (7% vs. 2% at 30 days). (14) In a study of 30,136 patients, patients discharged from hospitals with higher rates of

follow-up within 7 days of discharge for heart failure had a lower risk of 30-day readmission (HR .91 between highest and lowest quartiles). (15)

While direct evidence on the relationship between timely follow-up and CAD and hypertension is extremely limited, both conditions are well recognized comorbidities and predictors of heart failure. (16, 17, 18) Logically, because a strong body of evidence links timely follow-up after heart failure to lower readmissions, timely follow-up for CAD and hypertension can also be reasoned to lower readmissions.

• COPD: In a study of Medicare beneficiaries with COPD, 10% of whom also had asthma as a primary diagnosis code, the 30-day rates of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. Within the study population, patients who had a timely follow-up visit had a significantly reduced risk of ED visit (HR=.86) and readmission (HR=.91). (8) A population-based study of 7829 patients with asthma or COPD found that follow-up visits within 30 days of discharge were associated with a significant reduction in the 90-day risk of an emergency readmission (RR=.79). (7)

• Diabetes: A review article suggests that better management of hyperglycemia in the ED, along with proper follow-up to ensure continued appropriate management, improves clinical outcomes and prevents readmission. (19) Another study found that sentinel ED visits in diabetic patients are a warning sign for future readmission for hyperglycemia, and recommended that clinicians "provide clear discharge instructions for follow-up and glucose management to prevent further hyperglycemic emergencies from occurring." (20)

The gap in care described above represents one reason these particular 6 conditions were selected for inclusion in the measure. These measures were also selected because they are ambulatory care-sensitive conditions that respond well to timely primary care, as reflected in broad body of evidence (described in the evidence form) linking timely follow-up for these conditions to improved health outcomes. Additionally, these conditions were selected based on the impact they have on patients and health systems, measured in both condition prevalence as well as the costs and resources associated with appropriate treatment. For example, the CDC finds that approximately 30 million Americans have diabetes, and complications were estimated to cost the US about $245 billion in 2012. (21) Nearly 27 million Americans are living with asthma, which is responsible for about 440,000 inpatient discharges, as well as 1.7 million ED visits and 11 million physician's office visits per year, with a total economic burden estimated at about $53 billion. (22, 23)

The prevalence of hypertension is even higher, with about 75 million adults (1 in 3) suffering from high blood pressure and an estimated $50 billion in economic burden. (24) About 15.7 million Americans suffer from COPD, with mortality rates as high as 62.8 per 100,000 in Kentucky and an overall economic burden of nearly $50 billion. (25, 26) About 15 million Americans suffer from CAD, and about 6 million suffer from heart failure. Heart disease remains the leading cause of death in the United States and represents an immense burden on the health system as a whole and on patients and their families individually. (27, 28)

While the prevalence of these 6 conditions is not uniformly high, they incur considerable cost to health systems as well as individual patients, and are conditions that represent degraded patient health and risk of mortality. The following statistics summarize incidence and cost data gathered from the 2014 Healthcare Cost and Utilization Project. Please note that ED cost data were not available.

Hypoglycemia

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.92
  - ED Rate per 100,000 persons: 8.06
  - Estimate of Mean Inpatient Costs: $7,177
- Any DX
  - Inpatient Rate per 100,000 persons: 12.55
  - ED Rate per 100,000 persons: 28.75

Hyperglycemia

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.16
  - ED Rate per 100,000 persons: 10.06
  - Estimate of Mean Inpatient Costs: $4,595
- Any DX
  - Inpatient Rate per 100,000 persons: 21.4
  - ED Rate per 100,000 persons: 509.26

Diabetic Ketoacidosis

- Principal DX
  - Inpatient Rate per 100,000 persons: 45.79
  - ED Rate per 100,000 persons: 48.77
  - Estimate of Mean Inpatient Costs: $7,280
- Any DX
  - Inpatient Rate for 100,000 persons: 59.75
  - ED Rate per 100,000 persons: 63.29

Acute Myocardial Infarction

- Principal DX
  - Inpatient Rate per 100,000 persons: 69.28
  - ED Rate per 100,000 persons: 64.31
  - Estimate of Mean Inpatient Costs: $21,380
- Any DX
  - Inpatient Rate per 100,000 persons: 95.79
  - ED Rate per 100,000 persons: 87.13

Angina

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.44
  - ED Rate per 100,000 persons: 0.59
  - Estimate of Mean Inpatient Costs: $7,693
- Any DX
  - Inpatient Rate per 100,000 persons: 1.61
  - ED Rate per 100,000 persons: 1.93

Acute Asthma Attack

- Principal DX
  - Inpatient Rate per 100,000 persons: 68.60
  - ED Rate per 100,000 persons: 462.98
  - Estimate of Mean Inpatient Costs: $6,555
- Any DX
  - Inpatient Rate per 100,000 persons: 471.03

- ED Rate per 100,000 persons: 1824.8

Hypotension

- Principal DX
  - Inpatient Rate per 100,000 persons: 8.25
  - ED Rate per 100,000 persons: 21.06
  - Estimate of Mean Inpatient Costs: $7,088
- Any DX
  - Inpatient Rate per 100,000 persons: 156.05
  - ED Rate per 100,000 persons: 143.97

Shortness of Breath

- Principal DX
  - Inpatient Rate per 100,000 persons: 1.04
  - ED Rate per 100,000 persons: 61.44
  - Estimate of Mean Inpatient Costs: $5,486
- Any DX
  - Inpatient Rate per 100,000 persons: 17.18
  - ED Rate per 100,000 persons: 386.76

Acute Pulmonary Edema

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.67
  - ED Rate per 100,000 persons: 0.88
  - Estimate of Mean Inpatient Costs: $9,011
- Any DX
  - Inpatient Rate per 100,000 persons: 5.51
  - ED Rate per 100,000 persons: 4.35

Peripheral Edema

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.7
  - ED Rate per 100,000 persons: 47.12
  - Estimate of Mean Inpatient Costs: $6,468
- Any DX
  - Inpatient Rate per 100,000 persons: 40.48
  - ED Rate per 100,000 persons: 128.07

Acute Left Ventricular Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.07
  - ED Rate per 100,000 persons: 0.08
  - Estimate of Mean Inpatient Costs: $8,839
- Any DX

- Inpatient Rate per 100,000 persons: 0.78
- ED Rate per 100,000 persons: 1.23

Pleural Effusion

- Principal DX
  - Inpatient Rate per 100,000 persons: 4.29
  - ED Rate per 100,000 persons: 7.31
  - Estimate of Mean Inpatient Costs: $12,562
- Any DX
  - Inpatient Rate per100,000 persons: 65.46
  - ED Rate per 100,000 persons: 60.93

Hypertension

- Principal DX
  - Inpatient Rate per 100,000 persons: 7.33
  - ED Rate per 100,000 persons: 138.68
  - Estimate of Mean Inpatient Costs: $5,699
- Any DX
  - Inpatient Rate per 100,000 persons: 1280.37
  - ED Rate per 100,000 persons: 3710.17

Chronic Kidney Disease and Systolic Heart Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 2.54
  - ED Rate per 100,000 persons: 6.81
  - Estimate of Mean Inpatient Costs: $19,507
- Any DX
  - Inpatient Rate per 100,000 persons: 366.59
  - ED Rate per 100,000 persons: 440.07

Pulmonary Congestion and Hypostasis

- Principal DX
  - Inpatient Rate per 100,000 persons: .63
  - ED Rate per 100,000 persons: 1.62
  - Estimate of Mean Inpatient Costs: $7,724
- Any DX
  - Inpatient Rate per 100,000 persons: 11.10
  - ED Rate per 100,000 persons: 11.68

Congestive Heart Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 13.67
  - ED Rate per 100,000 persons: 33.86
  - Estimate of Mean Inpatient Costs: $11,374

- Any DX
  - Inpatient Rate per 100,000 persons: 353.30
  - ED Rate per 100,000 persons: 462.81

COPD

- Principal DX
  - Inpatient Rate per 100,000 persons: 70.49
  - ED Rate per 100,000 persons: 407.92
  - Estimate of Mean Inpatient Costs: $8,019
- Any DX
  - Inpatient Rate per 100,000 persons: 467.86
  - ED Rate per 100,000 persons: 1119.54

CITATIONS:

1) Centers for Medicare and Medicaid Services. CMS Quality Strategy. 2016. Available from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Downloads/CMS-Quality-Strategy.pdf

2) California Healthcare Foundation. The Post-Hospital Follow-Up Visit: A Physician Checklist to Reduce Readmissions. October 2010. Available from https://www.chcf.org/wp-content/uploads/2017/12/PDF-PostHospitalFollowUpVisit.pdf.

3) Chassin M, et al. Accountability Measures – Using Measurement to Promote Quality Improvement. N Engl J Med 2010; 363:683-688. doi: 10.1056/NEJMsb1002320.

4) Misky GJ, et al. Post-hospitalization transitions: Examining the effects of timing of primary care provider follow-up. J Hosp Med. 2010 Sep;5(7):392-7. doi: 10.1002/jhm.666.

5) Shen E, et al. Association of a Dedicated Post–Hospital Discharge Follow-up Visit and 30-Day Readmission Risk in a Medicare Advantage Population. JAMA Intern Med. 2017;177(1):132-135. doi:10.1001/jamainternmed.2016.7061.

6) Schatz M, et al. Follow-Up After Acute Asthma Episodes: What Improves Future Outcomes? J Emerg Med. 2009 Aug;37(2 Suppl):S42-50. doi: 10.1016/j.jemermed.2009.06.109.

7) Sin D et al. The impact of follow-up physician visits on emergency readmissions for patients with asthma and chronic obstructive pulmonary disease: a population-based study. Am J Med. 2002 Feb 1;112(2):120-5. doi: 10.1016/S0002-9343(01)01079-8.

8) Sharma G, et al. Outpatient Follow-up Visit and 30-Day ED Visit and Readmission in Patients Hospitalized for Chronic Obstructive Pulmonary Disease. Arch Intern Med. 2010 Oct 11; 170(18): 1664–1670. doi: 10.1001/archinternmed.2010.345.

9) Lee K, et al. Post-Discharge Follow-up Characteristics Associated with 30-Day Readmission After Heart Failure Hospitalization. Med Care. 2016 Apr; 54(4): 365–372. doi: 10.1097/MLR.0000000000000492.

10) Faridi KF, et al. Timing of First Postdischarge Follow-up and Medication Adherence After Acute Myocardial Infarction. JAMA Cardiol. 2016 May 1;1(2):147-55. doi: 10.1001/jamacardio.2016.0001.

11) Hospital 2 Home: Improving Transitions from the Hospital to Community Settings. 2013. Slide Deck. Available from https://cvquality.acc.org/docs/default-source/h2h/1-h2h-overview-slides.pdf?sfvrsn=c6d58fbf_2.

12) Baker H, et al. Regional Hospital Collaboration and Outcomes in Medicare Heart Failure Patients: See You in 7. JACC Heart Fail. 2015 Oct;3(10):765-73. doi: 10.1016/j.jchf.2015.06.007.

13) Tung YC, et al. Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure. PLoS One. 2017 Jan 27;12(1):e0170061. doi: 10.1371/journal.pone.0170061.

14) Ezekowitz JA, et al. Impact of specialist follow-up in outpatients with congestive heart failure. CMAJ. 2005 Jan 18;172(2):189-94. doi: 10.1503/cmaj.1032017.

15) Hernandez AF, et al. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. JAMA. 2010 May 5;303(17):1716-22. doi: 10.1001/jama.2010.533.

16) Velagaleti R, et al. Heart Failure in the 21st Century: Is it a Coronary Artery Disease Problem or Hypertension Problem? Cardiol Clin. 2007 Nov; 25(4): 487–v. doi: 10.1016/j.ccl.2007.08.010.

17) Lala A, et al. The role of coronary artery disease in heart failure. Heart Fail Clin. 2014 Apr;10(2):353-65. doi: 10.1016/j.hfc.2013.10.002.

18) Messerli F, et al. The Transition from Hypertension to Heart Failure. JACC: Heart Failure Aug 2017, 5 (8) 543-551; DOI: 10.1016/j.jchf.2017.04.012.

19) Echouffo-Tcheugui J, et al. Management of Hyperglycemia and Diabetes in the ED. Curr Diab Rep (2017) 17: 56. https://doi.org/10.1007/s11892-017-0883-2.

20) Yan JW, Gushulak KM, Columbus MP, Hamelin AL, Wells GA, Stiell IG. Sentinel visits in ED patients with diabetes mellitus as a warning sign for hyperglycemic emergencies. CJEM. 2017 Jul 25:1-8. doi: 10.1017/cem.2017.338.

21) ADA. Economic Costs of Diabetes in the U.S. in 2012. Diabetes Care. 2013 Apr; 36(4): 1033–1046.

22) Asthma and Allergy Foundation of America. Costs of Asthma on Society. Available from http://www.aafa.org/page/cost-of-asthma-on-society.aspx.

23) Centers for Disease Control and Prevention. Most Recent Asthma Data. May 2018. Available from https://www.cdc.gov/asthma/most_recent_data.htm.

24) Centers for Disease Control and Prevention. High Blood Pressure. Available from https://www.cdc.gov/bloodpressure/index.htm.

25) Guarascio A, et al. The clinical and economic burden of chronic obstructive pulmonary disease in the USA. Clinicoecon Outcomes Res. 2013; 5: 235–245. Published online 2013 Jun 17. doi: 10.2147/CEOR.S34321.

26) Centers for Disease Control and Prevention. Chronic Obstructive Pulmonary Disease (COPD). Available from https://www.cdc.gov/copd/data.html.

27) Sanchis-Gomar F, et al A. Epidemiology of coronary heart disease and acute coronary syndrome. Ann Transl Med. 2016 Jul; 4(13): 256. doi: 10.21037/atm.2016.06.33.

28) Roger V. Epidemiology of Heart Failure. Circ Res. 2013 Aug 30; 113(6): 646–659. doi: 10.1161/CIRCRESAHA.113.300268.

S.4. Numerator Statement: The numerator is the sum of the issuer-product-level denominator events (Emergency Room [ED], observation hospital stay or inpatient hospital stay) for acute exacerbation of hypertension, asthma, heart failure (HF), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), or diabetes where follow-up was received within the timeframe recommended by clinical practice guidelines, as detailed below:

- Hypertension: Within 7 days of the date of discharge
- Asthma: Within 14 days of the date of discharge
- HF: Within 14 days of the date of discharge
- CAD: Within 14 days of the date of discharge
- COPD: Within 30 days of the date of discharge

- Diabetes: Within 30 days of the date of discharge

S.6. Denominator Statement: The denominator is the sum of the plan-product-level acute exacerbations that require either an ED visit, observation stay, or inpatient stay (i.e., acute events) for any of the six conditions listed above (hypertension, asthma, HF, CAD, COPD, or diabetes).

S.8. Denominator Exclusions: The measure excludes events with:

1. Subsequent acute events that occur two days after the prior discharge, but still during the follow-up interval of the prior event for the same reason. To prevent double-counting, only the first acute event will be included in the denominator.
2. Acute events after which the patient does not have continuous enrollment for 30 days in the same product.
3. Acute events where the discharge status of the last claim is not "to community" ("Left against medical advice" is not a discharge to community.)
4. Acute events for which the calendar year ends before the follow-up window ends (e.g., acute asthma events ending fewer than 14 days before December 31)
5. Acute events where the patient enters a skilled nursing facility (SNF), non-acute care, or hospice care within the follow-up interval

De.1. Measure Type: Process

S.17. Data Source: Claims

S.20. Level of Analysis: Health Plan, Other

Preliminary Analysis: New Measure

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Evidence Summary

- Developer provides a logic model seeking to establish the process-outcome link between each of the six conditions addressed by the measure leading to admission to the ED and the appropriate followup period.
- Developer references evidence from multiple guidelines, USPSTF recommendations, articles, and a 2017 NQF Technical Expert Panel, including the strength of the evidence and recommendation when determinable.

- Developer says that they consulted with subject matter experts to resolve conflicting recommendations between guidelines.

Exception to evidence

- NQF staff rated this measure as "insufficient" for evidence, as described below.
- Developer provided as evidence clinical guidelines that represent current best practice for treatment of the chronic conditions under consideration, which include some evidence which is difficult to assess, as well as clear recommendations for follow-up in the post-acute phase. While the recommended course of action did not have an adequate number of studies, it is a best practice recommended by the guideline authors. This may warrant an exception to the evidence criterion.

Questions for the Committee:

- The guideline citations for evidence provided by the developer carry a range of evidence and recommendation strength both for individual conditions and between the conditions. There are instances where guidelines state that evidence is low, but give their strongest recommendation that follow-up occurs. How does the Committee rate the evidence and recommendation strength for each condition given this variation?
    - Condition 1: COPD—3 Clinical Practice Guidelines and 1 Systematic Review
    - Condition 2: CAD—2 Clinical Practice Guidelines
    - Condition 3: HF—1 Clinical Practice Guideline
    - Condition 4: HTN—1 Clinical Practice Guideline (JNC7 & JNC8)
    - Condition 5: Asthma—2 Clinical Practice Guidelines and 1 Systematic Review
    - Condition 6: Diabetes—1 Clinical Practice Guideline
    - Additional evidence found in research for each condition (except asthma) was provided
- Does the Committee feel the developer's approach of expert consultation to resolve differences of recommendations between guidelines was appropriate and sufficiently explained?
- Does the Committee agree that it is acceptable (or beneficial) to hold providers accountable considering the evidence provided?
- This measure has characteristics of a composite measure, though the developer has submitted it as a process measure (after serveral discussions with NQF staff). Measure developer provides a limited rationale for why these specific conditions were included in the measure and not others, citing care gaps, ambulatory care-sensitivity, and evidence-based linkages associated with follow-up. Are these the right conditions, is there a reason for them to be combined together, and are there other conditions to include as well?
- Some conditions warranted a "moderate" evidence rating, while others were rated "insufficient" based on the algorithm from NQF Guidance for Evaluating the Clinical Evidence. Does the committee feel conditions 4 and 6 should be rated as insufficient evidence with exception, based on guideline recommendations?

Guidance from the Evidence Algorithm

- Condition 1: Process measure based on systematic review (Box 3) → QQC not presented (Box 4) → Quantity: high; Quality: moderate; Consistency: moderate (Box 5) → Moderate (Box 5b) → Moderate
- Condition 2: Process measure based on systematic review (Box 3) → QQC presented (Box 4) → Quantity: high; Quality: high; Consistency: low/mod (Box 5) → Moderate (Box 5b) → Moderate
- Condition 3: Process measure based on systematic review (Box 3) → QQC presented (Box 4) → Quantity: moderate; Quality: high; Consistency: high (Box 5) → Moderate (Box 5b) → Moderate

- Condition 4: Process measure based on systematic review (Box 3) → QQC not presented (Box 7) → Empirical evidence not submitted (Box 10) → No PMs of other evidence-based process (Box 11)→ Evidence of systematic assessment of expert opinion (Box 12) → Insufficient
- Condition 5: Process measure based on systematic review (Box 3) → QQC not presented (Box 7) → Empirical evidence submitted with grading (Box 8) → Includes all studies (Box 9) → Benefits outweigh undesirable effects → Moderate
- Condition 6: Process measure based on systematic review (Box 3) → QQC not presented (Box 7) → Empirical evidence not submitted (Box 10) → No PMs of other evidence-based process (Box 11)→ Evidence of systematic assessment of expert opinion (Box 12) → Insufficient

Preliminary rating for evidence: ☐ **High** ☐ **Moderate** ☐ **Low** ☒ **Insufficient**

RATIONALE:

- The evidence provided by the developer for each of the conditions varied.
- Conditions 1, 2, 3, & 5 had moderate supporting evidence, but Conditions 4 & 6 relied on assertions from guidelines based on expert opinion, but insufficiently supported by empirical evidence from studies.
- Because of this, the entire measure is rated as "Insufficient", but may warrant an exception.

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Measure has not been implemented; no performance data available except from testing.
- Developer briefly summarizes evidence for impact of follow up for asthma, COPD, heart failure and diabetes, but not for hypertension or CAD in rationale (though discussed in Evidence). Developer provides statistics related to high prevalence of these conditions and their respective high costs.
- Developer notes paucity of studies related to follow-up rates, but provides several examples demonstrating very low follow-up rates, with most studies demonstrating medium to large performance gaps.
- Developer defines a logical construct for the measure need:
    1. These chronic conditions outlined are prevalent, high-cost, and associated with mortality and morbidity
    2. Outcomes of readmission for these conditions can be influenced through timely follow-up
    3. Timely follow-up has large performance gaps as found in studies within the literature

Disparities

- Developer performed stratified analyes by gender, age and LIS status.
- Did not produce a Cramer's V score greater than 0.05 for any variable, indicating that subpopulations are not experiencing disparities in follow-up rates.
- Developer offers rationale why their data may differ from existing literature.

Questions for the Committee:

- Does the Committee consider the gap in care to be one that warrants a national performance measure?

Preliminary rating for opportunity for improvement: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

Evidence: *1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure."*

- Combining these diverse conditions into a single measure does introduce some conceptual and evidence-base challenges. I share some concern that evidence may be insufficient. Although clinical practice guidelines exist, at times they may represent expert opinion rather than high-quality evidence. I know for some conditions, at times paradoxical relationships exist (where for example, follow-up may be associated with higher readmissions) - the evidence base would be strongest if systematic reviews draw conclusions which have thoroughly considered all studies, with positive and negative relationships accounted for.
- Rationale and Gaps are clear. Evidence is sufficient (by exception for Hypertension & Diabetes). Measure would contribute to evidence base for CAD and Hypertensions especially if stratified (reported out) by condition.
- Lacking of evidence to support the interventions. Possible variation among providers in the implementation of suggested interventions - difficult to evaluate.

Performance Gap: *(1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?)*

- Gaps appear to be substantiated.
- Gaps are clear.Numerator/Denominator are clear.Understand the debate regarding composite vs process...stratified reporting would support actionable responses. Identified summary gap report was helpful especially since identified by condition. The opportunity based approach may mitigate need for stratification. Race was not addressed.
- Evaluation for identifying a gap are limited to asthma, COPD, Heart Failure and diabetes.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications **and** Testing

2b. Validity: Testing**;** Exclusions**;** Risk-Adjustment**;** Meaningful Differences**;** Comparability**;** Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

Evaluation of Reliability and Validity (and composite construction, if applicable):

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The SNQF staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The NQF staff is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Evaluation A: Scientific Acceptability

Measure Number: 3455

Measure Title: Timely Follow-Up After Acute Exacerbations of Chronic Conditions

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☒ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**

☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**

☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☐ **Facility** ☒ **Health Plan**

☐ **Population: Community, County or City** ☐ **Population: Regional and State**

☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☒ **New** ☐ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   ***NOTE**: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**
   - No concerns

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   - The developer assessed the measure score reliability using a signal-to-noise ratio. Reliability is calculated using the beta-binomial model.

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - The developer cites Adams' estimation that a measure is typically considered reliable if the average signal-to-noise reliability is greater than or equal to 0.7. In Medicare Advantage plans, the analysis shows an average signal-to-noise ratio of 0.84 or higher in all three years reported. In qualified health plans/commercial plans, the signal-to-noise ratio is 0.66, which is below the 0.7 threshold that is generally considered reliable, meaning that over half of the plans in the sample would not be reliably ranked.
   - Developer excluded QHPs with denominators fewer than 30. This is not one of the exclusion criteria, and developer did not offer a rationale for selecting this cut off. Developer acknowledges that the bulk of events occurred inside of one large QHP.

Medicare Advantage Signal-to-Noise Reliability

Year	N	Mean	Median	Min	Max	Std. Dev	IQR	P10	P25	P50	P75	P90
2014	657	0.891	0.972	0.142	1.000	0.178	0.118	0.631	0.876	0.972	0.994	0.999
2015	690	0.854	0.960	0.105	1.000	0.203	0.209	0.521	0.782	0.960	0.990	0.997
2016	647	0.839	0.954	0.090	1.000	0.224	0.220	0.484	0.769	0.954	0.989	0.997

Qualified Health Plan/Commercial Signal-to-Noise Reliabiltiy

Year	N	Mean	Median	Min	Max	Std. Dev	IQR	P10	P25	P50	P75	P90
2015	6	0.658	0.678	0.298	0.931	0.244	0.360	0.298	0.503	0.678	0.863	0.931

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.
   Submission document: Testing attachment, section 2a2.2

   ☒**Yes**

   ☐**No**

   ☐**Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?
   Submission document: Testing attachment, section 2a2.2

   ☐**Yes**

   ☐**No**

   ☒**Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☐**High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒**Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐**Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐**Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    (Box 1) Yes → (Box 2) Yes → (Box 4) Yes → (Box 5) Yes → (Box 6b) Yes → Moderate.

    This was rated as moderate, but with a number of concerns.

    - Data for the Medicare Advantage plans was good, which gives some confidence that the measure is reliable under conditions with sufficient events and adequate performance variation.
    - There were a number of problems with the QHP data
        - The sample size for the qualified health plan/commercial plans was small and average signal-to-noise ratio is below 0.7 in those plans, implying little confidence in the ranking of plans against one another for over half of the plans for this level of performance variation.
        - Reliability was artificially inflated by removing plans with n < 30. Without including this as an exclusion criteria, this would be rated as "insufficient".
        - Poor reliability is likely an artifact of the number of events per plan and the variation in performance, but without seeing the data it is difficult to say.
    - Poor reliability for the QHPs is mitigated by the performance on Medicare Advantage plans; it is reasonable to expect comparable reliability for QHPs, if we can assume comparable variability in performace between plans and sufficient events per plan.
        - The concern is that those assumptions are not valid for QHPs.
        - It could simply be that the populations in QHPs do not have the same rate of ER and inpatient utilization as MA plans' populations (which is reasonable to assume given that QHP beneficiaries are generally healthier than MA beneficiaries), or that their performance has too little variation.
        - We cannot draw the conclusion that this is reliable for QHPs based on the data presented.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

Submission document: Testing attachment, section 2b2.

The measure excludes patients in the following cases:

1. For subsequent acute events that occur during the follow-up interval of a prior acute event for the same reason; only the first acute event in the series will be included in the denominator.
2. Acute events after which the beneficiary does not have continuous enrollment for 30 days in the same product.
3. Acute events where the discharge status of the last claim is not "to community". ("Left against medical advice" is not a discharge to community.)
4. Acute events for which the calendar year ends before the follow-up window ends (e.g., acute asthma events ending fewer than 14 days before December 31).
5. Acute events where the patient enters a skilled nursing facility (SNF), non-acute care, or hospice care within the follow-up interval.

The developer provides analysis for these exclusions, noting occurences for each exclusion by plan type:

Medicare Advantage

Exclusion	Occurrence 2014	Occurrence 2015	Occurrence 2016
Acute events before exclusions	**1,585,942**	**1,832,079**	**2,033,319**
Patient is not discharged to community at end of event	36.9%	37.7%	28.9%
Calendar year ends within follow-up window	7.0%	5.8%	6.7%
Event followed by nonacute visit within follow-up window	9.3%	9.6%	12.7%
Event followed by hospice within follow-up window	0.0%	0.0%	0.1%
Event is a readmission within the follow-up window	4.9%	5.2%	5.1%
Non-continuous enrollment	5.4%	4.9%	4.9%
Denominator after exclusions	**810,183**	**935,526**	**1,109,555**

Qualitfied Health Plan/Commercial

Exclusion	Occurrence 2015
Acute events before exclusions	**4,496**
Patient is not discharged to community at end of event	15%
Calendar year ends within follow-up window	6%
Event followed by nonacute visit within follow-up window	30%
Event followed by hospice within follow-up window	0%
Event is a readmission within the follow-up window	4%
Non-continuous enrollment	39%
Denominator after exclusions	**1,499**

In QHP/Commercial Plans, two-thirds of patients are excluded due to the high percentage of patients who do not have continuous enrollment and who have an event followed by nonacute visits within the follow-up window.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

Submission document: Testing attachment, section 2b4.

- No concerns

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
Submission document: Testing attachment, section 2b5.

- N/A

15. **Please describe any concerns you have regarding missing data.**

Submission document: Testing attachment, section 2b6.

- Developer excluded QHPs with denominators fewer than 30. This is not one of the exclusion criteria, and developer did not offer a rationale for selecting this cut off. Developer acknowledges that the bulk of events occurred inside of one large QHP.

16. **Risk Adjustment**

16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

16d.**Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

VALIDITY: TESTING

17. **Validity testing level:** ☒ **Measure score** ☐ **Data element** ☐ **Both**
18. **Method of establishing validity of the measure score:**

☒ **Face validity**

☒ **Empirical validity testing of the measure score**

☐ **N/A (score-level testing not conducted)**

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

- The developer both assessed the measure's face validity and provided empirical analysis of the measure score.
- **Face Validity:** The developer systematically assessed the face validity of the measure by engaging a technical expert panel, comprised of multiple key experts including clinicans, payors, patien representatives, and medical researchers, at multiple stages of the measure development process.
- **Empirical Validity:** The developer also assessed the convergent validity of the measure scores. Convergent validity refers to the degree to which multiple measures of a single underlying

concept are interrelated. For this measure, the developer tested and expected to find MA products with high rates of follow-up after emergency department (ED) visits to have higher rates of follow-up after hospitalizations as well.

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

Face Validity

The developer states that 92% of the experts who reviewed the measure agreed that the measure as specified in this application distinguishes good from poor plan-level quality and recommended the measure as specified to be included in the Health Insurance Exchange program.

Empirical Validity

The developer reports that the correlations are not particularly strong, ranging from 0.1 to 0.4, however the developer does not expect correlations here to reach the 0.7 threshold to be considered meaningful. This is because they are looking at correlation between a product's score in two related but different measures. The developer states that follow-up in 2014 should be a very strong predictor of follow-up in 2015 (test-retest reliability); survey responses of beneficiary perceived care coordination are a predictor of follow-up in 2015, though not as strong as the prior years' follow-up rate.

Follow-Up After Acute Event, Process Measures Affecting Follow-up

Measure	Pearson's Correlation Coefficient**		
	2014	2015	2016
Access to Primary Care Doctor Visits	0.1794	0.3028	0.2556
Getting Appointments and Care Quickly	0.3401	0.2620	0.3263
Getting Needed Care	0.1959	0.1897	0.2129
Medication Reconciliation Post-Discharge	*	0.1527	0.3055

*Measure not included in Star Ratings for this year

**All correlations in this table are statistically significant at p<0.05

Follow-Up After Acute Event, Measures of Care Coordination

Measure	Pearson's Correlation Coefficient**		
	2014	2015	2016
Follow-up visit after Hospital Stay for Mental Illness	0.2706	0.2570	0.2289
Care Coordination	0.2303	0.2739	0.2216

**All correlations in this table are statistically significant at p<0.05

Follow-Up After Acute Event, Outcomes Measures Follow-up Effects

Measure	Pearson's Correlation Coefficient**		
	2014	2015	2016
Annual Flu Vaccine	0.3984	0.2903	0.2595
Asthma Medication Ratio	*	0.2180	0.3062
Continuous Beta Blocker Treatment	0.2694	0.2204	0.2388
Diabetes Care – Eye Exam	0.2163	0.1891	0.1262
Pharmacotherapy Management of COPD Exacerbation – Systemic Corticosteroid	0.2416	0.2550	0.2397
Statin Therapy for Patients With Cardiovascular Disease	*	0.1012	0.1209
Testing to Confirm Chronic Obstructive Pulmonary Disease	0.1856	0.1531	0.2288

*Measure not included in Star Ratings for this year

**All correlations in this table are statistically significant at p<0.05

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

    Submission document: Testing attachment, section 2b1.

    ☒**Yes**

    ☐**No**

    ☐**Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?** *NOTE that data element validation from the literature is acceptable.*

    Submission document*: Testing attachment, section 2b1.*

    ☐**Yes**

    ☐**No**

    ☒**Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

    ☐**High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒**Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

    ☐**Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

    ☐**Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

(Box 1) Yes → (Box 2) Yes → (Box 5) Yes → (Box 6) Yes → (Box 7b) Yes → Moderate Rating

Face validity testing was appropriate. Measures selected for construct validity were not appropriate, resulting in low Pearson's Correlation Coefficients between compared metrics, but this was not required when submitting a new measure, so developer should not be penalized. If measure comes up for maintenance, developer must select measures that are closer in construct in order to do this type of validity testing.

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**
    - Committee must resolve concern related to reliability for MA plans versus QHPs.

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

Reliability- Specifications: *(2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?)*

- It's a bit complicated, but appears to be reliable as proposed.

- Expert consult approach is acceptable. NOT holding providers accountable for guideline directed follo-up care is inacceptable. Agree that measure is a complex process measure though my first inclination was that is was a composite. Reliability is demonstrated however Qualified Health Plans with denominators < 30 may need to be excluded.
- It appears that the interventions may be variable and I would be looking for reliability.

Reliability-Testing: *"2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?"*

- Appropriate notes re: comparability of Medicare Advantage to QHP data...
- Just for QHP's as mentioned in 6.2a1
- Yes due to the variation in interventions and co-morbidities.

Validity-Testing: *"2b1. Validity -Testing: Do you have any concerns with the testing results?"*

- Face validity is presented.
- Validity is demonstrated if QHP's addressed.
- Depending on the source of the data. This will likely be self reported and backed up by clinical documentation.

Validity- Threats to Validity: *"2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?"*

- Question re: risk adjustment.
- One approach to address QHP validity is exclusion of Plans with a denominator < 30.
- There is likely to be missing date for components of the measure due to the variation of conditions addressed.

Other Threats to Validity: *"2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment)2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure?2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?"*

- This may need committee discussion - proposed no risk adjustment.
- Risk adjustment not required
- There is no risk adjustment.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- Coded by someone other than person obtaining orginal information
- All data elements in defined fields
- No fees or licensing requirements

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?

- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Is the data collection strategy ready to be put into operational use?

Preliminary rating for feasibility: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

Feasibility: *"3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?"*

- Varying follow-up intervals adds a layer of complexity to the measure, but it is likely not overly problematic.
- Measure is feasibile
- This is likely to be administrative data so it is feasible.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details

- New measure; developed for use in Medicare Part C and D Star Ratings program and the Quality Rating System program, after endorsement

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- The measure hasn't been used yet, but stakeholders were consulted during development via a TEP of clinicians, payers, patient representatives, researchers, and other experts in the field; the TEP provided feedback early in the development process and again at the end. Workgroup members were asked to assess face validity at the end of the process, and the measure received 92% yes, 8% no on the question of the measure "can be used to distinguish good from poor plan-level quality"
- The measure was also put out for public comment; limited comments were received and focused on evidential strength concerns, risk adjustment, communication barriers, and link to improvement; the developer provided responses to these comments in the submission form

Additional Feedback:

N/A

Questions for the Committee:

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ Pass ☐ No Pass

RATIONALE:

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- The measure is not yet in use so there is no performance data. The developer states that "there is strong reason to believe that this measure, by incentivizing health plans (insurance products) to ensure patients receive appropriate follow-up care, will promote development and implementation of innovative quality improvement activities to improve high performance on this measure. Evidence has shown that quality measures are an effective method in driving quality improvement"
- As this measure is intended to be used in payment programs, the reported measure scores will influence public reporting.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- None available yet – measure is not in use

Potential harms

- None available yet – measure is not in use

Additional Feedback:

- None available yet – measure is not in use

Questions for the Committee:

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- Do the benefits of the measure outweigh any potential unintended consequences?

Preliminary rating for Usability and use: ☐ High ☒ Moderate ☐ Low ☐ Insufficient

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

Use: *"4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided?4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?"*

- New measure. appears to have integrated feedback during development
- Feedback on use Pass
- Not currently publicly reported.

Usability: *"4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations?4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them."*

- Many factors influence the presence/absence of follow-up. follow-up can take many forms beyond the face-to-face clinic visit. it's not obvious that these processes are easily improved by providers / health plans / health systems.
- Measure not implemented yet though usability is moderate
- The data would be usable it there is a link to intervention and outcome.

Criterion 5: Related and Competing Measures

Related or competing measures

Endorsed (related):

- 0229 : Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following heart failure (HF) hospitalization
- 1789 : Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)
- 1891 : Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization

Not endorsed (related):

- Follow-Up After Emergency Department Visit for People with High-Risk Multiple Chronic Conditions. (Steward: NCQA)
- Proportion of persons with a chronic condition that have a potentially avoidable complication during a calendar year. ("PAC") NQF Measure #0709*
- Post-Discharge Appointment for Heart Failure Patients. NQF Measure #2439*
- Heart Failure (HF): Detailed discharge instructions. NQF Measure #0136*

*NQF measures that are no longer endorsed

Harmonization

The developer states the measures are harmonized to the extent possible, and that this measure has either a different focus and target population from the other measures, as well as a different level of analysis.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

Related and Competing: *"5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?"*

- Not to my knowledge.
- No

Public and Member Comments

NQF received no comments on this measure as of: February 1, 2019

Brief Measure Information

NQF #: 3455

Corresponding Measures:

De.2. Measure Title: Timely Follow-Up After Acute Exacerbations of Chronic Conditions

Co.1.1. Measure Steward: IMPAQ International

De.3. Brief Description of Measure: The percentage of issuer-product-level acute events requiring either an emergency department (ED) visit or hospitalization for one of the following 6 chronic conditions: hypertension, asthma, heart failure (HF), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), or diabetes mellitus (Type I or Type II), where follow-up was received within the timeframe recommended by clinical practice guidelines in a non-emergency outpatient setting.

1b.1. Developer Rationale: Follow-up care is an element of care coordination, which has recently been highlighted as a priority area within the CMS Quality Strategy goal of Eliminating Disparities. (1) Follow-up care allows providers to perform a number of important activities, including ensuring that patients understand and are adhering to their medication regimen, monitoring them for adverse events, and educating them to recognize warning signs. (2) To date, there are scant published studies or data, especially within the US, that provide aggregate rates of timely follow-up care after ED or hospitalization for the conditions considered in this measure, making it imperative to fill in the gap in the healthcare community.

The body of evidence described below suggests, broadly and for the specific conditions included in this measure, that failing to obtain timely follow-up care is empirically linked to higher readmission rates, which are known to be associated with increased costs and decreased patient satisfaction and are widely considered an indicator of sub-optimal patient care. Furthermore, in the attached evidence tables we provide clinical guidelines for each condition, recommending that patients receive follow-up care after inpatient or ED discharge within the respective timeframe set for each chronic condition in the measure.

The goal of this measure is to improve the quality of care provided and improve patient outcomes by incentivizing health plans (insurance products) to ensure patients receive appropriate follow-up care following acute exacerbations of chronic conditions. To achieve this goal, the measure will identify health plans that have significantly lower rates of appropriate follow-up visits for acute conditions, relative to other health plans with the same acute conditions for similar patient populations. In doing so, this measure will prompt health plans to carefully evaluate care processes and implement quality improvement strategies. (3) Ultimately, this measure will provide an opportunity for health plans to become aware of and to improve rates of appropriate follow-up visits following acute events leading to decreased morbidity and mortality for patients with any of the 6 conditions covered by this measure.

By incentivizing health plans (insurance products) to improve follow-up rates for these conditions, this measure will improve outcomes as demonstrated in the following logic model:

1. Patient presents to the hospital/ED with an acute exacerbation of 1 of the 6 conditions included in the measure.
2. Patient is treated (either admitted to the hospital or treated only in the emergency department) and discharged to the community.
3. Health plan (insurance products) encourages follow-up visit/care through strategies such as incentives to providers,* reminders to patients, providing data/reports and continuing education to providers, etc.
4. Patient receives follow-up visit/care based upon evidence-based clinical guidelines, conditions are appropriately managed, with improved patient health and function.

5. ED and hospital cost and utilization are reduced by preventing avoidable readmissions.

* Note: If a health plan provides bonuses or other financial incentives to providers based on quality performance, an unintended consequence of holding plans accountable could be that providers are financially penalized if the plan performs poorly. However, providers, in this case, have every incentive to help patients receive timely follow-up care, which could render the net effects being positive.

This logic model is made available in visual format as an appendix to the evidence form.

In general, the evidence demonstrates the presence of a link between timely follow-up care after discharge from the hospital/ED and lower rates of readmission. (4, 5) Specifically, the evidence also supports the link between follow-up care and improved outcomes for each of the conditions included in this measure, as described below.

• Asthma: A systematic review states that evidence from randomized controlled trials (RCTs) and non-RCTs suggests that timely follow-up with specialists reduces subsequent asthma exacerbations, as well as fewer symptoms and improved quality of life. (6) A population-based study on 7,829 patients with asthma or COPD found that follow-up visits within 30 days of discharge were associated with a significant reduction in the 90-day risk of an emergency readmission (RR=.79). (7) In a study of Medicare beneficiaries with COPD, 10% of whom also had asthma as a primary diagnosis code, the 30-day rate of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. Within the study population, patients who had a timely follow-up visit had a significantly reduced risk of ED visit (HR=.86) and readmission (HR=.91). (8)

• Heart Failure: Outpatient follow-up from a cardiology or general medicine provider within 7 days for a heart failure patient after hospitalization is associated with a lower chance of 30-day readmission (OR=.81). (9) Delayed outpatient follow-up after myocardial infarction has been associated with worse short-term and long-term patient medication adherence. Logically, this will add to morbidity and very likely both mortality and subsequent cardiovascular events/subsequent hospital admissions. (10) In the United States, the Hospital to Home (H2H) program, a national quality improvement initiative, also recommends that all heart failure and myocardial infarction patients have a follow-up appointment or cardiac rehab referral scheduled within 7 days of discharge. (11) Furthermore, the H2H program implemented a structured improvement project called the "See You in 7" initiative at 10 hospitals in Southeast Michigan, which included follow-up within a week of discharge as a core concept. After one year of participation, the adjusted 30-day readmission rates at collaborating hospitals decreased compared to non-participating hospitals (2.6% decrease vs. 0.6% decrease). (12) Heart failure and myocardial infarction patients in Taiwan were found to have a lower risk of 30-day readmission if they received outpatient visit with a physician within 7 days of discharge (HR=.54). (13) In a study of 3,136 patients, those who received cardiovascular follow-up saw fewer ED (38% vs. 80%) and hospital (13% vs. 94%) readmissions for cardiovascular reasons within the year, and lower unadjusted mortality (7% vs. 2% at 30 days). (14) In a study of 30,136 patients, patients discharged from hospitals with higher rates of follow-up within 7 days of discharge for heart failure had a lower risk of 30-day readmission (HR .91 between highest and lowest quartiles). (15)

While direct evidence on the relationship between timely follow-up and CAD and hypertension is extremely limited, both conditions are well recognized comorbidities and predictors of heart failure. (16, 17, 18) Logically, because a strong body of evidence links timely follow-up after heart failure to lower readmissions, timely follow-up for CAD and hypertension can also be reasoned to lower readmissions.

• COPD: In a study of Medicare beneficiaries with COPD, 10% of whom also had asthma as a primary diagnosis code, the 30-day rates of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. Within the study population, patients who had a timely follow-up visit had a significantly reduced risk of ED visit (HR=.86) and readmission (HR=.91). (8) A population-based study of 7829 patients with asthma or COPD found that follow-

up visits within 30 days of discharge were associated with a significant reduction in the 90-day risk of an emergency readmission (RR=.79). (7)

• Diabetes: A review article suggests that better management of hyperglycemia in the ED, along with proper follow-up to ensure continued appropriate management, improves clinical outcomes and prevents readmission. (19) Another study found that sentinel ED visits in diabetic patients are a warning sign for future readmission for hyperglycemia, and recommended that clinicians "provide clear discharge instructions for follow-up and glucose management to prevent further hyperglycemic emergencies from occurring." (20)

The gap in care described above represents one reason these particular 6 conditions were selected for inclusion in the measure. These measures were also selected because they are ambulatory care-sensitive conditions that respond well to timely primary care, as reflected in broad body of evidence (described in the evidence form) linking timely follow-up for these conditions to improved health outcomes. Additionally, these conditions were selected based on the impact they have on patients and health systems, measured in both condition prevalence as well as the costs and resources associated with appropriate treatment. For example, the CDC finds that approximately 30 million Americans have diabetes, and complications were estimated to cost the US about $245 billion in 2012. (21) Nearly 27 million Americans are living with asthma, which is responsible for about 440,000 inpatient discharges, as well as 1.7 million ED visits and 11 million physician's office visits per year, with a total economic burden estimated at about $53 billion. (22, 23)

The prevalence of hypertension is even higher, with about 75 million adults (1 in 3) suffering from high blood pressure and an estimated $50 billion in economic burden. (24) About 15.7 million Americans suffer from COPD, with mortality rates as high as 62.8 per 100,000 in Kentucky and an overall economic burden of nearly $50 billion. (25, 26) About 15 million Americans suffer from CAD, and about 6 million suffer from heart failure. Heart disease remains the leading cause of death in the United States and represents an immense burden on the health system as a whole and on patients and their families individually. (27, 28)

While the prevalence of these 6 conditions is not uniformly high, they incur considerable cost to health systems as well as individual patients, and are conditions that represent degraded patient health and risk of mortality. The following statistics summarize incidence and cost data gathered from the 2014 Healthcare Cost and Utilization Project. Please note that ED cost data were not available.

Hypoglycemia

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.92
  - ED Rate per 100,000 persons: 8.06
  - Estimate of Mean Inpatient Costs: $7,177
- Any DX
  - Inpatient Rate per 100,000 persons: 12.55
  - ED Rate per 100,000 persons: 28.75

Hyperglycemia

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.16
  - ED Rate per 100,000 persons: 10.06
  - Estimate of Mean Inpatient Costs: $4,595
- Any DX
  - Inpatient Rate per 100,000 persons: 21.4
  - ED Rate per 100,000 persons: 509.26

Diabetic Ketoacidosis

- Principal DX
  - Inpatient Rate per 100,000 persons: 45.79
  - ED Rate per 100,000 persons: 48.77
  - Estimate of Mean Inpatient Costs: $7,280
- Any DX
  - Inpatient Rate for 100,000 persons: 59.75
  - ED Rate per 100,000 persons: 63.29

Acute Myocardial Infarction

- Principal DX
  - Inpatient Rate per 100,000 persons: 69.28
  - ED Rate per 100,000 persons: 64.31
  - Estimate of Mean Inpatient Costs: $21,380
- Any DX
  - Inpatient Rate per 100,000 persons: 95.79
  - ED Rate per 100,000 persons: 87.13

Angina

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.44
  - ED Rate per 100,000 persons: 0.59
  - Estimate of Mean Inpatient Costs: $7,693
- Any DX
  - Inpatient Rate per 100,000 persons: 1.61
  - ED Rate per 100,000 persons: 1.93

Acute Asthma Attack

- Principal DX
  - Inpatient Rate per 100,000 persons: 68.60
  - ED Rate per 100,000 persons: 462.98
  - Estimate of Mean Inpatient Costs: $6,555
- Any DX
  - Inpatient Rate per 100,000 persons: 471.03
  - ED Rate per 100,000 persons: 1824.8

Hypotension

- Principal DX
  - Inpatient Rate per 100,000 persons: 8.25
  - ED Rate per 100,000 persons: 21.06
  - Estimate of Mean Inpatient Costs: $7,088
- Any DX
  - Inpatient Rate per 100,000 persons: 156.05
  - ED Rate per 100,000 persons: 143.97

Shortness of Breath

- Principal DX
  - Inpatient Rate per 100,000 persons: 1.04
  - ED Rate per 100,000 persons: 61.44
  - Estimate of Mean Inpatient Costs: $5,486
- Any DX
  - Inpatient Rate per 100,000 persons: 17.18
  - ED Rate per 100,000 persons: 386.76

Acute Pulmonary Edema

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.67
  - ED Rate per 100,000 persons: 0.88
  - Estimate of Mean Inpatient Costs: $9,011
- Any DX
  - Inpatient Rate per 100,000 persons: 5.51
  - ED Rate per 100,000 persons: 4.35

Peripheral Edema

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.7
  - ED Rate per 100,000 persons: 47.12
  - Estimate of Mean Inpatient Costs: $6,468
- Any DX
  - Inpatient Rate per 100,000 persons: 40.48
  - ED Rate per 100,000 persons: 128.07

Acute Left Ventricular Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.07
  - ED Rate per 100,000 persons: 0.08
  - Estimate of Mean Inpatient Costs: $8,839
- Any DX
  - Inpatient Rate per 100,000 persons: 0.78
  - ED Rate per 100,000 persons: 1.23

Pleural Effusion

- Principal DX
  - Inpatient Rate per 100,000 persons: 4.29
  - ED Rate per 100,000 persons: 7.31
  - Estimate of Mean Inpatient Costs: $12,562
- Any DX
  - Inpatient Rate per100,000 persons: 65.46

- ED Rate per 100,000 persons: 60.93

Hypertension

- Principal DX
  - Inpatient Rate per 100,000 persons: 7.33
  - ED Rate per 100,000 persons: 138.68
  - Estimate of Mean Inpatient Costs: $5,699
- Any DX
  - Inpatient Rate per 100,000 persons: 1280.37
  - ED Rate per 100,000 persons: 3710.17

Chronic Kidney Disease and Systolic Heart Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 2.54
  - ED Rate per 100,000 persons: 6.81
  - Estimate of Mean Inpatient Costs: $19,507
- Any DX
  - Inpatient Rate per 100,000 persons: 366.59
  - ED Rate per 100,000 persons: 440.07

Pulmonary Congestion and Hypostasis

- Principal DX
  - Inpatient Rate per 100,000 persons: .63
  - ED Rate per 100,000 persons: 1.62
  - Estimate of Mean Inpatient Costs: $7,724
- Any DX
  - Inpatient Rate per 100,000 persons: 11.10
  - ED Rate per 100,000 persons: 11.68

Congestive Heart Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 13.67
  - ED Rate per 100,000 persons: 33.86
  - Estimate of Mean Inpatient Costs: $11,374
- Any DX
  - Inpatient Rate per 100,000 persons: 353.30
  - ED Rate per 100,000 persons: 462.81

COPD

- Principal DX
  - Inpatient Rate per 100,000 persons: 70.49
  - ED Rate per 100,000 persons: 407.92
  - Estimate of Mean Inpatient Costs: $8,019
- Any DX

- Inpatient Rate per 100,000 persons: 467.86
- ED Rate per 100,000 persons: 1119.54

CITATIONS:

1) Centers for Medicare and Medicaid Services. CMS Quality Strategy. 2016. Available from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Downloads/CMS-Quality-Strategy.pdf

2) California Healthcare Foundation. The Post-Hospital Follow-Up Visit: A Physician Checklist to Reduce Readmissions. October 2010. Available from https://www.chcf.org/wp-content/uploads/2017/12/PDF-PostHospitalFollowUpVisit.pdf.

3) Chassin M, et al. Accountability Measures – Using Measurement to Promote Quality Improvement. N Engl J Med 2010; 363:683-688. doi: 10.1056/NEJMsb1002320.

4) Misky GJ, et al. Post-hospitalization transitions: Examining the effects of timing of primary care provider follow-up. J Hosp Med. 2010 Sep;5(7):392-7. doi: 10.1002/jhm.666.

5) Shen E, et al. Association of a Dedicated Post–Hospital Discharge Follow-up Visit and 30-Day Readmission Risk in a Medicare Advantage Population. JAMA Intern Med. 2017;177(1):132-135. doi:10.1001/jamainternmed.2016.7061.

6) Schatz M, et al. Follow-Up After Acute Asthma Episodes: What Improves Future Outcomes? J Emerg Med. 2009 Aug;37(2 Suppl):S42-50. doi: 10.1016/j.jemermed.2009.06.109.

7) Sin D et al. The impact of follow-up physician visits on emergency readmissions for patients with asthma and chronic obstructive pulmonary disease: a population-based study. Am J Med. 2002 Feb 1;112(2):120-5. doi: 10.1016/S0002-9343(01)01079-8.

8) Sharma G, et al. Outpatient Follow-up Visit and 30-Day ED Visit and Readmission in Patients Hospitalized for Chronic Obstructive Pulmonary Disease. Arch Intern Med. 2010 Oct 11; 170(18): 1664–1670. doi: 10.1001/archinternmed.2010.345.

9) Lee K, et al. Post-Discharge Follow-up Characteristics Associated with 30-Day Readmission After Heart Failure Hospitalization. Med Care. 2016 Apr; 54(4): 365–372. doi: 10.1097/MLR.0000000000000492.

10) Faridi KF, et al. Timing of First Postdischarge Follow-up and Medication Adherence After Acute Myocardial Infarction. JAMA Cardiol. 2016 May 1;1(2):147-55. doi: 10.1001/jamacardio.2016.0001.

11) Hospital 2 Home: Improving Transitions from the Hospital to Community Settings. 2013. Slide Deck. Available from https://cvquality.acc.org/docs/default-source/h2h/1-h2h-overview-slides.pdf?sfvrsn=c6d58fbf_2.

12) Baker H, et al. Regional Hospital Collaboration and Outcomes in Medicare Heart Failure Patients: See You in 7. JACC Heart Fail. 2015 Oct;3(10):765-73. doi: 10.1016/j.jchf.2015.06.007.

13) Tung YC, et al. Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure. PLoS One. 2017 Jan 27;12(1):e0170061. doi: 10.1371/journal.pone.0170061.

14) Ezekowitz JA, et al. Impact of specialist follow-up in outpatients with congestive heart failure. CMAJ. 2005 Jan 18;172(2):189-94. doi: 10.1503/cmaj.1032017.

15) Hernandez AF, et al. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. JAMA. 2010 May 5;303(17):1716-22. doi: 10.1001/jama.2010.533.

16) Velagaleti R, et al. Heart Failure in the 21st Century: Is it a Coronary Artery Disease Problem or Hypertension Problem? Cardiol Clin. 2007 Nov; 25(4): 487–v. doi: 10.1016/j.ccl.2007.08.010.

17) Lala A, et al. The role of coronary artery disease in heart failure. Heart Fail Clin. 2014 Apr;10(2):353-65. doi: 10.1016/j.hfc.2013.10.002.

18) Messerli F, et al. The Transition from Hypertension to Heart Failure. JACC: Heart Failure Aug 2017, 5 (8) 543-551; DOI: 10.1016/j.jchf.2017.04.012.

19) Echouffo-Tcheugui J, et al. Management of Hyperglycemia and Diabetes in the ED. Curr Diab Rep (2017) 17: 56. https://doi.org/10.1007/s11892-017-0883-2.

20) Yan JW, Gushulak KM, Columbus MP, Hamelin AL, Wells GA, Stiell IG. Sentinel visits in ED patients with diabetes mellitus as a warning sign for hyperglycemic emergencies. CJEM. 2017 Jul 25:1-8. doi: 10.1017/cem.2017.338.

21) ADA. Economic Costs of Diabetes in the U.S. in 2012. Diabetes Care. 2013 Apr; 36(4): 1033–1046.

22) Asthma and Allergy Foundation of America. Costs of Asthma on Society. Available from http://www.aafa.org/page/cost-of-asthma-on-society.aspx.

23) Centers for Disease Control and Prevention. Most Recent Asthma Data. May 2018. Available from https://www.cdc.gov/asthma/most_recent_data.htm.

24) Centers for Disease Control and Prevention. High Blood Pressure. Available from https://www.cdc.gov/bloodpressure/index.htm.

25) Guarascio A, et al. The clinical and economic burden of chronic obstructive pulmonary disease in the USA. Clinicoecon Outcomes Res. 2013; 5: 235–245. Published online 2013 Jun 17. doi: 10.2147/CEOR.S34321.

26) Centers for Disease Control and Prevention. Chronic Obstructive Pulmonary Disease (COPD). Available from https://www.cdc.gov/copd/data.html.

27) Sanchis-Gomar F, et al A. Epidemiology of coronary heart disease and acute coronary syndrome. Ann Transl Med. 2016 Jul; 4(13): 256. doi: 10.21037/atm.2016.06.33.

28) Roger V. Epidemiology of Heart Failure. Circ Res. 2013 Aug 30; 113(6): 646–659. doi: 10.1161/CIRCRESAHA.113.300268.

S.4. Numerator Statement: The numerator is the sum of the issuer-product-level denominator events (Emergency Room [ED], observation hospital stay or inpatient hospital stay) for acute exacerbation of hypertension, asthma, heart failure (HF), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), or diabetes where follow-up was received within the timeframe recommended by clinical practice guidelines, as detailed below:

- Hypertension: Within 7 days of the date of discharge
- Asthma: Within 14 days of the date of discharge
- HF: Within 14 days of the date of discharge
- CAD: Within 14 days of the date of discharge
- COPD: Within 30 days of the date of discharge
- Diabetes: Within 30 days of the date of discharge

S.6. Denominator Statement: The denominator is the sum of the plan-product-level acute exacerbations that require either an ED visit, observation stay, or inpatient stay (i.e., acute events) for any of the six conditions listed above (hypertension, asthma, HF, CAD, COPD, or diabetes).

S.8. Denominator Exclusions: The measure excludes events with:

1. Subsequent acute events that occur two days after the prior discharge, but still during the follow-up interval of the prior event for the same reason. To prevent double-counting, only the first acute event will be included in the denominator.
2. Acute events after which the patient does not have continuous enrollment for 30 days in the same product.
3. Acute events where the discharge status of the last claim is not “to community” (“Left against medical advice” is not a discharge to community.)

4. Acute events for which the calendar year ends before the follow-up window ends (e.g., acute asthma events ending fewer than 14 days before December 31)
5. Acute events where the patient enters a skilled nursing facility (SNF), non-acute care, or hospice care within the follow-up interval

De.1. Measure Type: Process

S.17. Data Source: Claims

S.20. Level of Analysis: Health Plan, Other

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? N/A – Our measure is not paired or grouped.

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

NQF_Evidence_Form_Follow-Up_Measure_10.25.2018_FOR_SUBMISSION.docx

1a.1 <u>For Maintenance of Endorsement:</u> Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

No

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:**

Measure Title: NQF# 3455 Timely Follow-Up After Acute Exacerbations of Chronic Conditions

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: N/A

Date of Submission: Click here to enter a date

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*

- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Outcome: a Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence b that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: c a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence b that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence b that the measured structure leads to a desired health outcome.
- Efficiency: d evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

a. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

b. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

c. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

d. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome:

☐Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

☐ Intermediate clinical outcome (*e.g., lab value*):

☒ Process: The percentage of issuer-product-level[1] acute events requiring either an emergency department (ED) visit or hospitalization for one of the following 6 chronic conditions: hypertension, asthma, heart failure (HF), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), or diabetes mellitus (Type I or Type II), where follow-up was received within the timeframe recommended by clinical practice guidelines in a non-emergency outpatient setting.

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

The logic model establishing the process-outcome link for this measure concept is listed below; the process step corresponding to the measure concept appears in **bold**:

1. Patient presents to the hospital/ED with an acute exacerbation of 1 of the 6 conditions included in the measure.
2. Patient is treated (either admitted to the hospital or treated only in the emergency department) and discharged to the community.
3. Health plan (insurance products) encourages follow-up visit/care through strategies such as incentives to providers,[2] reminders to patients, providing data/reports and continuing education to providers, etc.

---

[1] This measure is defined at the issuer-product level, meaning that results are aggregated for each qualified insurance issuer and for each product. For clarity, a product is a discrete package of health insurance coverage benefits that issuers offer in the context of a particular network type, such as health maintenance organization (HMO), preferred provider organization (PPO), exclusive provider organization (EPO), point of service (POS), or indemnity. Issuers are broadly defined as health insurance providers who participate in the Federally-facilitated Marketplaces and health insurance contracts offered in the Medicare Advantage market.

[2] Note: If a health plan provides bonuses or other financial incentives to providers based on quality performance, an unintended consequence of holding plans accountable could be that providers are financially penalized if the plan performs poorly. However, providers, in this case, have every incentive to help patients receive timely follow-up care, which could render the net effects being positive.

4. **Patient receives follow-up visit/care based upon evidence-based clinical guidelines, conditions are appropriately managed, with improved patient health and function.**
5. ED and hospital cost and utilization are reduced by preventing avoidable readmissions.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

N/A; this measure is not derived from patient report.

****RESPOND TO ONLY ONE SECTION BELOW – (EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

N/A

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (FOR INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☒ Clinical Practice Guideline recommendation (with evidence review)

☒ US Preventive Services Task Force Recommendation

☒ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☒ Other

Exhibit A below summarizes the follow-up timeframes that were set for each of the 6 chronic conditions included in this measure. These follow-up timeframes were derived from the follow-up periods identified by clinical practice guideline(s) or systematic review(s) of the literature. In the event of inconsistent follow-up periods across guidelines or reviews, we consulted subject matter experts and the Technical Expert Panel (TEP) for an overall recommendation. **Exhibit B** (on the next page) summarizes the follow-up recommendations, by condition as described in the clinical practice guideline(s) or systematic review(s). Following Exhibits A and B are the full evidence tables, grouped by condition.

Exhibit A: Summary of Follow-up Periods Used in the Measure Specification, by Condition

Chronic Condition	Follow-up Period Used in Measure Specification
Chronic Obstructive Pulmonary Disease (COPD)	Within 30 days
Coronary Artery Disease (CAD)	Within 14 days
Heart Failure (HF)	Within 14 days
Hypertension	Within 7 days
Asthma	Within 14 days
Diabetes	Within 30 days

Exhibit B: Summary of Follow-up Periods from Systematic Reviews, Clinical Practice, and USPSTF Guidelines (by Condition)

Table #	Clinical Practice Guideline	Systematic Review	Year and Source	Grading and Strength of:		Follow-up Period Recommendation
				Evidence	Recommendation	
Chronic Condition #1: Chronic Obstructive Pulmonary Disease (COPD)				**Follow-up Period in Measure Specification: Within 30 days**		
1	X		2015 American College of Chest Physicians and Canadian Thoracic Society	C = low or very low quality evidence	1 = strong recommendation	monthly
2	X		2016 University of Michigan	D = individual observational studies	I = generally should be performed.	moderate or low risk patients within one month of discharge
3	X and USPSTF		2014 Department of Veterans Affairs/Department of Defense 2014 USPSTF Graded Recommendation	Not provided	GRADE: Strong For USPSTF Grade B: Service is recommended.	30-90 days
4		X	2017 Health Quality Ontario	Low and very low	Not provided	30 days
Chronic Condition #2: Coronary Artery Disease (CAD)				**Follow-up Period in Measure Specification: Within 14 days**		
5	X		2014 American College of Cardiology/American Heart Association Task Force	B = limited populations evaluated, data from single randomized controlled trial (RCT) or non-RCTs	Class IIa = benefit > risk; it is reasonable to perform procedure/ administer treatment.	Within 3 days
6	X		2012 American College of Cardiology Foundation/American Heart Association Task Force	C = very limited populations evaluated; only consensus opinion of experts, case studies, or standard of care	Class I = benefit > risk; procedure/treatment should be performed/administered.	Within 2-6 weeks or within 14 days for higher risk patients
Chronic Condition #3: Heart Failure (HF)				**Follow-up Period in Measure Specification: Within 14 days**		
7	X		2013 American College of Cardiology Foundation/American Heart Association Task Force	B = limited populations evaluated, data from single RCT or non-RCTs	Class IIa = benefit > risk; it is reasonable to perform procedure/ administer treatment.	Within 7-14 days
Chronic Condition #4: Hypertension				**Follow-up Period in Measure Specification: Within 7 days**		
8	X		2004 National Heart, Lung, and Blood Institute. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC7) 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults Report from the Panel Members Appointed to the Eighth Joint National Committee (JNC 8)	Not provided	Not provided	"Within several days"
Chronic Condition #5: Asthma				**Follow-up Period in Measure Specification: Within 14 days**		
9	X		2007 Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma	B = RCTs; limited body of data	Strong recommendation	1-4 weeks
10	X		2018 Global Strategy for Asthma Management and Prevention (GINA) Report	D = panel consensus	Not provided	2-7 days
11		X	2009 Journal of Allergy and Clinical Immunology	D (definition not provided by the author)	Strong	Within one week
Chronic Condition #6: Diabetes				**Follow-up Period in Measure Specification: Within 30 days**		
12	X		2017 American Diabetes Association	B = supportive evidence from well-conducted cohort studies	Not provided	Within one month

Chronic Condition #1: Chronic Obstructive Pulmonary Disease (COPD)

(3 Clinical Practice Guidelines and 1 Systematic Review)

Table 1: Clinical Practice Guideline

Source of Systematic Review (Clinical Practice Guideline): • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Prevention of acute exacerbations of COPD: American College of Chest Physicians and Canadian Thoracic Society Guideline** • **Criner GJ, Bourbeau J, Diekemper RL et al.** • **April 2015** • **Criner GJ, Bourbeau J, Diekemper RL, et al. Prevention of acute exacerbations of COPD: American College of Chest Physicians and Canadian Thoracic Society Guideline. Chest. 2015 Apr;147(4):894-942** • **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4388124/pdf/chest_147_4_894.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Recommendation #8: In patients with COPD with a previous or recent history of exacerbations, the panel recommends education and case management that includes direct access to a health-care specialist **at least monthly** to prevent severe acute exacerbations of COPD, as assessed by decreases in hospitalizations."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	1C: Strong recommendation, low- or very-low-quality evidence; evidence for at least one critical outcome from observational studies, case series, or RCTs with serious flaws or indirect evidence
Provide all other grades and definitions from the evidence grading system.	Quality of Body of Evidence (A, B, C, or CB) A = High-quality evidence; consistent evidence RCTs without important limitations or exceptionally strong evidence from observational studies B = Moderate-quality evidence; evidence from RCTs with important limitations (inconsistent results, methodologic flaws, indirect, or imprecise) or very strong evidence from observational studies C = low- or very-low-quality evidence; evidence for at least one critical outcome from observational studies, case series, or RCTs with serious flaws or indirect evidence CB (consensus based) = Insufficient evidence for a graded recommendation
Grade assigned to the **recommendation** with definition of the grade	1C: Strong recommendation, low- or very-low-quality evidence Recommendation can apply to most patients in many circumstances. Higher-quality research is likely to have an important impact on confidence in the estimate of effect and may well change the estimate.
Provide all other grades and definitions from the recommendation grading system.	Strength of Recommendation (Level 1 or 2) 1: Strong recommendation; benefits clearly outweigh risk and burden or vice versa. 2: Weak recommendation; Benefits closely balance with risks and burden.

Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: The Guidelines International Network (GIN) Library and National Guideline Clearinghouse were used to search for guidelines on COPD. GIN search found 26 guidelines and National Guideline Clearinghouse search found 24 guidelines. In total, 8 guidelines were considered relevant and assessed. PubMed and the Cochrane Library were used to search for Systematic reviews and primary literature. The search of the Cochrane Library resulted in 127 systematic reviews, and an additional 14 systematic reviews were found in the PubMed search. • Quality: Systematic reviews include randomized controlled trials (randomized controlled prospective parallel-group study, randomized clinical trial, prospective unblended randomized controlled trial).
Estimates of benefit and consistency across studies	1C grade states "benefits clearly outweigh risk and burdens or vice versa."
What harms were identified?	None of the studies reported any adverse events related to the intervention.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new clinical guidelines related to COPD follow-up from American College of Chest Physicians and Canadian Thoracic Society

Table 2: Clinical Practice Guideline

Source of Systematic Review (Clinical Practice Guideline): • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Care of the Hospitalized Patient with Acute Exacerbation of COPD** • **University of Michigan COPD Guideline Team (Sagana RL, Wesorick DH, Bassin BS, et al.)** • **May 2016** • **Sagana RL, Wesorick DH, Bassin BS, et al. Care of the Hospitalized Patient with Acute Exacerbation of COPD. UMHS Chronic Obstructive Pulmonary Disease. May 2016:1-28** • **http://www.med.umich.edu/1info/FHP/practiceguides/InptCOPD/COPD.final.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"A comprehensive approach to discharge is recommended (I, D). Key elements of the hospital discharge are summarized in Table 6. (Table 6): At UH [University Hospital], we recommend that patients at high risk for readmission be seen 7-10 days after discharge and patients at moderate or low risk be seen **within one month** of discharge."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Level of evidence supporting a diagnostic method or an intervention D = individual observation studies (case or case series)
Provide all other grades and definitions from the evidence grading system	A = systematic reviews of randomized controlled trials (RCTs) B = RCTs C = systematic review of non-RCTs or observational studies, non-RCTs, group observation studies (e.g., cohort, cross-sectional, case control) D = individual observation studies (case or case series) E = opinion of expert panel

Grade assigned to the **recommendation** with definition of the grade	Strength of recommendation I= generally should be performed.
Provide all other grades and definitions from the recommendation grading system	I = generally should be performed. II = may be reasonable to perform. III = generally should not be performed.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: The main search retrieved 860 references. The MEDLINE In-Process database was also searched, retrieving 386 documents. The Cochrane database of systematic reviews was searched, producing 22 reviews. • Quality: The main search focused on guidelines, clinical trials, and cohort studies. The base results are as follows: 13 results for COPD Guidelines, 215 results for COPD Clinical Trials, and 167 results for COPD cohort studies. The MEDLINE In-Process database search found 7 results for COPD guidelines, 85 results for clinical trials, and 114 results for cohort studies.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided
What harms were identified?	Harms were not identified
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	The University of Michigan published 2017 COPD Guidelines for Clinical Care in the Ambulatory (outpatient) Setting (see link below); but these do not address recommendations for follow-up visits after discharge from the ED or inpatient setting, and do not change the conclusions from this SR. http://www.med.umich.edu/1info/FHP/practiceguides/copd/copd.pdf

Table 3: Clinical Practice Guideline/ USPSTF Recommendation

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **VA/DoD Clinical Practice Guideline for the Management of Chronic Obstructive Pulmonary Disease** • **Management of Chronic Obstructive Pulmonary Disease Working Group** • **December 2014** • **Management of Chronic Obstructive Pulmonary Disease Working Group. VA/DoD Clinical Practice Guideline for the Management of Chronic Obstructive Pulmonary Disease. Department of Veterans Affairs/Department of Defense. December 2014: p. 19, 36, and 73.** • **https://www.healthquality.va.gov/guidelines/CD/copd/VADoDCOPDCPG2014.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Recommendation 28: We recommend that patients discharged home from hospitalization with acute transitional oxygen therapy are evaluated for the need for LTOT [long-term oxygen therapy] **within 30-90 days** after discharge. LTOT should not be discontinued if patients continue to meet the above criteria."

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A; grading was only provided for strength of the recommendation.
Provide all other grades and definitions from the evidence grading system.	N/A; grading was only provided for strength of the recommendation.
Grade assigned to the **recommendation** with definition of the grade	Assigned GRADE and USPSTF Grades below: • GRADE Strength of Recommendation: Strong For (or "We recommend offering this option ...") • USPSTF Grade: B: 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.
Provide all other grades and definitions from the recommendation grading system.	The GRADE of a recommendation is based on the following elements: • Relative strength (Strong or Weak): A strong recommendation indicates that the Work Group is highly confident that desirable outcomes outweigh undesirable outcomes. If the Work Group is less confident of the balance between desirable and undesirable outcomes, they present a weak recommendation. • Direction (For or Against). • The four choices are: o Strong For (or "We recommend offering this option …") o Weak For (or "We suggest offering this option …") o Weak Against (or "We suggest not offering this option …") o Strong Against (or "We recommend against offering this option …") USPSTF Grades A = The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B = 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. C= The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences. There is at least moderate certainty that the net benefit is small. D = The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I = The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: In this systematic review, 454 articles were reviewed; 94 articles were included. • Quality: As per the VA/DoD Guideline for Guidelines document, risk-of-bias (or study quality) of individual studies and previous systematic reviews was assessed using the USPSTF method. Each study was assigned a rating of "Good, Fair, or Poor" based on sets of criteria that vary depending on the study design.

Estimates of benefit and consistency across studies	USPSTF Grade B: "There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial."
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new clinical guidelines related to COPD follow-up from VA/DoD.

Table 4: Systematic Review

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Effect of Early Follow-Up After Hospital Discharge on Outcomes in Patients With Heart Failure or Chronic Obstructive Pulmonary Disease: A Systematic Review.** • **Song J, Walter M.** • **2017** • **Health Quality Ontario. Song J, Walter M. Effect of Early Follow-Up After Hospital Discharge on Outcomes in Patients With Heart Failure or Chronic Obstructive Pulmonary Disease: A Systematic Review. Ontario Health Technology Assessment Series. 2017;17(8):1-37.** • **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5466361/**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Based on low- and very low-quality evidence, **follow-up within 7 days and within 30 days of discharge** from hospitalization for heart failure or COPD—compared with usual care or no follow-up—were both associated with a **reduced risk of all-cause readmission, ED visits, and mortality**. Overall, there is a lack of large, methodologically robust studies specifically focusing on the effectiveness of 7-day follow-up after discharge in improving patient outcomes." "From studies comparing a 30-day follow-up with usual care or no follow-up in patients with heart failure or COPD: • Low-quality evidence showed a significantly reduced risk of 30-day all-cause readmission, ED visits, and death; of 3-month all-cause readmission; and of a composite measure of 6-month all-cause readmission or death. • Very low-quality evidence showed a significant difference in rates of 3-month ED visits."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	**Low**: Our confidence in the estimate is limited: the true prognosis (probability of future events) may be substantially different from the estimate • Low quality evidence for: 30-day all-cause readmission, ED visits, and death; of 3-month all-cause readmission; and of a composite measure of 6-month all-cause readmission or death. **Very low**: We have very little confidence in the estimate: the true prognosis (probability of future events) is likely to be substantially different from the estimate • Very low-quality evidence for: rates of 3-month ED visits.

Provide all other grades and definitions from the evidence grading system.	The quality of the body of evidence for each outcome was examined according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group criteria. High = We are very confident that the true prognosis (probability of future events) lies close to that of the estimate. Moderate = We are moderately confident that the true prognosis (probability of future events) is likely to be close to the estimate, but there is a possibility that it is substantially different. Low = Our confidence in the estimate is limited: the true prognosis (probability of future events) may be substantially different from the estimate. Very low = We have very little confidence in the estimate: the true prognosis (probability of future events) is likely to be substantially different from the estimate.
Grade assigned to the **recommendation** with definition of the grade	N/A; only gave quality of evidence grading.
Provide all other grades and definitions from the recommendation grading system.	N/A; only gave quality of evidence grading.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Quantity: From a total of 3,228 unique citations, we identified 10 eligible studies: 1 RCT, 2 non-RCTs, and 7 observational studies. • 4 studies were specifically on 7-day follow-up and 30-day health outcomes. The other 6 studies were on 30-day follow-up and more variable time to health outcomes. Quality: • We appraised the quality of RCTs using the Effective Practice and Organization of Care (EPOC) criteria and the quality of observational studies using the Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I).
Estimates of benefit and consistency across studies	"The findings generally show that early follow-up after hospital discharge is associated with improved patient outcomes, though the evidence is inconsistent and of low quality. Compared with patients who do not receive early follow-up care, those who do receive it have a lower risk of readmission or ED use within 30 days of discharge, the two outcomes with the most supporting studies." "The available studies do not demonstrate a clear difference in the effects of follow-up within 7 versus 30 days, but this may be due to the small number and low quality of studies on a 7-day follow-up. Of the 4 studies on a 7-day follow-up that we included, 1 did not adjust for covariates, 1 large study had improper adjustment, 1 study used aggregate data which yielded results that are difficult to interpret, and 1 study was too small."
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies since this systematic review that would change the conclusions.

Chronic Condition #2: Coronary Artery Disease (CAD)

(2 Clinical Practice Guidelines)

Table 5: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **2014 AHA/ACC Guideline for the Management of Patients With Non–ST-Elevation Acute Coronary Syndromes: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines** • **Amsterdam EA, et al.** • **2014** • **Amsterdam EA, Wenger NK, Brindis RG, et al. 2014 ACC/AHA guideline for the management of patients with non–ST-elevation acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014. p. e156 and e176** • **https://ac.els-cdn.com/S0735109714062792/1-s2.0-S0735109714062792-main.pdf?_tid=4993bd9a-67e2-4f96-9c43-b9537a9dfd19&acdnat=1532454208_ab1f5d3814fb9af943ef85bd5c42c5a4** • **Data Supplement 8. Discharge from ED or Chest Pain Unit (p. 20): http://jaccjacc.cardiosource.com/acc_documents/2014_NSTE-ACS_Data_Supplement_Tables.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"3.5.1. Discharge From the ED or Chest Pain Unit: Recommendations Class IIa 2. It is reasonable for patients with possible ACS [acute coronary syndrome] who have normal serial ECGs [electrocardiography] and cardiac troponins to have a treadmill ECG (Level of Evidence: A), stress myocardial perfusion imaging, or stress echocardiography before discharge or within **72 hours after discharge** (Level of Evidence B)."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	For the 3.5.1 recommendation: Level B: limited populations evaluated; data derived from a single RCT or non-RCTs
Provide all other grades and definitions from the evidence grading system.	Level A = Multiple populations evaluated; data derived from multiple RCTs or meta-analyses Level B = Limited populations evaluated; data derived from a single RCT or non-RCTs Level C = Very limited populations evaluated; only consensus opinion of experts, case studies, or standard of care
Grade assigned to the **recommendation** with definition of the grade	Class IIa = benefit > risk; additional studies with focused objectives needed; it is reasonable to perform procedure/administer treatment.

Provide all other grades and definitions from the recommendation grading system.	Class I = benefit > risk; Procedure/treatment should be performed/administered. Class IIa = benefit > risk; additional studies with focused objectives needed; it is reasonable to perform procedure/administer treatment. Class IIb = benefit ≥ risk; additional studies with broad objectives needed; additional registry data would be helpful; procedure/treatment may be considered. Class III No Benefit = procedure/test is not helpful and there is no proven benefit to treatment Class III Harm = procedure/test is excess cost without benefit or harmful; treatment is harmful to patients.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	An extensive evidence review was conducted through October 2012, and other selected references published through April 2014 were reviewed by the GWC. Literature included was derived from research involving human subjects and indexed in MEDLINE (through PubMed), EMBASE, the Cochrane Library, Agency for Healthcare Research and Quality Reports, and other selected databases relevant to this CPG. Per Data Supplement 8 (hyperlink included in citation at the beginning of the table): • Quantity: 9 studies • Quality: 2 single-center prospective RCTs, 1 observational single-center study, 2 prospective single-center studies, 3 prospective multi-center RCTs, and one observational cohort study
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new clinical guidelines related to CAD follow-up from ACC/AHA.

Table 6: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **2012 ACCF/AHA Focused Update Incorporated Into the ACCF/AHA 2007 Guidelines for the Management of Patients With Unstable Angina/Non–ST-Elevation Myocardial Infarction: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines** • **American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines** • **June 4, 2013** • **American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, 2012 ACCF/AHA Focused Update Incorporated Into the ACCF/AHA 2007 Guidelines for the Management of Patients With Unstable Angina/Non–ST-Elevation Myocardial Infarction: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, *Circulation*: p. e756.** • **http://circ.ahajournals.org/content/early/2013/04/29/CIR.0b013e31828478ac** • **Online data supplement: https://www.ahajournals.org/action/downloadSupplement?doi=10.1161%2FCIR.0b013e31828478ac&file=online_data_supplement.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"5.3. Postdischarge Follow-Up: Low-risk medically treated patients and revascularized patients should return in **2 to 6 weeks**, and higher risk patients should return within **14 d[ays]**."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Level C: Very limited populations evaluated; only consensus opinion of experts, case studies, or standards of care
Provide all other grades and definitions from the evidence grading system.	Level A = Multiple populations evaluated; data derived from multiple RCTs or meta-analyses Level B = Limited populations evaluated; data derived from a single RCT or non-RCTs Level C = Very limited populations evaluated; only consensus opinion of experts, case studies, or standard of care
Grade assigned to the **recommendation** with definition of the grade	Class I: benefit > risk; Procedure/treatment should be performed/administered.

Provide all other grades and definitions from the recommendation grading system.	Class I = benefit > risk; Procedure/treatment should be performed/administered. Class IIa = benefit > risk; additional studies with focused objectives needed; it is reasonable to perform procedure/administer treatment. Class IIb = benefit ≥ risk; additional studies with broad objectives needed; additional registry data would be helpful; procedure/treatment may be considered. Class III No Benefit = procedure/test is not helpful and there is no proven benefit to treatment. Class III Harm = procedure/test is excess cost without benefit or harmful; treatment is harmful to patients.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: 14 studies included in data supplement table • Quality: Not all study-types provided , but the data supplement table does identify at least 1 prospective randomized trial, 1 randomized factorial trial, 1 randomized double-blind trial, 1 randomized clinical trial, and 2 meta-analyses of randomized controlled trials.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	ACC/AHA released 2014 guidelines, which are included in Table 1 above. The 2014 guidelines are more specific to the treatments that should occur post-discharge, whereas 2012 guidelines are general for follow-up recommendations.

Chronic Condition #3: Heart Failure (HF)

(1 Clinical Practice Guideline)

Table 7: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines** • **Yancy CW, Jessup M, Bozkurt B, et al.** • **October 15, 2013** • **Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation. 2013; p. e289-e290** • **http://circ.ahajournals.org/content/128/16/e240**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	Table 29 (p. e289): "A follow-up visit **within 7 to 14 days** and/or a telephone follow-up within 3 days of hospital discharge are reasonable." p. e290: "Scheduling an early follow-up visit (**within 7 to 14 days**) and early telephone follow-up (within 3 days) of hospital discharge are reasonable."

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Level B: Limited population evaluated, data derived from a single RCT or non-RCTs
Provide all other grades and definitions from the evidence grading system.	Level A = Multiple populations evaluated; data derived from multiple RCTs or meta-analyses Level B = Limited populations evaluated; data derived from a single RCT or non-RCTs Level C = Very limited populations evaluated; only consensus opinion of experts, case studies, or standard of care
Grade assigned to the **recommendation** with definition of the grade	Class IIa = benefit > risk; additional studies with focused objectives needed; it is reasonable to perform procedure/administer treatment.
Provide all other grades and definitions from the recommendation grading system.	Class I = benefit > risk; Procedure/treatment should be performed/administered. Class IIa = benefit > risk; additional studies with focused objectives needed; it is reasonable to perform procedure/administer treatment. Class IIb = benefit ≥ risk; additional studies with broad objectives needed; additional registry data would be helpful; procedure/treatment may be considered Class III No Benefit = procedure/test is not helpful, and there is no proven benefit to treatment. Class III Harm = procedure/test is excess cost without benefit or harmful; treatment is harmful to patients.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	An extensive evidence review was conducted through October 2011 and includes selected other references through April 2013. Searches were extended to studies, reviews, and other evidence conducted in human subjects that were published in English from PubMed, EMBASE, Cochrane, Agency for Healthcare Research and Quality Reports, and other selected databases relevant to this guideline. • Quantity: 2 studies • Quality: 1 observational analysis; the other study type was not provided.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Keane K. Lee, Jingrong Yang, Adrian F. Hernandez, Anthony E. Steimle, Alan S. Go. (2016) Post-discharge Follow-up Characteristics Associated With 30-Day Readmission After Heart Failure Hospitalization. Medical Care 54:4, 365-372. Yu-Chi Tung, Guann-Ming Chang, Hsien-Yen Chang, Tsung-Hsien Yu, Giuseppe Andò. (2017) Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure. *PLOS ONE* 12:1, e0170061. • These studies both associate outpatient follow-up within 7 days with decreased readmissions. Keane et al. notes, however, that later follow-up (within 8-30 days) was not significantly associated with readmission. • The ACC/AHA Task Force on Clinical Practice Guidelines released a 2017 Focused update of these 2013 guidelines; however, the guidelines related to follow-up did not change and therefore do not change the conclusions from the 2013 guidelines.

Chronic Condition #4: Hypertension

(1 Clinical Practice Guideline)

Table 8: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **JNC 7: The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure** • **U.S. Department of Health and Human Services, National Institutes of Health: National Heart, Lung, and Blood Institute** • **August 2004** • **U.S. Department of Health and Human Services, National Institutes of Health: National Heart, Lung, and Blood Institute. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. 2004. NIH Publication No. 04-5230: p. 54.** • **https://www.nhlbi.nih.gov/files/docs/guidelines/jnc7full.pdf** **JNC8 was released in 2014, but does not include updated guidelines for follow-up. The scope of topics for JNC8 was narrower than JNC7. For JNC8, the evidence review of RCTs addressed a limited number of questions, judged by the panel to be of the highest priority. JNC8 did not impact the conclusions of the JNC 7 guidelines above.)** **James PA, Oparil S, Carter BL, et al. 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults Report From the Panel Members Appointed to the Eighth Joint National Committee (JNC 8). *JAMA.* 2014;311(5):507–520. doi:10.1001/jama.2013.284427**

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Most importantly, patients should not leave the ER without a confirmed follow-up visit **within several days**."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Not graded and classification of evidence for this recommendation was not included.
Provide all other grades and definitions from the evidence grading system.	Various systems of grading the evidence were considered, and the classification scheme used in JNC 6 and other NHBPEP clinical guidelines was selected. This scheme classifies studies according to a process adapted from Last and Abramson (see Scheme Used for Classification of the Evidence below). • M Meta-analysis; use of statistical methods to combine the results from clinical trials • RA RCTs; also known as experimental studies • RE Retrospective analyses; also known as case-control studies • F Prospective studies; also known as cohort studies, including historical or prospective follow-up studies • X Cross-sectional surveys; also known as prevalence studies • PR Previous review or position statements • C Clinical interventions (nonrandomized)
Grade assigned to the **recommendation** with definition of the grade	Not graded
Provide all other grades and definitions from the recommendation grading system.	No additional grading of recommendations
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The Executive Committee developed relevant medical subject headings (MeSH) terms and keywords to further review the scientific literature. These MeSH terms were used to generate MEDLINE searches that focused on English-language, peer-reviewed, scientific literature from January 1997 through April 2003. The Executive Committee focused its deliberations on evidence pertaining to outcomes of importance to patients and with effects of sufficient magnitude to warrant changes in medical practice ("patient-oriented evidence that matters," or POEMs). • Quantity: 4 studies • Quality: 3 previous reviews or position statements and 1 cross-sectional survey; also known as prevalence study
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Updated evidence-based guidelines for the management of hypertension were released in 2014 (JNC 8). Updated guidelines for follow-up were not included and therefore do not change the conclusions of the JNC 7 guidelines above. James PA, Oparil S, Carter BL, et al. 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults. Report From the Panel Members Appointed to the Eighth Joint National Committee (JNC 8). *JAMA.* 2014;311(5):507–520. doi:10.1001/jama.2013.284427

Chronic Condition #5: Asthma

(2 Clinical Practice Guidelines and 1 Systematic Review)

Table 9: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma: Full Report 2007** • **HHS, NIH, NHLBI** • **August 28, 2007** • **National Heart, Lung, and Blood Institute, Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma: Full Report 2007. 2007: p. 102, 373, 388.** • **https://www.nhlbi.nih.gov/files/docs/guidelines/asthgdln.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	p. 102: "The Expert Panel recommends that: At the time of discharge from the ED, patients be referred for follow-up asthma care appointment (either PCP or asthma specialist) **within 1–4 weeks** (Evidence B)." p. 373: "Management of asthma exacerbations requiring urgent medical care (e.g., in the urgent care setting or ED) includes: preventing relapse of the exacerbation or recurrence of another exacerbation by providing: referral to follow-up asthma care **within 1–4 weeks** (Evidence B)." p. 388: "Figure 5-6: Discharge Home: Before discharge, schedule follow-up appointment with primary care provider and/or asthma specialist in **1–4 weeks**." p. 400: "The Expert Panel recommends the following actions for discharging patients from the ED: Emphasize the need for continual, regular care in an outpatient setting, and refer the patient for a follow-up asthma care appointment (either primary care provider (PCP) or asthma specialist) **within 1–4 weeks** (Evidence B). If appropriate, consider referral to an asthma self-management education program (Evidence B)."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Evidence Category B: RCTs, limited body of data Evidence is from endpoints of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of RCTs, or meta-analysis of RCTs. In general, category B pertains when few RCTs exist; they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent.

Provide all other grades and definitions from the evidence grading system	Evidence Category A: RCTs, rich body of data. Evidence is from endpoints of well-designed RCTs that provide a consistent pattern of findings in the population for which the recommendation is made. Category A requires substantial numbers of studies involving substantial numbers of participants. Evidence Category B: RCTs, limited body of data. Evidence is from endpoints of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of RCTs, or meta-analysis of RCTs. In general, category B pertains when few RCTs exist; they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent. Evidence Category C: Non-RCTs and observational studies Evidence is from outcomes of uncontrolled or non-RCTs or from observational studies. Evidence Category D: Panel consensus judgment. This category is used only in cases where the provision of some guidance was deemed valuable, but the clinical literature addressing the subject was insufficient to justify placement in one of the other categories. The Panel consensus is based on clinical experience or knowledge that does not meet the criteria for categories A through C.
Grade assigned to the **recommendation** with definition of the grade	Strong recommendation as the guideline states that it is recommended "When a certain clinical practice "**is recommended,"** this indicates a strong recommendation by the panel."
Provide all other grades and definitions from the recommendation grading system	When a certain clinical practice "is recommended," this indicates a strong recommendation by the panel. When a certain clinical practice "should, or may, be considered," this indicates that the recommendation is less strong.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: The literature review was conducted in 3 cycles over an 18-month period (September 2004 to March 2006). Of the abstracts, 2,122 were advanced for full-text review, which resulted in 1,654 articles serving as a bibliography of references used to update the guidelines, available on the NHLBI Web site. For Section 5, "Managing Exacerbations of Asthma" (where the guideline is included) there are 261 full text reviews. • Quality: Observational studies, interventional studies and RCTs
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	The National Heart, Lung, and Blood Institute (NHLBI) has not yet published EPR-4, so EPR-3 contains the most recent guidelines. There are no new guidelines from the NHLBI that change the conclusions of this SR.

Table 10: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Global Strategy for Asthma Management and Prevention** • **Global Initiative for Asthma** • **2018** • **Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention, 2018. Available from: www.ginaasthma.org; p. 87-88, and 121.** • **https://ginasthma.org/2018-gina-report-global-strategy-for-asthma-management-and-prevention/**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	p. 87: "Prior to discharge from the ED or hospital to home, arrangements should be made for a follow-up appointment **within 1 week**, and strategies to improve asthma management including medications, inhaler skills and written asthma action plan, should be addressed." p. 88: Box 4-5: "A follow-up appointment **within 2–7 days of discharge** should be made with the patient's usual healthcare provider, to ensure that treatment is continued, that asthma symptoms are well controlled, and that the patient's lung function reaches their personal best (if known)." p. 121: "Prior to discharge from the ED or hospital, family/caregivers should receive the following advice and information (all are Evidence D): A follow-up appointment **within 2–7 days and another within 1–2 months,** depending on the clinical, social and practical context of the exacerbation."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	For recommendation on p. 121, graded D. The page 87 and 88 recommendations do not list grading. D: This category is used only in cases where the provision of some guidance was deemed valuable, but the clinical literature addressing the subject was insufficient to justify placement in one of the other categories. The Panel Consensus is based on clinical experience or knowledge that does not meet the above listed criteria.

Provide all other grades and definitions from the evidence grading system.	A= RCTs and meta-analyses. Rich body of data. Evidence is from endpoints of well-designed RCTs or meta-analyses of relevant studies that provide a consistent pattern of findings in the population for which the recommendation is made. Category A requires a substantial number of studies involving substantial number of participants. B = RCTs and meta-analyses. Limited body of data. Evidence is from endpoints of intervention studies that include only a limited number of patients, post hoc or subgroup analysis of RCTs or meta-analysis of such RCTs. In general, Category B pertains when few RCTs exist, they are small in size, they were undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent. C = Non-RCTs. Non-RCT observational studies. Evidence is from outcomes of uncontrolled or non-RCTs or from observational studies. D = Panel consensus judgement. This category is used only in cases where the provision of some guidance was deemed valuable, but the clinical literature addressing the subject was insufficient to justify placement in one of the other categories. The Panel Consensus is based on clinical experience or knowledge that does not meet the above listed criteria.
Grade assigned to the **recommendation** with definition of the grade	N/A; strength of recommendation not included
Provide all other grades and definitions from the recommendation grading system.	Grading system for recommendations not used, only for the evidence.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: 65 publications • Quality: 35 clinical trials and 26 meta-analyses
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	This is the most recent GINA report, so there are no new studies to change the conclusions from this SR.

Table 11: Systematic Review

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Follow-up after acute asthma episodes: What improves future outcomes?** • **Schatz, Michael et al.** • **August 2009** • **Schatz, Michael et al. Follow-up after acute asthma episodes: What improves future outcomes? Journal of Allergy and Clinical Immunology, August 2009, Volume 124, Issue 2, p. S41.** • **http://www.jacionline.org/article/S0091-6749(09)00798-2/fulltext** • **Introduction (includes methodology): https://www.jacionline.org/article/S0091-6749(09)00705-2/pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Strong: recommend that the follow-up visit with the primary care physician, asthma specialist, or specialized asthma clinic be **within 1 week** of the ED visit (Evidence Category D)."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Evidence Category D: Task force consensus judgment based on clinical experience and other nonsystematic clinical observations.
Provide all other grades and definitions from the evidence grading system.	The task force used the same system to describe the level of evidence as used by the EPR3 Expert Panel (Table 9) • Evidence Category A: randomized controlled trials, rich body of data • Evidence Category B: randomized controlled trials, limited body of data • Evidence Category C: nonrandomized trials and observational studies • Evidence Category D: task force consensus judgment based on clinical experience and other nonsystematic clinical observations.
Grade assigned to the **recommendation** with definition of the grade	Strong
Provide all other grades and definitions from the recommendation grading system	Strong: Clinical Practice is recommended Conditional: Clinical Practice should or might be considered
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity and Quality: The literature search produced 25 RCTs and 6 meta-analyses. 10 RCTs were deemed relevant for this review, with none of the meta-analyses considered relevant.

Estimates of benefit and consistency across studies	"Effective and timely outpatient care of asthma can prevent adverse asthma outcomes, specifically ED visits and hospitalizations." "… appropriate follow-up is essential to optimize outcomes after acute asthma."
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies conducted that change the conclusions from this SR.

Chronic Condition #6: Diabetes

(1 Clinical Practice Guideline)

Table 12: Clinical Practice Guideline

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Diabetes Care in the Hospital** • **American Diabetes Association** • **2017** • **American Diabetes Association. Diabetes Care in the Hospital. Sec. 14. In Standards of Medical Care in Diabetes 2017. Diabetes Care 2017;40(Suppl. 1):S120, S124-125.** • **https://professional.diabetes.org/sites/professional.diabetes.org/files/media/dc_40_s1_final.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	p. S124: "An outpatient follow-up visit with the primary care provider, endocrinologist, or diabetes educator **within 1 month of discharge** is advised for all patients having hyperglycemia in the hospital. If glycemic medications are changed or glucose control is not optimal at discharge, an earlier appointment (in 1–2 weeks) is preferred, and frequent contact may be needed to avoid hyperglycemia and hypoglycemia." p. S120: "There should be a structured discharge plan tailored to the individual patient with diabetes. Evidence Grade: B"
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The statement on p. S124 is not graded. The statement on S120 is graded B. B = Supportive evidence from well-conducted cohort studies
Provide all other grades and definitions from the evidence grading system.	ADA Evidence Grading System for "Standards of Medical Care in Diabetes" A = Clear evidence from well-conducted, generalizable RCTs that are adequately powered B = Supportive evidence from well-conducted cohort studies C = Supportive evidence from poorly controlled or uncontrolled studies E = Expert consensus or clinical experience
Grade assigned to the **recommendation** with definition of the grade	N/A; only grading of evidence

Provide all other grades and definitions from the recommendation grading system	N/A; only grading of evidence
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: 67 references for "Section 14. Diabetes Care in the Hospital" where the recommendations are found. • Quality: Not specified.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Yan JW, Gushulak KM, Columbus MP, Hamelin AL, Wells GA, Stiell IG. Sentinel visits in emergency department patients with diabetes mellitus as a warning sign for hyperglycemic emergencies. CJEM. 2017 Jul 25:1-8. doi: 10.1017/cem.2017.338. • In Yan et al, diabetic patients with a sentinel ED visit often returned and required subsequent admission for hyperglycemia and the study recommends that clinicians "provide clear discharge instructions for follow-up and glucose management to prevent further hyperglycemic emergencies from occurring." Echouffo-Tcheugui JB, Garg R. Management of Hyperglycemia and Diabetes in the Emergency Department. Curr Diab Rep. 2017 Aug; 17(8):56. doi: 10.1007/s11892-017-0883-2. Both studies support follow-up visits after a diabetic ER visit for reduced readmissions and improved clinical outcomes, but do not recommend a specific timeframe for follow-up and do not change the conclusions from this SR. • In Echouffo et al, evidence suggests that better management of hyperglycemia in the ED with proper follow-up improves clinical outcomes and prevents readmission.

Exhibit C (on the next page) summarizes the timeframes for follow-up as described in other evidence the team collected (research articles, consensus/summary reports, quality improvement initiatives, etc.).

Exhibit C: Summary of Follow-up Periods from Other Evidence Sources (by Condition)

Table #	Identify the type of "other" evidence	Year and Source	Grading and Strength of:		Follow-up Period Recommendation
			Evidence	Recommend-ation	
Condition #1: Chronic Obstructive Pulmonary Disease (COPD)			**Follow-up Period in Measure Specification: Within 30 days**		
13	Consensus Report	2017 Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease (GOLD) Report	Statements 1 & 2 are not graded; Statement 3 is B = RCTs with a limited number of patients.	Not provided	30 days
Chronic Condition #2: Coronary Artery Disease (CAD)			**Follow-up Period in Measure Specification: Within 14 days**		
14	National Quality Improvement Initiative	2013 American College of Cardiology Hospital to Home	Not provided	Not provided	Within 7 days
Chronic Condition #3: Heart Failure (HF)			**Follow-up Period in Measure Specification: Within 14 days**		
15	Research Article	Yu-Chi Tung, Guann-Ming Chang, Hsien-Yen Chang, Tsung-Hsien Yu, Giuseppe Andò. (2017) Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure. PLOS ONE 12:1, e0170061.	Not provided	Not provided	Within 7 days
16	Research Article	Keane K. Lee, Jingrong Yang, Adrian F. Hernandez, Anthony E. Steimle, Alan S. Go. (2016) Post-discharge Follow-up Characteristics Associated With 30-Day Readmission After Heart Failure Hospitalization. *Medical Care* 54:4, 365-372.	Not provided	Not provided	Within 7 days
Chronic Condition #4: Hypertension			**Follow-up Period in Measure Specification: Within 7 days**		
17	Summary Paper	Kessler CS, Joudeh Y. Evaluation and treatment of severe asymptomatic hypertension. 2010. American Family Physician 81(4): 475.	Not provided	Not provided	Within 1 to 7 days for severe uncontrolled hypertension
Chronic Condition #5: Asthma			**Follow-up Period in Measure Specification: Within 14 days**		
None					
Chronic Condition #6: Diabetes			**Follow-up Period in Measure Specification: Within 30 days**		
18	Research Article	Yan JW, Gushulak KM, Columbus MP, van Aarsen, K, Hamelin AL, Wells GA, Stiell IG. Risk factors for recurrent emergency department visits for hyperglycemia in patients with diabetes mellitus. *International Journal of Emergency Medicine.* 2017; 10(23): 1-8.	Not provided	Not provided	Timeframe not included

Table 19: In addition to the articles above that discuss specific follow-up timeframes, we also include information pertaining to the importance of general follow-up from the 2017 National Quality Forum Emergency Department Transitions of Care: A Quality Measurement Framework Final Report. Specific follow-up timeframes are not included in the report.

Chronic Condition #1: Chronic Obstructive Pulmonary Disease (COPD)

(1 consensus report)

Table 13: Consensus Report

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report** • **Global Initiative for Chronic Obstructive Lung Disease** • **2017** • **Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: 2017 Report. p. 61, 108-109** • **https://goldcopd.org/gold-2017-global-strategy-diagnosis-management-prevention-copd/**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"Early follow-up **(within 1 month)** following discharge should be undertaken when possible and has been related to less exacerbation-related readmissions. There are many patient issues that prevent early follow-up; those not attending early follow-up have increased 90-day mortality." "Table 5.7: Ensure follow-up arrangements: early follow-up < 4 weeks…" "Table 3.8. Pulmonary rehabilitation reduces hospitalizations among patients who have had a recent exacerbation (≤ 4 weeks from prior hospitalization) (Evidence B)"
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The first two statements are not graded, the third is Evidence B (Evidence is from RCTs that include only a limited number of patients, post hoc or subgroup analyses of RCTs, or meta-analyses of RCTs. Also pertains when few RCTs exist, or important limitations are evident).
Provide all other grades and definitions from the evidence grading system.	Levels of Evidence have been assigned to evidence-based recommendations where appropriate. A = Evidence is from endpoints of well-designed RCTs that provide consistent findings in the population for which the recommendation is made without any important limitations. Requires high quality evidence from more than or equal to 2 clinical trials involving a substantial number of subjects, or a single high quality RCT involving substantial numbers of patients without any bias. B = Evidence is from RCTs that include only a limited number of patients, post hoc or subgroup analyses of RCTs, or meta-analyses of RCTs. Also pertains when few RCTs exist, or important limitations are evident (methodologic flaws, small numbers, short duration, undertaken in a population that differs from the target population of the recommendation, or the results are somewhat inconsistent). C = Evidence is from outcomes of uncontrolled or non-RCTs or from observational studies. D = panel consensus judgement; provision of guidance is deemed valuable, but clinical literature addressing the subject is insufficient. Panel consensus is based on clinical experience or knowledge that does not meet the above stated criteria.

Grade assigned to the **recommendation** with definition of the grade	N/A; only gave quality of evidence grading
Provide all other grades and definitions from the recommendation grading system.	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	To produce the GOLD report, a PubMed search was completed using search fields established by the Committee: *COPD, All fields, Adult, 19+ years, only items with abstracts, Clinical Trial, meta-analyses, Human.* The literature included for this 2017 update was published from 2015 to 2016. • The quantity and quality of the studies was not provided in the report.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new recommendations related to COPD follow-up from the Global Initiative for Chronic Obstructive Lung Disease.

Chronic Condition #2: Heart Failure (HF)

(1 National Quality Improvement Initiative)

Table 14: National Quality Improvement Initiative

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **American College of Cardiology's Hospital to Home (H2H) Initiative** • **American College of Cardiology** • **2013** • **Improving Transitions from the Hospital to Community Settings. (2013). Retrieved from https://cvquality.acc.org/docs/default-source/h2h/1-h2h-overview-slides.pdf?sfvrsn=2** • **https://cvquality.acc.org/docs/default-source/h2h/1-h2h-overview-slides.pdf?sfvrsn=c6d58fbf_2**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"See You In 7 Challenge. Goal: All patients discharged with a diagnosis of HF [heart failure] and MI [myocardial infarction] have a scheduled follow-up appointment or cardiac rehab referral made **within 7 days** of discharge."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A
Provide all other grades and definitions from the evidence grading system.	N/A
Grade assigned to the **recommendation** with definition of the grade	N/A

Provide all other grades and definitions from the recommendation grading system.	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	N/A
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	The H2H program implemented a structured improvement project called the "See You in 7" initiative at 10 hospitals in Southeast Michigan. After 1 year of participation, the adjusted 30-day readmission rates at collaborating hospitals decreased substantially compared to non-participating hospitals. The conclusions from this study support the H2H recommendations of follow-up within 7 days. Baker, H., Oliver-Mcneil, S., Deng, L., & Hummel, S. L. (2015). Regional Hospital Collaboration and Outcomes in Medicare Heart Failure Patients. JACC: Heart Failure, 3(10), 765-773. doi:10.1016/j.jchf.2015.06.007

Chronic Condition #3: Coronary Artery Disease (CAD)

(2 research articles)

Table 15: Research Article

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure** • **Yu-Chi Tung et al.** • **2017** • **Yu-Chi Tung, Guann-Ming Chang, Hsien-Yen Chang, Tsung-Hsien Yu, Giuseppe Andò. (2017) Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure. PLOS ONE 12:1, e0170061.** • **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5271349/pdf/pone.0170061.pdf**

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	• Early physician follow-up (defined as a visit with a physician **within 7 days after discharge**) was associated with a lower hazard ratio of readmission compared with no early physician follow-up for patients with NSTEMI (hazard ratio [HR], 0.47; 95% confidence interval [CI], 0.39±0.57), and for patients with heart failure (HR, 0.54; 95% CI, 0.48±0.60). Early follow-up was defined as whether discharged patients had an outpatient visit. The length of 7 days was selected to be consistent with current efforts to improve transitional care (the authors cited ACC's Hospital to Home Initiative). • This study showed that 7-day physician follow-up was associated with a lower 30-day readmission rate, and physician continuity (7-day same physician follow-up) was associated with a much lower 30-day readmission rate for patients with NSTEMI and those with heart failure. • This study may provide evidence to support guidelines recommending scheduling an early follow-up visit after discharge, and may provide an evidenced-based approach to improve 30-day readmission following NSTEMI and heart failure.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A
Provide all other grades and definitions from the evidence grading system.	N/A
Grade assigned to the **recommendation** with definition of the grade	N/A
Provide all other grades and definitions from the recommendation grading system.	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The authors used nationwide population-based data to examine associations between 7-day physician follow-up and 30-day readmission rates, and further associations of 7-day same physician (during the index hospitalization and at follow-up) and cardiologist follow-up with 30-day readmission rates for non-ST-segment-elevation myocardial infarction (NSTEMI) or heart failure. The authors analyzed all patients 18 years or older with NSTEMI and heart failure and discharged from hospitals in 2010 in Taiwan through Taiwan's National Health Insurance Research Database. Cox proportional hazard models with robust sandwich variance estimates and propensity score weighting were performed after adjustment for patient and hospital characteristics to test associations between 7-day physician follow-up and 30-day readmission.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies were conducted that changed the conclusions of this study.

Table 16: Research Article

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Post-discharge Follow-up Characteristics Associated With 30-Day Readmission After Heart Failure Hospitalization.** • **Lee et al.** • **2016** • **Keane K. Lee, Jingrong Yang, Adrian F. Hernandez, Anthony E. Steimle, Alan S. Go. (2016) Post-discharge Follow-up Characteristics Associated With 30-Day Readmission After Heart Failure Hospitalization. *Medical Care* 54:4, 365-372.** • **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4800825/pdf/nihms743092.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	• Among 11,985 eligible adults, early initial outpatient contact **within 7 days** after discharge was associated with lower odds of readmission, whereas later outpatient contact, between 8 and 30 days after hospital discharge, was not significantly associated with readmission. • In adults discharged to home after hospitalization for HF, outpatient follow-up with a cardiology or general medicine provider **within 7 days** was associated with a lower chance of 30-day readmission.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A
Provide all other grades and definitions from the evidence grading system.	N/A
Grade assigned to the **recommendation** with definition of the grade	N/A
Provide all other grades and definitions from the recommendation grading system.	N/A

Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Nested matched case-control study (January 1, 2006 to June 30, 2013). The study was conducted within Kaiser Permanente Northern California, a large, integrated healthcare delivery system that provides comprehensive care to more than 3.7 million members within the San Francisco Bay Area. The authors included subjects hospitalized from January 1, 2006 through June 30, 2013 with a primary discharge diagnosis of HF (*International Classification of Diseases, Ninth Edition* [ICD-9] codes 398.91, 402.01, 402.11, 402.91, 428.0, 428.1, or 428.9), which has been shown to have a positive predictive value of >95% for clinical heart failure.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies were conducted that changed the conclusions of this study.

Chronic Condition #4: Hypertension

(1 summary paper)

Table 17: Summary Paper

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Evaluation and treatment of severe asymptomatic hypertension** • **Kessler CS, Joudeh Y.** • **February 15, 2010** • **Kessler CS, Joudeh Y. Evaluation and treatment of severe asymptomatic hypertension. American Family Physician 81(4): 475.** • **http://www.aafp.org/afp/2010/0215/p470.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	Table 3: "Severe uncontrolled hypertension: Initiate treatment and follow-up **within one to seven days** of presentation."
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A
Provide all other grades and definitions from the evidence grading system.	N/A
Grade assigned to the **recommendation** with definition of the grade	N/A
Provide all other grades and definitions from the recommendation grading system.	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	• Quantity: 5 studies • Quality: N/A

Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies to change the conclusion.

Chronic Condition #5: Asthma

None

Chronic Condition #6: Diabetes

(1 research article)

Table 18: Research Article

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Risk factors for recurrent emergency department visits for hyperglycemia in patients with diabetes mellitus** • **Yan et al.** • **2017** • **Yan JW, Gushulak KM, Columbus MP, van Aarsen, K, Hamelin AL, Wells GA, Stiell IG. Risk factors for recurrent emergency department visits for hyperglycemia in patients with diabetes mellitus. *International Journal of Emergency Medicine*. 2017; 10(23): 1-8.** • **https://intjem.biomedcentral.com/track/pdf/10.1186/s12245-017-0150-y**

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	• "Factors independently associated with a recurrent hyperglycemia visit within 30 days included a previous hyperglycemia visit in the past month (odds ratio [OR] 3.5, 95% confidence interval [CI] 2.1–5.8)." • "**The results of this unique multicenter study may aid emergency physicians in making follow-up and disposition decisions for patients presenting with hyperglycemia, given the lack of evidence in this area at present**." • "Return ED visits for hyperglycemia represent negative outcomes for patients with diabetes and impact the healthcare system overall." • "**Patients who had already had a previous hyperglycemia visit within the past month were most at risk for a recurrent unplanned ED visit for hyperglycemia**. It is unsurprising that individuals who may utilize the ED for management of a recurring, chronic condition such as diabetes are likely to do so on an ongoing basis." • "A future prospective study examining a patient's ability to access follow-up with a healthcare provider and transition their care from emergency to primary care would help to determine if improved access to a family physician would lead to reduced unnecessary ED return visits for hyperglycemia. Indeed, studies of other chronic disease entities such as congestive heart failure or chronic obstructive pulmonary disease have demonstrated that access to follow-up is associated with a decreased 30-day risk of ED visits and readmission to hospital. "
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A
Provide all other grades and definitions from the evidence grading system.	N/A
Grade assigned to the **recommendation** with definition of the grade	N/A
Provide all other grades and definitions from the recommendation grading system.	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	"Due to the retrospective nature of this health records review, this study is limited by the data that were recorded on patient charts. It is possible that some patients in our study period were missed if the treating physician's final diagnoses did not include an ICD-10 code related to hyperglycemia, diabetic ketoacidosis, or hyperosmolar hyperglycemic state, particularly if they were perceived to have a more important diagnosis such as cardiac arrest, acute coronary syndrome, or stroke."
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.

What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	No new studies to change the conclusion

Overall follow-up

(1 NQF Report)

Table 19: NQF Report

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Emergency Department Transitions of Care: A Quality Measurement Framework: Final Report** • **National Quality Forum** • **August 30, 2017** • **National Quality Forum, Emergency Department Transitions of Care: A Quality Measurement Framework: Final Report, 2017, 1-59.** • **http://www.qualityforum.org/Publications/2017/08/Emergency_Department_Transitions_of_Care_-_A_Quality_Measurement_Framework_Final_Report.aspx**

Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	p. 6: "**After an ED visit, follow-up is a high-risk time** where patients may experience important gaps in care that may result in missed diagnoses, and potentially avoidable healthcare use, such as return ED visits." p. 11: "There is frequently a request by the ED provider for **follow-up within a specific time period** (e.g., two to three days). However, it is sometimes not clear whether the primary care physician has the capacity to re-evaluate the patient within that period, or whether there are sufficient resources for ongoing management of the patient's condition." p. 11-12: "The Panel identified several key information elements that may be transmitted during ED transitions including: **Follow-up plan of care** – As a patient transitions out of the ED, EDs should communicate explicit follow-up plans with the patient and receiving provider with clear contingencies as the patient's condition evolves." p. 15: "Potential concepts to address gaps in the area of key information and its transmission and effective communication and shared decision making include: **Follow-up appointment scheduled for patients who lack a designated primary care provider.**" p. 18: "The Panel proposed 4 subdomains that incorporate these different perspectives: **4. Follow-Up and Safety Outcomes**: During a transition in care, the extent to which there are institutional processes to ensure appropriate care during the ED visit and **appropriate follow-up after the ED visit**" p. 19: "During the Panel's discussions, the following key themes were identified: **Patient follow-up including return visits to the ED**: In transitions of care, one of the important assessments is whether the patient received any necessary follow-up after discharge." p. 19: "Given the lack of outcome measures for transitions in care, the Panel focused on measure concepts that could fill identified measure gaps. Potential concepts to support gaps in this area include: • **Follow-up with patients after discharge from the ED**." Concepts/measures the Panel identified related to follow-up in Appendix C were: • #27: Existing Measure: Patients with a transient ischemic event ER visit that had a follow-up office visit (NQF #0644) [**EXISTING MEASURE]** • #28: Follow-up occurred after patient leaves the ED (e.g., visit or phone call) [**CONCEPT]** • #29: The percentage of high-risk ED discharges (as designated by the ED provider) where there is contact (in-person follow-up or other) within a specified period of time by the primary care physician or responsible specialist [**CONCEPT]**
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	N/A; not clinical guidelines but an NQF recommendation report
Provide all other grades and definitions from the evidence grading system.	N/A

Grade assigned to the **recommendation** with definition of the grade	N/A; not clinical guidelines but an NQF recommendation report
Provide all other grades and definitions from the recommendation grading system.	N/A
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The primary purpose of the scan was to assist in development of the measurement framework, and to identify an initial set of measures and measure concepts to consider for inclusion in the framework. The scan identified a total of 136 measures and 42 measure concepts. NQF staff then sorted the measures by relevance: 29 measures were directly relevant to the ED, 30 measures were potentially relevant, 36 measures were indirectly relevant, and 41 measures were not relevant. Appendix C identified these measures/concepts under the various domains/subdomains; #27, 28, and 29 in Appendix C related to follow-up.
Estimates of benefit and consistency across studies	Estimates of benefits and consistencies across studies were not provided.
What harms were identified?	Harms were not identified.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	N/A; report completed the end of August 2017 and there are no new NQF reports of this type.

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

N/A as evidence for 1a.3 (Systematic Review of the Evidence) is provided.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

APPENDIX: Measure Logic Model

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

Follow-up care is an element of care coordination, which has recently been highlighted as a priority area within the CMS Quality Strategy goal of Eliminating Disparities. (1) Follow-up care allows providers to perform a number of important activities, including ensuring that patients understand and are adhering to their medication regimen, monitoring them for adverse events, and educating them to recognize warning signs. (2) To date, there are scant published studies or data, especially within the US, that provide aggregate rates of timely follow-up care after ED or hospitalization for the conditions considered in this measure, making it imperative to fill in the gap in the healthcare community.

The body of evidence described below suggests, broadly and for the specific conditions included in this measure, that failing to obtain timely follow-up care is empirically linked to higher readmission rates, which are known to be associated with increased costs and decreased patient satisfaction and are widely considered an indicator of sub-optimal patient care. Furthermore, in the attached evidence tables we provide clinical guidelines for each condition, recommending that patients receive follow-up care after inpatient or ED discharge within the respective timeframe set for each chronic condition in the measure.

The goal of this measure is to improve the quality of care provided and improve patient outcomes by incentivizing health plans (insurance products) to ensure patients receive appropriate follow-up care following acute exacerbations of chronic conditions. To achieve this goal, the measure will identify health plans that have significantly lower rates of appropriate follow-up visits for acute conditions, relative to other health plans with the same acute conditions for similar patient populations. In doing so, this measure will prompt health plans to carefully evaluate care processes and implement quality improvement strategies. (3) Ultimately, this measure will provide an opportunity for health plans to become aware of and to improve rates of appropriate follow-up visits following acute events leading to decreased morbidity and mortality for patients with any of the 6 conditions covered by this measure.

By incentivizing health plans (insurance products) to improve follow-up rates for these conditions, this measure will improve outcomes as demonstrated in the following logic model:

1. Patient presents to the hospital/ED with an acute exacerbation of 1 of the 6 conditions included in the measure.
2. Patient is treated (either admitted to the hospital or treated only in the emergency department) and discharged to the community.
3. Health plan (insurance products) encourages follow-up visit/care through strategies such as incentives to providers,* reminders to patients, providing data/reports and continuing education to providers, etc.
4. Patient receives follow-up visit/care based upon evidence-based clinical guidelines, conditions are appropriately managed, with improved patient health and function.
5. ED and hospital cost and utilization are reduced by preventing avoidable readmissions.

* Note: If a health plan provides bonuses or other financial incentives to providers based on quality performance, an unintended consequence of holding plans accountable could be that providers are financially

penalized if the plan performs poorly. However, providers, in this case, have every incentive to help patients receive timely follow-up care, which could render the net effects being positive.

This logic model is made available in visual format as an appendix to the evidence form.

In general, the evidence demonstrates the presence of a link between timely follow-up care after discharge from the hospital/ED and lower rates of readmission. (4, 5) Specifically, the evidence also supports the link between follow-up care and improved outcomes for each of the conditions included in this measure, as described below.

• Asthma: A systematic review states that evidence from randomized controlled trials (RCTs) and non-RCTs suggests that timely follow-up with specialists reduces subsequent asthma exacerbations, as well as fewer symptoms and improved quality of life. (6) A population-based study on 7,829 patients with asthma or COPD found that follow-up visits within 30 days of discharge were associated with a significant reduction in the 90-day risk of an emergency readmission (RR=.79). (7) In a study of Medicare beneficiaries with COPD, 10% of whom also had asthma as a primary diagnosis code, the 30-day rate of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. Within the study population, patients who had a timely follow-up visit had a significantly reduced risk of ED visit (HR=.86) and readmission (HR=.91). (8)

• Heart Failure: Outpatient follow-up from a cardiology or general medicine provider within 7 days for a heart failure patient after hospitalization is associated with a lower chance of 30-day readmission (OR=.81). (9) Delayed outpatient follow-up after myocardial infarction has been associated with worse short-term and long-term patient medication adherence. Logically, this will add to morbidity and very likely both mortality and subsequent cardiovascular events/subsequent hospital admissions. (10) In the United States, the Hospital to Home (H2H) program, a national quality improvement initiative, also recommends that all heart failure and myocardial infarction patients have a follow-up appointment or cardiac rehab referral scheduled within 7 days of discharge. (11) Furthermore, the H2H program implemented a structured improvement project called the "See You in 7" initiative at 10 hospitals in Southeast Michigan, which included follow-up within a week of discharge as a core concept. After one year of participation, the adjusted 30-day readmission rates at collaborating hospitals decreased compared to non-participating hospitals (2.6% decrease vs. 0.6% decrease). (12) Heart failure and myocardial infarction patients in Taiwan were found to have a lower risk of 30-day readmission if they received outpatient visit with a physician within 7 days of discharge (HR=.54). (13) In a study of 3,136 patients, those who received cardiovascular follow-up saw fewer ED (38% vs. 80%) and hospital (13% vs. 94%) readmissions for cardiovascular reasons within the year, and lower unadjusted mortality (7% vs. 2% at 30 days). (14) In a study of 30,136 patients, patients discharged from hospitals with higher rates of follow-up within 7 days of discharge for heart failure had a lower risk of 30-day readmission (HR .91 between highest and lowest quartiles). (15)

While direct evidence on the relationship between timely follow-up and CAD and hypertension is extremely limited, both conditions are well recognized comorbidities and predictors of heart failure. (16, 17, 18) Logically, because a strong body of evidence links timely follow-up after heart failure to lower readmissions, timely follow-up for CAD and hypertension can also be reasoned to lower readmissions.

• COPD: In a study of Medicare beneficiaries with COPD, 10% of whom also had asthma as a primary diagnosis code, the 30-day rates of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. Within the study population, patients who had a timely follow-up visit had a significantly reduced risk of ED visit (HR=.86) and readmission (HR=.91). (8) A population-based study of 7829 patients with asthma or COPD found that follow-up visits within 30 days of discharge were associated with a significant reduction in the 90-day risk of an emergency readmission (RR=.79). (7)

• Diabetes: A review article suggests that better management of hyperglycemia in the ED, along with proper follow-up to ensure continued appropriate management, improves clinical outcomes and prevents

readmission. (19) Another study found that sentinel ED visits in diabetic patients are a warning sign for future readmission for hyperglycemia, and recommended that clinicians "provide clear discharge instructions for follow-up and glucose management to prevent further hyperglycemic emergencies from occurring." (20)

The gap in care described above represents one reason these particular 6 conditions were selected for inclusion in the measure. These measures were also selected because they are ambulatory care-sensitive conditions that respond well to timely primary care, as reflected in broad body of evidence (described in the evidence form) linking timely follow-up for these conditions to improved health outcomes. Additionally, these conditions were selected based on the impact they have on patients and health systems, measured in both condition prevalence as well as the costs and resources associated with appropriate treatment. For example, the CDC finds that approximately 30 million Americans have diabetes, and complications were estimated to cost the US about $245 billion in 2012. (21) Nearly 27 million Americans are living with asthma, which is responsible for about 440,000 inpatient discharges, as well as 1.7 million ED visits and 11 million physician's office visits per year, with a total economic burden estimated at about $53 billion. (22, 23)

The prevalence of hypertension is even higher, with about 75 million adults (1 in 3) suffering from high blood pressure and an estimated $50 billion in economic burden. (24) About 15.7 million Americans suffer from COPD, with mortality rates as high as 62.8 per 100,000 in Kentucky and an overall economic burden of nearly $50 billion. (25, 26) About 15 million Americans suffer from CAD, and about 6 million suffer from heart failure. Heart disease remains the leading cause of death in the United States and represents an immense burden on the health system as a whole and on patients and their families individually. (27, 28)

While the prevalence of these 6 conditions is not uniformly high, they incur considerable cost to health systems as well as individual patients, and are conditions that represent degraded patient health and risk of mortality. The following statistics summarize incidence and cost data gathered from the 2014 Healthcare Cost and Utilization Project. Please note that ED cost data were not available.

Hypoglycemia

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.92
  - ED Rate per 100,000 persons: 8.06
  - Estimate of Mean Inpatient Costs: $7,177
- Any DX
  - Inpatient Rate per 100,000 persons: 12.55
  - ED Rate per 100,000 persons: 28.75

Hyperglycemia

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.16
  - ED Rate per 100,000 persons: 10.06
  - Estimate of Mean Inpatient Costs: $4,595
- Any DX
  - Inpatient Rate per 100,000 persons: 21.4
  - ED Rate per 100,000 persons: 509.26

Diabetic Ketoacidosis

- Principal DX
  - Inpatient Rate per 100,000 persons: 45.79
  - ED Rate per 100,000 persons: 48.77

- Estimate of Mean Inpatient Costs: $7,280
- Any DX
  - Inpatient Rate for 100,000 persons: 59.75
  - ED Rate per 100,000 persons: 63.29

Acute Myocardial Infarction

- Principal DX
  - Inpatient Rate per 100,000 persons: 69.28
  - ED Rate per 100,000 persons: 64.31
  - Estimate of Mean Inpatient Costs: $21,380
- Any DX
  - Inpatient Rate per 100,000 persons: 95.79
  - ED Rate per 100,000 persons: 87.13

Angina

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.44
  - ED Rate per 100,000 persons: 0.59
  - Estimate of Mean Inpatient Costs: $7,693
- Any DX
  - Inpatient Rate per 100,000 persons: 1.61
  - ED Rate per 100,000 persons: 1.93

Acute Asthma Attack

- Principal DX
  - Inpatient Rate per 100,000 persons: 68.60
  - ED Rate per 100,000 persons: 462.98
  - Estimate of Mean Inpatient Costs: $6,555
- Any DX
  - Inpatient Rate per 100,000 persons: 471.03
  - ED Rate per 100,000 persons: 1824.8

Hypotension

- Principal DX
  - Inpatient Rate per 100,000 persons: 8.25
  - ED Rate per 100,000 persons: 21.06
  - Estimate of Mean Inpatient Costs: $7,088
- Any DX
  - Inpatient Rate per 100,000 persons: 156.05
  - ED Rate per 100,000 persons: 143.97

Shortness of Breath

- Principal DX
  - Inpatient Rate per 100,000 persons: 1.04

- ED Rate per 100,000 persons: 61.44
- Estimate of Mean Inpatient Costs: $5,486

- Any DX
  - Inpatient Rate per 100,000 persons: 17.18
  - ED Rate per 100,000 persons: 386.76

Acute Pulmonary Edema

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.67
  - ED Rate per 100,000 persons: 0.88
  - Estimate of Mean Inpatient Costs: $9,011
- Any DX
  - Inpatient Rate per 100,000 persons: 5.51
  - ED Rate per 100,000 persons: 4.35

Peripheral Edema

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.7
  - ED Rate per 100,000 persons: 47.12
  - Estimate of Mean Inpatient Costs: $6,468
- Any DX
  - Inpatient Rate per 100,000 persons: 40.48
  - ED Rate per 100,000 persons: 128.07

Acute Left Ventricular Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 0.07
  - ED Rate per 100,000 persons: 0.08
  - Estimate of Mean Inpatient Costs: $8,839
- Any DX
  - Inpatient Rate per 100,000 persons: 0.78
  - ED Rate per 100,000 persons: 1.23

Pleural Effusion

- Principal DX
  - Inpatient Rate per 100,000 persons: 4.29
  - ED Rate per 100,000 persons: 7.31
  - Estimate of Mean Inpatient Costs: $12,562
- Any DX
  - Inpatient Rate per100,000 persons: 65.46
  - ED Rate per 100,000 persons: 60.93

Hypertension

- Principal DX

- Inpatient Rate per 100,000 persons: 7.33
- ED Rate per 100,000 persons: 138.68
- Estimate of Mean Inpatient Costs: $5,699

- Any DX
  - Inpatient Rate per 100,000 persons: 1280.37
  - ED Rate per 100,000 persons: 3710.17

Chronic Kidney Disease and Systolic Heart Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 2.54
  - ED Rate per 100,000 persons: 6.81
  - Estimate of Mean Inpatient Costs: $19,507
- Any DX
  - Inpatient Rate per 100,000 persons: 366.59
  - ED Rate per 100,000 persons: 440.07

Pulmonary Congestion and Hypostasis

- Principal DX
  - Inpatient Rate per 100,000 persons: .63
  - ED Rate per 100,000 persons: 1.62
  - Estimate of Mean Inpatient Costs: $7,724
- Any DX
  - Inpatient Rate per 100,000 persons: 11.10
  - ED Rate per 100,000 persons: 11.68

Congestive Heart Failure

- Principal DX
  - Inpatient Rate per 100,000 persons: 13.67
  - ED Rate per 100,000 persons: 33.86
  - Estimate of Mean Inpatient Costs: $11,374
- Any DX
  - Inpatient Rate per 100,000 persons: 353.30
  - ED Rate per 100,000 persons: 462.81

COPD

- Principal DX
  - Inpatient Rate per 100,000 persons: 70.49
  - ED Rate per 100,000 persons: 407.92
  - Estimate of Mean Inpatient Costs: $8,019
- Any DX
  - Inpatient Rate per 100,000 persons: 467.86
  - ED Rate per 100,000 persons: 1119.54

CITATIONS:

1) Centers for Medicare and Medicaid Services. CMS Quality Strategy. 2016. Available from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/Downloads/CMS-Quality-Strategy.pdf

2) California Healthcare Foundation. The Post-Hospital Follow-Up Visit: A Physician Checklist to Reduce Readmissions. October 2010. Available from https://www.chcf.org/wp-content/uploads/2017/12/PDF-PostHospitalFollowUpVisit.pdf.

3) Chassin M, et al. Accountability Measures – Using Measurement to Promote Quality Improvement. N Engl J Med 2010; 363:683-688. doi: 10.1056/NEJMsb1002320.

4) Misky GJ, et al. Post-hospitalization transitions: Examining the effects of timing of primary care provider follow-up. J Hosp Med. 2010 Sep;5(7):392-7. doi: 10.1002/jhm.666.

5) Shen E, et al. Association of a Dedicated Post–Hospital Discharge Follow-up Visit and 30-Day Readmission Risk in a Medicare Advantage Population. JAMA Intern Med. 2017;177(1):132-135. doi:10.1001/jamainternmed.2016.7061.

6) Schatz M, et al. Follow-Up After Acute Asthma Episodes: What Improves Future Outcomes? J Emerg Med. 2009 Aug;37(2 Suppl):S42-50. doi: 10.1016/j.jemermed.2009.06.109.

7) Sin D et al. The impact of follow-up physician visits on emergency readmissions for patients with asthma and chronic obstructive pulmonary disease: a population-based study. Am J Med. 2002 Feb 1;112(2):120-5. doi: 10.1016/S0002-9343(01)01079-8.

8) Sharma G, et al. Outpatient Follow-up Visit and 30-Day ED Visit and Readmission in Patients Hospitalized for Chronic Obstructive Pulmonary Disease. Arch Intern Med. 2010 Oct 11; 170(18): 1664–1670. doi: 10.1001/archinternmed.2010.345.

9) Lee K, et al. Post-Discharge Follow-up Characteristics Associated with 30-Day Readmission After Heart Failure Hospitalization. Med Care. 2016 Apr; 54(4): 365–372. doi: 10.1097/MLR.0000000000000492.

10) Faridi KF, et al. Timing of First Postdischarge Follow-up and Medication Adherence After Acute Myocardial Infarction. JAMA Cardiol. 2016 May 1;1(2):147-55. doi: 10.1001/jamacardio.2016.0001.

11) Hospital 2 Home: Improving Transitions from the Hospital to Community Settings. 2013. Slide Deck. Available from https://cvquality.acc.org/docs/default-source/h2h/1-h2h-overview-slides.pdf?sfvrsn=c6d58fbf_2.

12) Baker H, et al. Regional Hospital Collaboration and Outcomes in Medicare Heart Failure Patients: See You in 7. JACC Heart Fail. 2015 Oct;3(10):765-73. doi: 10.1016/j.jchf.2015.06.007.

13) Tung YC, et al. Relationship between Early Physician Follow-Up and 30-Day Readmission after Acute Myocardial Infarction and Heart Failure. PLoS One. 2017 Jan 27;12(1):e0170061. doi: 10.1371/journal.pone.0170061.

14) Ezekowitz JA, et al. Impact of specialist follow-up in outpatients with congestive heart failure. CMAJ. 2005 Jan 18;172(2):189-94. doi: 10.1503/cmaj.1032017.

15) Hernandez AF, et al. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. JAMA. 2010 May 5;303(17):1716-22. doi: 10.1001/jama.2010.533.

16) Velagaleti R, et al. Heart Failure in the 21st Century: Is it a Coronary Artery Disease Problem or Hypertension Problem? Cardiol Clin. 2007 Nov; 25(4): 487–v. doi: 10.1016/j.ccl.2007.08.010.

17) Lala A, et al. The role of coronary artery disease in heart failure. Heart Fail Clin. 2014 Apr;10(2):353-65. doi: 10.1016/j.hfc.2013.10.002.

18) Messerli F, et al. The Transition from Hypertension to Heart Failure. JACC: Heart Failure Aug 2017, 5 (8) 543-551; DOI: 10.1016/j.jchf.2017.04.012.

19) Echouffo-Tcheugui J, et al. Management of Hyperglycemia and Diabetes in the ED. Curr Diab Rep (2017) 17: 56. https://doi.org/10.1007/s11892-017-0883-2.

20) Yan JW, Gushulak KM, Columbus MP, Hamelin AL, Wells GA, Stiell IG. Sentinel visits in ED patients with diabetes mellitus as a warning sign for hyperglycemic emergencies. CJEM. 2017 Jul 25:1-8. doi: 10.1017/cem.2017.338.

21) ADA. Economic Costs of Diabetes in the U.S. in 2012. Diabetes Care. 2013 Apr; 36(4): 1033–1046.

22) Asthma and Allergy Foundation of America. Costs of Asthma on Society. Available from http://www.aafa.org/page/cost-of-asthma-on-society.aspx.

23) Centers for Disease Control and Prevention. Most Recent Asthma Data. May 2018. Available from https://www.cdc.gov/asthma/most_recent_data.htm.

24) Centers for Disease Control and Prevention. High Blood Pressure. Available from https://www.cdc.gov/bloodpressure/index.htm.

25) Guarascio A, et al. The clinical and economic burden of chronic obstructive pulmonary disease in the USA. Clinicoecon Outcomes Res. 2013; 5: 235–245. Published online 2013 Jun 17. doi: 10.2147/CEOR.S34321.

26) Centers for Disease Control and Prevention. Chronic Obstructive Pulmonary Disease (COPD). Available from https://www.cdc.gov/copd/data.html.

27) Sanchis-Gomar F, et al A. Epidemiology of coronary heart disease and acute coronary syndrome. Ann Transl Med. 2016 Jul; 4(13): 256. doi: 10.21037/atm.2016.06.33.

28) Roger V. Epidemiology of Heart Failure. Circ Res. 2013 Aug 30; 113(6): 646–659. doi: 10.1161/CIRCRESAHA.113.300268.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

N/A. This measure has not yet been implemented. However, the performance of entities within our testing sample sets is detailed in the testing form.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

There are relatively few published studies or data, especially within the US, that provide aggregate rates of follow-up visits from the ED or hospitalizations for the conditions in this measure—meaning that there are limited data to outline and quantitatively assess the gaps in care or opportunities for improvement in follow-up care for these patients, as it has not been empirically studied to date. But despite a connection between timely follow-up after acute events and lower readmissions and other positive outcomes, and recommendations in multiple sets of clinical guidelines, the available data point to a gap in care: on a broad level, one study found that overall, 32% of 300 discharged ED patients who were specifically believed to be at risk for clinical deterioration did not receive timely follow-up care as instructed. More than one third of those who did not obtain timely follow-up care reported the inability to obtain an appointment as the primary barrier. (1)

A population-based study on 7,829 patients found that only 31% of patients admitted to the ED for asthma or COPD received follow-up care within 30 days of discharge. (2) In a study of 62,746 patients hospitalized or admitted to the ED with COPD, approximately 33% of patients did not receive a follow-up visit within 30 days. (3)

In a study examining the use of inhaled corticosteroids in asthma patients, only 40% of 414 patients admitted to the ED for asthma exacerbations received timely follow-up visit with a primary care physician. (4) In a study

of Medicare beneficiaries with COPD, 10% of whom also had asthma, the 30-day rates of post-discharge ER visits in patients with timely follow-up to their PCP or pulmonologist was 21.7%, compared with 26.3% in those with no post-discharge follow-up. (3) Additionally, much of the body of evidence on asthma follow-up rates is focused on children. These studies found consistently low rates of follow-up care—such as a study of 3,435 patients ages 2-18, of whom only 12% completed a follow-up visit within two months of visiting the ED for asthma, and another study finding that only a third of children within a cohort of 561 received follow-up visit within 30-days of an ED visit for asthma. (5, 6)

A study of 30,136 Medicare patients from 255 hospitals admitted for heart failure found that the median frequency of follow-up within 2 weeks of discharge was 65%. (7)

In a study of 1,848 patients in Sweden with chest pain and elevated levels of highly sensitive cardiac troponin (hs-cTnT) levels and no myocardial infarction, patients rarely scheduled follow-up after inpatient discharge, suggesting that incentives can be embedded into health plans (insurance products) to address this type of inertia (8) Among 56,767 high-risk patients presenting to an Ontario ED with chest pain, approximately 20% did not receive any follow-up by a physician within 30 days. This sample included patients with a history of angina and other cardiovascular conditions. Surprisingly—and problematically—patients at the highest risk for future events by virtue of significant comorbid disease were even less likely to receive follow-up care. (9)

Within a primarily fee-for-service environment in Ontario, a study of 41,485 ED visits for heart failure, atrial fibrillation, and hypertension found a follow-up visit rate of 37% within 7 days. (10) In a previously cited study of heart failure, between 70%-75% of patients in the study population had a history of hypertension, and the median rate of patients who had follow-up care within 7 days (the recommended timeframe for hypertension) in that study was found to be only 38%. (7)

In an ED study of 5,317 in the Netherlands, less than 2% of hyperglycemic patients had follow-up arranged by the treating physician. (11) In another study of 221 patients within the UK, only half of patients with diabetes-related diagnoses had documented timely follow-up arrangements after discharge from the ED. (12) While our measure is focused on the actual reception of a timely follow-up visit, as opposed to simple documentation of a follow-up plan, we still believe the latter is a valuable lower bar indicator.

This literature demonstrating a performance gap is strongly reinforced by our own testing, which involved data from both private Qualified Health Plan issuers as well as the entire Medicare Advantage universe. This testing represents empirical evidence of our own with convergent findings from the literature: low rates of timely follow-up care that justify a measure incentivizing health plans (insurance products) to improve follow-up rates. The data are included in detail in the testing form, but the results from Medicare Advantage and one QHP issuer are listed below as a point of reference (QHP data not available for 2016):

Medicare Advantage and QHP IMPAQ Testing Summary-Gaps in Care (Please note that these data indicate the percentage of patients who received follow-up care within the appropriate timeframe for that specific condition, separated out by testing population.)

Asthma:

2014: MA (52.8%) QHP (33%)

2015: MA (51.4%) QHP (32%)

2016: MA (52.8%) QHP (N/A)

CAD:

2014: MA (63.1%) QHP (63%)

2015: MA (62.3%) QHP (62%)

2016: MA (64.6%) QHP (N/A)

Heart Failure:

2014: MA (65.5%) QHP (71%)

2015: MA (63.8%) QHP (67%)

2016: MA (66.1%) QHP (N/A)

COPD:

2014: MA (72.4%) QHP (45%)

2015: MA (70.8%) QHP (47%)

2016: MA (74.3%) QHP (N/A)

Diabetes:

2014: MA (73.6%) QHP (64%)

2015: MA (72.2%) QHP (59%)

2016: MA (76%) QHP (N/A)

Hypertension:

2014: MA (42%) QHP (38%)

2015: MA 40.9%) QHP (38%)

2016: MA (42.1%) QHP (N/A)

CITATIONS:

1) Barlas D, Homan CS, Rakowski J, Houck M, Thode HC. How well do patients obtain short-term follow-up after discharge from the emergency department? Ann Emerg Med. 1999 Nov;34(5):610-4.doi: 10.1016/S0196-0644(99)70163-0.

2) Sin D, Bell N, Svenson L, Man SF. The impact of follow-up physician visits on emergency readmissions for patients with asthma and chronic obstructive pulmonary disease: a population-based study. Am J Med. 2002 Feb 1;112(2):120-5. doi: 10.1016/S0002-9343(01)01079-8.

3) Sharma G, Kuo YF, Freeman J, Zhang D, Goodwin J. Outpatient Follow-up Visit and 30-Day Emergency Department Visit and Readmission in Patients Hospitalized for Chronic Obstructive Pulmonary Disease. Arch Intern Med. 2010 Oct 11; 170(18): 1664–1670. doi: 10.1001/archinternmed.2010.345.

4) Cydulka RK, Tamayo-Sarver JH, Wolf C, Herrick E, Gress S. Inadequate follow-up controller medications among patients with asthma who visit the emergency department. Ann Emerg Med. 2005 Oct;46(4):316-22.

5) Andrews AL, Teful RJ, Basco WT. Low rates of controller medication initiation and outpatient follow-up after emergency department visits for asthma. J Pediatr. 2012 Feb;160(2):325-30. doi: 10.1016/j.jpeds.2011.07.037.

6) Cabana MD, Bruckman D, Bratton Sl, Kemper AR, Clark NM. Association Between Outpatient Follow-Up and Pediatric Emergency Department Asthma Visits. J Asthma. 2003;40(7):741-9. doi: 10.1081/JAS-120023499.

7) Hernandez AF, Greiner MA, Fonarow GC, Hammil BG, Heidenreich PA, Yancy CW, Peterson ED, Curtis LH. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. JAMA. 2010 May 5;303(17):1716-22. doi: 10.1001/jama.2010.533.

8) Roos A, Hellgren A, Rafatnia F, Hammarsten O, Ljung R, Carlsson AC, Holzmann MJ. Investigations, findings, and follow-up in patients with chest pain and elevated high-sensitivity cardiac troponin T levels but no myocardial infarction. Int J Cardiol. 2017 Apr 1;232:111-116. doi: 10.1016/j.ijcard.2017.01.044.

9) Wong M, Wang J, Czarnecki A, Koh M, Tu J, Schull M, Wijeysundera H, Lau C, Ko D. Factors associated with physician follow-up among patients with chest pain discharged from the emergency department. CMAJ. 2015 Mar 17; 187(5): E160–E168. doi: 10.1503/cmaj.141294.

10) Atzema CL, Yu B, Ivers NM, Rochon PA, Lee DS, Schull MJ, Austin PC. Predictors of obtaining follow-up care in the province of Ontario, Canada, following a new diagnosis of atrial fibrillation, heart failure, and hypertension in the emergency department. CJEM. 2018 May;20(3):377-391. doi: 10.1017/cem.2017.371.

11) van Veggel KM, Kruithof MK, Roelandse-Koop E, Eekhoff EM, Nanayakkara PW. Follow-up of patients with hyperglycemia in the emergency department without a history of diabetes mellitus. Eur J Intern Med. 2014 Dec;25(10):909-13. doi: 10.1016/j.ejim.2014.11.003.

12) Brackenridge A, Wallbank H, Lawrenson RA, Russell-Jones D. Emergency management of diabetes and hypoglycaemia. Emerg Med J. 2006 Mar; 23(3): 183–185. doi: 10.1136/emj.2005.026252.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

To evaluate disparities among populations, the team calculated numerator and denominator counts stratified by the following characteristics within the MA encounter data:

- Gender
- Low-income subsidy (LIS) status (as the proxy for socioeconomic status available in MA encounter data)
- Age

To determine if any differences are statistically significant, the team produced two statistics: the p-value (using a p-value of <0.05 as significant) and Cramer's V.

Because p-values are calculated with estimates of the standard error—which by definition decrease as sample sizes increase—with large enough sample sizes, even minuscule effects can become "statistically significant." The team has found that MA encounter data sample sizes are large enough for this to be an issue. As such, we also provide estimates of Cramer's V, which measures on a 0 to 1 scale the strength of the association between two variables (such as gender and measure score). If a population difference is "statistically significant" according to the p-value, but has a Cramer's V near 0, then the association is negligible. If it is statistically significant and has a Cramer's V close to 1, then the measure score is associated with the population characteristic of interest.

Our testing for each variable (gender, LIS status, age) did not reveal a Cramer's V greater than .041 for any variable in any year of testing data (2014-2016). From these results, we conclude that these subpopulations are not experiencing disparities in follow-up rates.

However, we acknowledge that these results may differ from the existing literature and we present three explanations that could account for the discrepancy. First, early studies that documented disparities in follow-up care across sub-populations may have relied on data that are outdated. Second, compared to the early work that found disparity using data at the micro-level, the current measure is defined at the issuer-by-product level, and lack of granularity could further account for the discrepancy. Third, we stress the importance of heterogeneity of the measures themselves even if the data being used share similarities. For instance, a recent study conducted by RAND (1) on follow-up care after hospital stays for mental illness in the Medicare Advantage program suggested differences between men and women. Our measure, on the other hand, documents timely follow-ups for 6 chronic conditions (hypertension, asthma, heart failure, coronary artery disease, chronic obstructive pulmonary disease, and diabetes) that differ substantively from mental health challenges. While disparities in some aspects of the care do not necessarily translate into disparities in others, we recommend that performance of subgroups (gender, age, LIS) continue to be monitored to ensure that high-quality care is received equitably by all groups.

CITATIONS

1) CMS Office of Minority Health in collaboration with the RAND Department. Racial, Ethnic, and Gender Disparities in Health Care in Medicare Advantage. April 2018. Web.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

N/A

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

De.6. Non-Condition Specific*(check all the areas that apply):*

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

http://impaqint.com/measure-information-timely-follow-after-acute-exacerbations-chronic-conditions

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **Attachment:** Value_Set_7-25-2018.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

No

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

This measure is not being considered for maintenance of endorsement.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The numerator is the sum of the issuer-product-level denominator events (Emergency Room [ED], observation hospital stay or inpatient hospital stay) for acute exacerbation of hypertension, asthma, heart failure (HF), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), or diabetes where follow-up was received within the timeframe recommended by clinical practice guidelines, as detailed below:

- Hypertension: Within 7 days of the date of discharge
- Asthma: Within 14 days of the date of discharge
- HF: Within 14 days of the date of discharge
- CAD: Within 14 days of the date of discharge
- COPD: Within 30 days of the date of discharge
- Diabetes: Within 30 days of the date of discharge

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

This measure is defined at the issuer-by-product level, meaning that results are aggregated for each qualified insurance issuer and for each product. For clarity, a product is a discrete package of health insurance coverage benefits that issuers offer in the context of a particular network type, such as health maintenance organization (HMO), preferred provider organization (PPO), exclusive provider organization (EPO), point of service (POS), or indemnity. Issuers are broadly defined as health insurance providers who participate in the Federally-facilitated Marketplaces and health insurance contracts offered in the Medicare Advantage market.

Timely follow-up is defined as a claim for the same patient after the discharge date of the acute event that is a non-emergency outpatient visit and has a CPT or HCPCS code indicating a visit that constitutes appropriate follow-up, as defined by clinical guidelines and clinical coding experts. The follow-up visit may be a general office visit or telehealth and take place in certain chronic care or transitional care management settings. The follow-up visit must occur within the condition-specific timeframe to be considered timely and for the conditions of the numerator/measure to be met. For a list of individual codes, please see the data dictionary attached in S.2b.

The follow-up visit timeframes for each of the 6 chronic conditions are based on evidence-based clinical practice guidelines (CPGs) as laid out in the evidence form.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

The denominator is the sum of the plan-product-level acute exacerbations that require either an ED visit, observation stay, or inpatient stay (i.e., acute events) for any of the six conditions listed above (hypertension, asthma, HF, CAD, COPD, or diabetes).

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Acute events are defined as either an ED visit, observation stay, or inpatient stay. If a patient is discharged and another claim begins for the same condition on the same day or the following day, the claims are considered

to be part of one continuous acute event. In this case, the discharge date of the last claim is the beginning of the follow-up interval. The final claim of the acute event must be a discharge to community.

An acute event is assigned to [condition] if:

1. The primary diagnosis is a sufficient code for [condition].

OR

2. The primary diagnosis is a related code for [condition] AND at least one additional diagnosis is a sufficient code for [condition].

a. In cases where the event has two or more conditions with a related code as the primary diagnosis and a sufficient code in additional diagnosis positions, assign the event to the condition with a sufficient code appearing in the "highest" (closest to primary) diagnosis position.

If the visits that make up an acute event are assigned different conditions, the event is assigned the condition that occurs last in the sequence. Following this methodology, only one condition is recorded in the denominator per acute event. For a list of individual codes, please see the data dictionary attached in S.2b.

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

The measure excludes events with:

1. Subsequent acute events that occur two days after the prior discharge, but still during the follow-up interval of the prior event for the same reason. To prevent double-counting, only the first acute event will be included in the denominator.
2. Acute events after which the patient does not have continuous enrollment for 30 days in the same product.
3. Acute events where the discharge status of the last claim is not "to community" ("Left against medical advice" is not a discharge to community.)
4. Acute events for which the calendar year ends before the follow-up window ends (e.g., acute asthma events ending fewer than 14 days before December 31)
5. Acute events where the patient enters a skilled nursing facility (SNF), non-acute care, or hospice care within the follow-up interval

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

For a list of individual codes, please see the data dictionary attached in S.2b.

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

We do not see a need to stratify. Since follow-up visits for ED or hospitalization have a strong correlation with better outcomes, this measure is designed to encourage health plans (insurance products) to improve follow-up of their patients of every socioeconomic status equally.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

1) Denominator events are identified by hospitalization, observation, and ED events with appropriate codes (i.e., codes identifying an acute exacerbation of 1 of the 6 included chronic conditions).

2) Exclusions are applied to the population from step 1) to produce the eligible patient population for the measure (i.e., the count of all qualifying events).

3) For each qualifying event, it is determined whether or not claims included a subsequent code that satisfies the follow-up requirement for that particular qualifying event (e.g., a diabetes event received follow-up within the appropriate timeframe for diabetes, from an appropriate provider). Each event for which the follow-up requirement was satisfied is counted as 'one' in the numerator. Each event for which the follow-up requirement was not satisfied is counted as a 'zero' in the numerator.

4) The percentage score is calculated as the numerator divided by the denominator.

Measure Scoring Logic

Following NQF's guideline, we employ Opportunity-Based Weighting to calculate the follow-up measure. (1) This means that each condition is weighted by the sum of acute exacerbations that require either an ED visit or an observation or inpatient stay for all the six conditions that occur, as reflected in the logic below.

[NUM(ASM) + NUM(CAD) + NUM(HF) + NUM (COPD) + NUM(DIAB) + NUM(HTN)] / [DENOM(ASM) + DENOM(CAD) + DENOM(HF) + DENOM (COPD) + DENOM(DIAB) + DENOM(HTN)]

***Please note that, while the development team designed the measure to aggregate each condition score in the manner described above into a single overall score, programs may choose to also calculate individual scores for each chronic condition when implementing the measure. Individual measure scores would simply be calculated by dividing the condition-specific numerator by the condition specific denominator, as in the example for heart failure below:

NUM(HF) / DENOM(HF)

Both methods capture the same quality information, with different levels of granularity. Below is an example of each scoring method:

Aggregate: 30 patients experience acute events. 25 events are heart failure, 5 events are COPD. Of these 30 patients, 25 receive appropriate follow-up. The measured entity receives a score of 83% (25/30).

Individual: The same 30 patients experience acute events. 25 events are for heart failure. 5 events are for COPD. 25 receive appropriate follow-up. This number included 20 of the patients who experienced heart failure, and all 5 patients who experienced COPD. The measured entity receives a heart failure score of 80% (20/25) and a COPD score of 100% (5/5).

The team considered several aggregation methods, including uniform weighting, opportunity-based weighting, and linear combination weighting for this measure. Each option has associated advantages and disadvantages.

The measure development team believes that opportunity-based weighting, described earlier in this section, is the best aggregation method for several reasons. First, sample sizes are relatively small, so rates for particular conditions may have high variance and produce erratic results. Second, with uniform weights (meaning each

condition's score contributes an equal amount to the overall score regardless of the number of events per condition), a change in the number of follow-ups for less prevalent conditions affects the aggregate score more than changing the number of follow-ups for more prevalent conditions. This gives an incentive to plans (insurance products) to focus on improving follow-up for the least prevalent conditions in order to improve their score., In contrast, opportunity-based weighting incentivizes plans to improve the number of follow-ups for each type of condition, because any penalty associated with the reduction in follow-ups of any condition is a function of the measure as a whole. (2) Furthermore, because there is no evidence that follow-ups for some of the 6 conditions are more important than others, opportunity-based weighting represents the simplest, fairest, and most easily interpretable and implementable weighting option for managed care organizations. There was no compelling evidence or rationale to use another, more complex weighting method.

It is important to note that this measure, while specified at the issuer-product-level and written to be applicable to various CMS payment programs, will still be required to go through a separate process to be fully operationalized into specific payment programs. These processes include publishing the measure in a Call Letter, soliciting public comment, and other activities to ensure the measure is appropriate for a given program.

1) National Quality Forum. Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures. 2009. Available from https://www.qualityforum.org/Publications/2009/08/Composite_Measure_Evaluation_Framework_and_National_Voluntary_Consensus_Standards_for_Mortality_and_Safety%E2%80%94Composite_Measures.aspx.

2) Shwartz, M., Restuccia, J. D., & Rosen, A. K. (2015). Composite Measures of Health Care Provider Performance: A Description of Approaches. The Milbank Quarterly, 93(4), 788–825. http://doi.org/10.1111/1468-0009.12165

Please note that the specifications of this measure have been slightly altered from what was submitted in the Intent to Submit form. These minor changes are intended to increase clarity.´

Citations:

1) National Quality Forum. Composite Measure Evaluation Framework and National Voluntary Consensus Standards for Mortality and Safety—Composite Measures. 2009. Available from https://www.qualityforum.org/Publications/2009/08/Composite_Measure_Evaluation_Framework_and_National_Voluntary_Consensus_Standards_for_Mortality_and_Safety%E2%80%94Composite_Measures.aspx.

2) Shwartz, M., Restuccia, J. D., & Rosen, A. K. (2015). Composite Measures of Health Care Provider Performance: A Description of Approaches. The Milbank Quarterly, 93(4), 788–825. http://doi.org/10.1111/1468-0009.12165

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

N/A – This measure is not based on a sample.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

N/A – This measure is not based on a survey or instrument.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Claims

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

The data source for this measure is administrative claims data. The specific claims data used will be dependent on the population being measured. For example, measurement in the Medicare Advantage population will use Medicare Advantage (MA) encounter data, while measurement in the Quality Rating System will use administrative claims data submitted by Qualified Health Plan issuers.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Health Plan, Other

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Emergency Department and Services, Inpatient/Hospital

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

NOTE: Given the nature of this measure, the development team carefully considered whether to classify it as a composite measure or a process measure. After discussing the question with NQF measurement staff, the development team believes this measure is most appropriately classified as a non-composite, complex process measure. The NQF measurement staff acknowledged the issue is nuanced, but agreed that the measure may be submitted as a process measure.

NQF defines a composite measure as a combination of two or more component measures, each of which individually reflects quality of care, into a single performance measure with a single score. We recognize that patients or other stakeholders may also want to see the scores within each condition if they are focused on a specific ailment. While our measure must be completed for all 6 conditions and reported as a single aggregate score, payment and quality measurement programs may also choose to report out the individual components (i.e., condition-level scores) for consumer, quality improvement, or other purposes.

Furthermore, composite measures are often scored by 'all-or-none' or 'any-or-none' scoring methodologies. Our measure does not employ either of these scoring methodologies. In addition, unlike existing NQF composite measures, all of the components or conditions covered by our measure do not need to be present for the measure to be calculated. Instead, our measure simply calculates the proportion score utilizing any of the 6 conditions that are present in the population being measured.

Our measure includes a form of opportunity-based weighting, simply meaning that the overall score is influenced by the number of patients who meet each of the possible denominator component criteria. The measure reports a single summary score, derived from the opportunity-weighted results of the 6 different chronic conditions, one for each of the following chronic conditions: COPD, asthma, CAD, HF, hypertension, and diabetes. An NQF consultant who has supported development of this measure concurs with the team that the numerator neither is an 'all-or-none' approach with multiple processes or outcomes being measured, nor includes other 'purposeful' weighting as one would expect in a composite measure; rather, it is weighted based on the number of patients who happen to meet the denominator criteria.

Finally, NQF composites tend to be aggregators of distinct indicators combined into one measure to indicate the overall quality of care for a certain domain. In contrast, our measure applies the exact same indicator to 6 different conditions.

2. Validity – See attached Measure Testing Submission Form

Follow-Up_Testing_Supplement_Form_FINAL.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Measure Title: Timely Follow-Up After Acute Exacerbations of Chronic Conditions
Date of Submission: 8/1/2018

Type of Measure:

☐ Outcome (*including PRO-PM*)	☐ Composite – ***STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b5** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this

form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing e demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **instrument-based measures** (including PRO-PMs) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. Validity testing f demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; g

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). h

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; i,j and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** 16 **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

e. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor

studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

f. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

g. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

h. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

i. Risk factors that influence outcomes should not be specified as exclusions.

j. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure*. *Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other:	☐ other:

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

CMS is considering the use of this measure in two plan measurement programs: Medicare Advantage (MA) Star Ratings and Health Insurance Exchange for Qualified Health Plans (QHPs). The Health Insurance Exchange for QHPs is also known as the Marketplace. CMS will hold a separate comment period to determine a formal process to operationalize this measure into interested programs. In these applications, the measured entities will be MA and QHP health insurance products.

The team selected MA data and QHP data for three reasons: (1) they are consistent with the target population and products measured in these target programs, (2) the six measured conditions [asthma, CAD, COPD, diabetes, heart failure, and hypertension] are prevalent in both the MA and QHP populations, and (3) testing across both of these data sets shows measure reliability and validity across age groups by using the MA data for the primarily 65 and over population and QHP data for the primarily under 65 population.

For testing, the team used four (4) data sets containing administrative claims:

- **One dataset of all MA encounter data**
  - This data is housed in the CMS Integrated Data Repository (IDR). This contains administrative encounter data for all MA plans.
- **Three Commercial and QHP data sets**
  - One QHP issuer's administrative claims data
  - One Qualified Entity's QHP claims database
  - One analytics firm's QHP and commercial claims database

The team used only one data set of MA data as the data set from CMS contained data from all MA products. Without a single data set containing QHP products available, the team obtained administrative claims data from three data sources. All of the selected QHP data sources contained administrative claims, but each provided additional measure entities (health insurance products/issuers). This aggregation ensures that a wider variety of measured entities (health insurance products/issuers) were included for testing and represent a larger geographical area. For simplicity, results for all products in the various QHP/Commercial data sources are aggregated under "QHP/Commercial" in the findings below.

1.3. What are the dates of the data used in testing? Calendar Years 2014, 2015, and 2016

1.4. What levels of analysis were tested? (*testing must be provided for <u>all</u> the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☒ health plan	☒ health plan
☐ other:	☐ other:

1.5. How many and which <u>measured entities</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Medicare Advantage

We used MA data from all 50 states and Washington D.C. to create a comprehensive, representative universe of the MA population. The team used all available MA data for these three years rather than a sub-set or sample of the MA data.

Because the MA data contain data for all products, they represent all geographic regions in the 50 states. This population is primarily 65 and over the age of 65, however younger individuals may qualify for MA due to disability or Lou Gehrig's disease.

Exhibit 1: Data Characteristics of MA Measured Entities (products)

	2014	2015	2016
Products	722	727	683
Patients	17,483,103	18,777,871	19,640,281
Avg product enrollment	25,091	26,786	29,737
Smallest product enrollment	1	12	7
Largest product enrollment	1,036,293	1,095,959	1,149,747
Region(s)	All CMS regions in the 50 states and D.C.	All CMS regions in the 50 states and D.C.	All CMS regions in the 50 states and D.C.

Qualified Health Plan/Commercial

We also used Qualified Health Plan (QHP) and commercial plan data from three separate sources. The team selected these databases for their availability, data quality, and multiple issuer representation. This ensures that data are representative of more than one issuer and geographic region. Measured entities (health products/issuers) in these data sets are located in the Midwest and on the West Coast.

Exhibit 2: Data Characteristics of Qualified Health Plan Measured Entities (products)

	2014	2015	2016
Products	10	19	10
Members/Beneficiaries	343,499	500,309	149,947
Avg benes per product	*	*	*
Smallest product enrollment	12	427	5417
Largest product enrollment	325,950	459,782	76,856
Region(s)	West Coast	West Coast; Midwest	West Coast; Midwest
Product Type	QHP	QHP	QHP

*Average enrollment not reported here, as QHP/commercial enrollment was heavily skewed to one larger product.

Together, this testing population was chosen to represent the two federal payer programs in which CMS intends to consider adoption of this measure (the Medicare Advantage Star Ratings and the Qualified Health Plan Health Insurance Exchange).

1.6. How many and which <u>patients</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Medicare Advantage

The MA data represents a wide range of age groups, as well as an even distribution of gender. A greater proportion of the members in the Medicare Advantage program are 65 or over the age of 65 due to eligibility requirements. However, a small fraction of patients were under 65 and were eligible due to disability. In addition, Low-Income Subsidy (LIS) status was available for patients also enrolled in a Part D prescription drug plan, allowing for some analysis of socio-economic status. Of the over 17 million patients in the MA population every year, approximately 5% have an ED visit or inpatient stay for one of these chronic conditions.

Exhibit 3: Medicare Advantage Patient Demographics

	2014		2015		2016	
Total Members	17,483,103		18,777,871		19,640,281	
Age*	**N**	**%**	**N**	**%**	**N**	**%**
Under 18	20	0.0%	20	0.0%	19	0.0%
18-25	21,084	0.1%	22,929	0.1%	21,555	0.1%
26-34	120,712	0.7%	133,117	0.7%	130,519	0.7%
35-44	268,043	1.5%	288,922	1.5%	287,144	1.5%
45-54	699,667	4.0%	736,971	3.9%	741,363	3.8%
55-64	2,061,081	11.8%	2,219,188	11.8%	2,361,751	12.0%
65-74	8,232,646	47.1%	8,891,217	47.3%	9,338,526	47.5%
75-84	4,427,125	25.3%	4,707,053	25.1%	4,900,141	24.9%
85+	1,652,725	9.5%	1,778,457	9.5%	1,859,291	9.5%
Sex						
Female	9,879,391	56.5%	10,603,838	56.5%	11,074,039	56.4%
Male	7,603,728	43.5%	8,174,032	43.5%	8,566,242	43.6%
Low Income Subsidy (LIS) status**						
No	16,865,030	96.5%	18,182,011	96.8%	19,006,575	96.8%
Yes	796,121	4.6%	767,617	4.1%	813,893	4.1%
Acute Events for the Condition						
Total	**810,183**	**4.6%**	**935,526**	**5.0%**	**1,109,555**	**5.6%**
Asthma	64,949	0.4%	73,763	0.4%	61,129	0.3%
CAD	159,745	0.9%	165,548	0.9%	217,659	1.1%
COPD	154,359	0.9%	176,038	0.9%	203,437	1.0%
Diabetes	175,746	1.0%	212,566	1.1%	251,409	1.3%
Heart Failure	102,141	0.6%	122,295	0.7%	136,054	0.7%
Hypertension	153,243	0.9%	185,316	1.0%	239,867	1.2%

*Age as of January 1 of the calendar year. One can be in the Medicare Advantage population if under 65 of age because of disability or Lou Gehrig's disease.

**LIS status = Y if the member had LIS status in at least one month of the year

Qualified Health Plan/Commercial

The QHP data also represents a wide range of age groups and an even distribution of gender. No socioeconomic status variables were collected in any of the QHP administrative claims data sets. Of the patients in the QHP population, just under 1% have an ED visit or acute event for these conditions. However, ED visits and inpatient visits are expensive, and such events for these conditions can greatly affect patients'

quality of life. [3,4,5]

Exhibit 4: Qualified Health Plan Patient Demographics

	2014		2015		2016	
Total Beneficiaries	343,499		500,309		149,947	
Age*	**N**	**%**	**N**	**%**	**N**	**%**
Under 18	17,317	5%	31,251	6%	18,418	12%
18-25	51,598	15%	65,289	13%	12,046	8%
26-34	47,243	14%	70,862	14%	24,799	17%
35-44	56,684	17%	77,062	15%	23,026	15%
45-54	91,742	27%	121,839	24%	25,281	17%
55-64	76,711	22%	129,137	26%	44,087	29%
65-74	1,748	1%	4,131	1%	2,062	1%
75-84	17,317	5%	31,251	6%	18,418	12%
85+	51,598	15%	65,289	13%	12,046	8%
Sex						
Female	178,565	52%	266,518	53%	83,736	56%
Male	164,934	48%	233,724	47%	66,121	44%
Acute Events for the Condition						
Total	**576**	**0.2%**	**2,091**	**0.4%**	**849**	**0.6%**
Asthma	89	0.0%	398	0.1%	111	0.1%
CAD	144	0.0%	412	0.1%	175	0.1%
COPD	42	0.0%	193	0.0%	102	0.1%
Diabetes	56	0.0%	341	0.1%	141	0.1%
Heart Failure	107	0.0%	297	0.1%	151	0.1%
Hypertension	138	0.0%	450	0.1%	169	0.1%

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

Where possible, the team used both QHP and MA data for all testing to ensure reliability and validity in both the under 65 and 65+ populations.

Exclusion testing: Both QHP and MA data were used to test frequency of exclusion criteria.

[3] Pocket Guide to COPD Diagnosis, Management, and Prevention: A Guide for Health Care Professionals: 2017 Report. Global Initiative for Chronic Obstructive Lung Disease 2017. http://goldcopd.org/wp-content/uploads/2016/12/wms-GOLD-2017-Pocket-Guide.pdf

[4] •Effect of Early Follow-Up After Hospital Discharge on Outcomes in Patients With Heart Failure or Chronic Obstructive Pulmonary Disease: A Systematic Review. Song J, Walter M. 2017. Health Quality Ontario. Ontario Health Technology Assessment Series. 2017;17(8):1-37. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5466361/

[5] Follow-up after acute asthma episodes: What improves future outcomes? August 2009. Schatz, Michael et al. Journal of Allergy and Clinical Immunology, Volume 124 , Issue 2 , S35 http://www.jacionline.org/article/S0091-6749(09)00798-2/fulltext

See evidence and overall submission forms for additional citations on frequency and cost of events and link between follow-up and improved outcomes.

Reliability testing: Both MA and QHP data were used to calculate signal-to-noise reliability testing. All years were used for MA data. For QHP data, the only year with a sufficient number of products to calculate reliability was 2015.

Validity testing: The team engaged a technical expert panel to assess face validity of the measure in both populations. Only MA data were used to conduct convergent validity testing. This is because Star Ratings are calculated only for MA products, and not for QHPs.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

MA data: Low income subsidy status (LIS). The team found that there was no meaningful difference in measure scores in the LIS population vs. the non-LIS population.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

For a measure to be reliable, the team must demonstrate that "the measure results are repeatable and the measurement error is acceptable, producing the same results a high proportion of the time when assessed in the same population in the same time period."[6] To this end, the team conducted signal-to-noise ratio analysis.

Signal-to-noise ratio: Signal-to-noise measures how much the differences between the product's measure scores are attributable to actual differences in the quality of those products (signal) and how much is attributable to measurement error (noise).[7,8] Reliability scores vary from 0.0 to 1.0, with a score of 0 indicating that all variation is attributable to measurement error (noise or variation across patients within providers) whereas a reliability of 1.0 implies that all variation is caused by a real difference in performance across accountable entities. Values above 0.7 are considered to be sufficient to see differences between some physicians or clinicians and the mean.

Reliability is calculated using the Beta-Binomial Model as[9]:

$$reliability_{contract\ i} = \frac{\sigma^2_{between-contract}}{\sigma^2_{between-contract} + \sigma^2_{within-contract\ i}}$$

Where $\sigma^2_{within-contract}$ is the variance of the individual product being measured, and $\sigma^2_{between-contract}$ is the variance in measure scores across all products.

---

[6] Centers for Medicare & Medicaid Services. (May 2017). Measures Management System Blueprint (the Blueprint) v. 13.0. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/MMS-Blueprint.html

[7] http://www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR653.pdf

[8] https://www.qualityforum.org/Publications/2013/10/Review_and_Update_of_Guidance_for_Evaluating_Evidence_and_Measure_Testing_-_Technical_Report.aspx

[9] http://www.rand.org/content/dam/rand/pubs/technical_reports/2009/RAND_TR653.pdf

Because the measure is a binomially distributed variable (sum of 0/1 responses divided by N), to estimate variance, the team used the method developed by RAND, fitting a beta binomial model to the data to evaluate reliability on a performance measure score level and looked at performance measure score differences by diverse products. Using these estimates, the team calculated reliability statistics for each measured product, regardless of product size or denominator size of the product.

The team reports the mean and median reliability scores. A measure is typically considered reliable if the average signal-to-noise reliability is greater than or equal to 0.7. Additionally, we report the distribution (minimum, 10th percentile, 25th percentile, 75th percentile, 90th percentile, and maximum) of reliability scores over all products.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

Medicare Advantage

Exhibit 5: Follow-Up After Acute Event, Signal-to-Noise Reliability

Year	N	Mean	Median	Min	Max	Standard Deviation	IQR	P10	P25	P50	P75	P90
2014	657	0.891	0.972	0.142	1.000	0.178	0.118	0.631	0.876	0.972	0.994	0.999
2015	690	0.854	0.960	0.105	1.000	0.203	0.209	0.521	0.782	0.960	0.990	0.997
2016	647	0.839	0.954	0.090	1.000	0.224	0.220	0.484	0.769	0.954	0.989	0.997

P=Percentile (10th 25th, 50th, 75th, and 90th); IQR=Interquartile range

Qualified Health Plan/Commercial

Exhibit 6: Follow-Up After Acute Event, Signal-to-Noise Reliability (2015)

Year	N	Mean	Median	Min	Max	Standard Deviation	IQR	P10	P25	P50	P75	P90
2015	6	0.658	0.678	0.298	0.931	0.244	0.360	0.298	0.503	0.678	0.863	0.931

* Due to the small sample size, reliability statistics for the QHP/Commercial population are calculated with only products having denominators >=30. These are the only products that scores would be displayed for, for the intended measure use in the QHP Health Insurance Exchange.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Average MA measure reliability exceeds the 0.7 threshold in all years. In addition, more than 75% of products have reliability scores above the 0.7 threshold. These levels indicate that measurement error (noise) is acceptably low and therefore higher performing products can be distinguished from lower performing products.

Finally, QHP/Commercial measure reliability is very near to the 0.7 threshold at 0.658. Reliability was calculated for only 6 of the 10 products, because only 6 products has denominators greater than or equal to 30.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☒ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Performance measure score: Systematic assessment of face validity

Face validity of the measure score as an indicator of quality was systematically assessed by engaging technical expert panels at multiple stages throughout measure development. These panels were comprised of multiple key experts including clinicians, payors, patient representatives, and medical researchers. The measure development team held initial workgroups with these expert panels to guide development and ultimately conducted a final workgroup during which the measure development team provided the expert panel with the detailed measure specifications (as specified in this application) and requested that they state (yes or no) whether they agreement with the following statement:

"The performance scores resulting from the measure Timely Follow-Up After Acute Exacerbations of Chronic Conditions, as specified, can be used to distinguish good from poor plan-level quality." (92% yes, 8% no)

Performance measure score: Empirical validity testing

To provide additional evidence beyond the technical expert panel assessment of face validity, the team also assessed convergent validity.

Convergent validity refers to the degree to which multiple measures of a single underlying concept are interrelated. For instance, we would expect MA products with high rates of follow-up after emergency department (ED) visits to have higher rates of follow-up after hospitalizations as well.

The team used publicly available product-level scores for MA Star Ratings measures that were hypothesized to be related to or measure the same underlying concept as the testing measures. The team used the following Star Ratings measures:

Exhibit 7: Star Ratings Measures Used in Convergent Validity Testing

Measure Name
Access to Primary Care Doctor Visits
Annual Flu Vaccine
Asthma Medication Ratio
Care Coordination
Continuous Beta Blocker Treatment
Diabetes Care–Eye Exam
Follow-up Visit After Hospital Stay for Mental Illness
Getting Appointments and Care Quickly
Getting Needed Care
Medication Management for People With Asthma
Medication Reconciliation Post-Discharge
Pharmacotherapy Management of COPD Exacerbation – Systemic Corticosteroid
Statin Therapy for Patients With Cardiovascular Disease
Testing to Confirm COPD

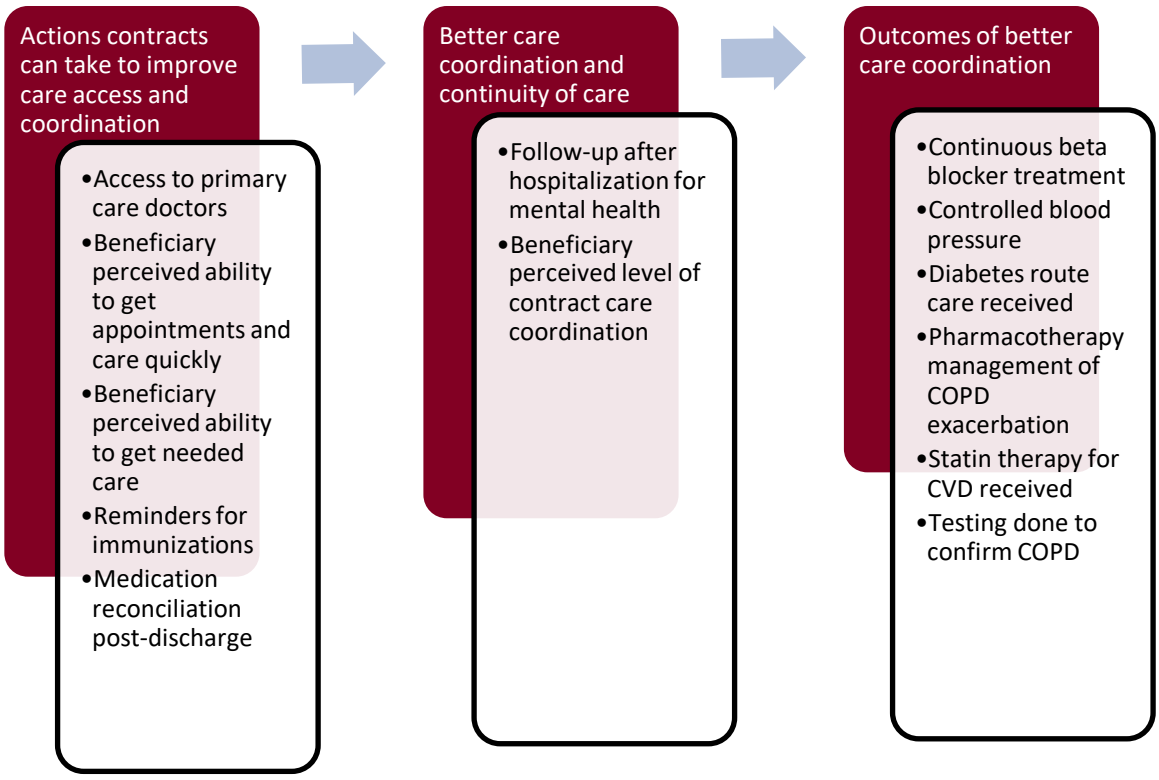
*The Star Rating Report Years used data from the Data Years to report for each measure.

The team matched the Star Ratings measure scores and testing measure scores for products based on data years. For example, each product's "Follow-up after acute event" scores for events occurring in 2014 were matched with that product's "Access to Primary Care Doctor Visits" scores for data year 2014, and the same process was used for each of the 2015 and 2016 program years. Then the team calculated Pearson's correlation coefficient and p-values to determine if the correlation between products' measure scores was statistically significant.

To show convergent validity of the measure, the team sought to show that:

1. there are actions a product can take that increase the likelihood of follow-up after acute events.
2. the follow-up measure is a measure of the underlying construct of "care coordination."
3. follow-up after acute events decreases the likelihood of adverse outcomes and increases continuity of care.

Exhibit 8: Follow-Up After Acute Event, Construct Model



The team hypothesized that actions such as making primary care easily and quickly accessible, contacting beneficiaries to remind them of necessary care, and incentivizing care coordination actions after a discharge can increase the likelihood of follow-up after discharge. If this is the case, then products that score well on the Star Ratings measures that quantify these activities will also score well on the follow-up measure.

The team additionally hypothesized that, if the follow-up measure is a measure of the construct of care coordination, then products that score well on the follow-up measure will also score well on the Star Ratings measures that target the same underlying construct of care coordination.

Finally, the team hypothesized that if follow-up leads to better outcomes, the products that scored better on the follow-up measure would also score well on several Star Ratings outcomes measures.

The team did not expect as high correlations as seen under reliability testing of this measure. Under reliability, one might expect a hypothetical "product 1", which scored well in year 1, to also score well in year 2 because a product's measure score for the same measure in two time periods are nearly identical measures. In validity

testing, we examine the scores of "product 1" in two related, but different measures. We expect to see statistically significant correlations, though perhaps not at the 0.7 same threshold.

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Exhibit 9: Follow-Up After Acute Event, Process Measures Affecting Follow-up

Measure	Pearson's Correlation Coefficient**		
	2014	2015	2016
Access to Primary Care Doctor Visits	0.1794	0.3028	0.2556
Getting Appointments and Care Quickly	0.3401	0.2620	0.3263
Getting Needed Care	0.1959	0.1897	0.2129
Medication Reconciliation Post-Discharge	*	0.1527	0.3055

*Measure not included in Star Ratings for this year

**All correlations in this table are statistically significant at p<0.05

Exhibit 10: Follow-Up After Acute Event, Measures of Care Coordination

Measure	Pearson's Correlation Coefficient**		
	2014	2015	2016
Follow-up visit after Hospital Stay for Mental Illness	0.2706	0.2570	0.2289
Care Coordination	0.2303	0.2739	0.2216

**All correlations in this table are statistically significant at p<0.05

Exhibit 11: Follow-Up After Acute Event, Outcomes Measures Follow-up Effects

Measure	Pearson's Correlation Coefficient**		
	2014	2015	2016
Annual Flu Vaccine	0.3984	0.2903	0.2595
Asthma Medication Ratio	*	0.2180	0.3062
Continuous Beta Blocker Treatment	0.2694	0.2204	0.2388
Diabetes Care – Eye Exam	0.2163	0.1891	0.1262
Pharmacotherapy Management of COPD Exacerbation – Systemic Corticosteroid	0.2416	0.2550	0.2397
Statin Therapy for Patients With Cardiovascular Disease	*	0.1012	0.1209
Testing to Confirm Chronic Obstructive Pulmonary Disease	0.1856	0.1531	0.2288

*Measure not included in Star Ratings for this year

**All correlations in this table are statistically significant at p<0.05

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Empirical Validity Testing:

Exhibit 9, Exhibit 10, and Exhibit 11 show the statistically significant (p<0.05) correlations between the follow-up measure scores and Star Ratings measure scores.

Products that scored well on the follow-up measure also scored well on process measures that may lead to more access to care for beneficiaries and therefore increased follow-up rates (Exhibit 9).

Products that scored well on the follow-up measure also scored well on other measures of care coordination, indicating that the follow-up measure also targets the underlying construct of care coordination (Exhibit 10).

In addition, products that scored well on the follow-up measure also scored better on several outcomes measures during the same time period, indicating that follow-up may lead to improved outcomes for beneficiaries (Exhibit 11). Notably, products that performed well on the follow-up measure also had higher rates of flu vaccinations, higher rates of continuous beta blocker treatment when necessary, higher rates of statin therapy for patients with CVD, higher rates of pharmacotherapy management of COPD after an exacerbation, and higher rates of testing to confirm COPD.

These correlations are not particularly strong, ranging from 0.1 to 0.4, however the team does not expect correlations here to reach the 0.7 threshold to be considered meaningful. This is because we are looking at correlation between a product's score in two related but different measures. Follow-up in 2014 should be a very strong predictor of follow-up in 2015 (test-retest reliability); survey responses of beneficiary perceived care coordination are a predictor of follow-up in 2015, though not as strong as the prior years' follow-up rate.

Systematic Assessment of Face Validity

The measure was evaluated by a group of experts, and 92% agreed that the measure as specified in this application distinguishes good from poor plan-level quality and recommended the measure as specified to be included in the Health Insurance Exchange program.

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section* 2b3

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

Exclusions were developed to prevent health insurance products/issuers from being pentalized when, logistically and logically, it would be difficult or impossible for a patient to seek follow-up care or to identify that outpatient follow-up using claims data. These exclusions are generally consistent with exclusions in similar NQF-endorsed measures of follow-up.

Patient is not discharged to community at end of event: Most events are excluded because the patient is not discharged to community at the apparent end of the event. Discharge status is determined as of the last claim of the event. If the status is not discharged to community (e.g., still patient, patient left against medical advice), the event is excluded. These events must be excluded because it may not be appropriate to expect patients who are not discharged to community to receive follow-up. One important consideration is the implicit exclusion of events where a patient left the emergency department (ED) or inpatient facility against medical advice. These events are excluded because the patients did not complete their medical care, so we are uncertain as to whether the patients received care that would warrant follow-ups. Furthermore, it is not likely that a product would be able to take action that would meaningfully affect patients who left against medical advice.

Calendar year ends within follow-up window: These events are excluded because claims for January of the following year may not be complete in the IDR at the time of measurement. As such, counts of follow-up in January of the following year may be inaccurate until data are available for the following calendar year.

Event followed by nonacute visit within follow-up window: As with events where a patient is not discharged to community, an event where a patient enters a nonacute care facility within the follow-up window is excluded because the patient would be unable to receive outpatient follow-up care while in the non-acute facility.

Event followed by hospice within follow-up window: As with events where a patient is not discharged to community, an event where a patient enters hospice within the follow-up window is excluded because the patient would be unable to receive outpatient follow-up care while in hospice.

Event is a readmission: Events are excluded that occur within the follow-up window of another event. We include the first event in the series of events because the patient should be receiving follow-up after the first

event. In particular, if a patient has two events in a short time period, this is likely an issue with the treatment received during the first event, and one way to reduce the probability of a subsequent event is to receive follow-up before the occurrence of the subsequent event.

Non-continuous enrollment: Events are excluded when the patient does not have continuous enrollment after the event because products have limited ability to influence follow-up for beneficiaries who switch to another insurer during the follow-up time window.

The above exclusions were analyzed for frequency and variability across products. In addition, chi squared analyses were performed to ensure that different populations (by age and gender) did not have differing rates of exclusions.

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Exhibit 12: Medicare Advantage Exclusion Testing Results

Exclusion	Occurrence 2014	Occurrence 2015	Occurrence 2016
Acute events before exclusions	**1,585,942**	**1,832,079**	**2,033,319**
Patient is not discharged to community at end of event	36.9%	37.7%	28.9%
Calendar year ends within follow-up window	7.0%	5.8%	6.7%
Event followed by nonacute visit within follow-up window	9.3%	9.6%	12.7%
Event followed by hospice within follow-up window	0.0%	0.0%	0.1%
Event is a readmission within the follow-up window	4.9%	5.2%	5.1%
Non-continuous enrollment	5.4%	4.9%	4.9%
Denominator after exclusions	**810,183**	**935,526**	**1,109,555**

Note: percentages of the population excluded are not mutually exclusive: for example a patient may be excluded because they both had a non-acute visit *and* non-continuous enrollment

The percentage of beneficiaries excluded for each exclusion did not differ significantly between groups, ether by age or by gender (< 0.05).

Exhibit 13: Qualified Health Plan/Commercial Exclusion Testing Results (2015)

Exclusion	Occurrence 2015
Acute events before exclusions	**4,496**
Patient is not discharged to community at end of event	15%
Calendar year ends within follow-up window	6%
Event followed by nonacute visit within follow-up window	30%
Event followed by hospice within follow-up window	0%
Event is a readmission within the follow-up window	4%
Non-continuous enrollment	39%
Denominator after exclusions	**1,499**

Note: percentages of the population excluded are not mutually exclusive: for example a patient may be excluded because they both had a non-acute visit *and* non-continuous enrollment

The percentage of beneficiaries excluded for each exclusion did not differ significantly between groups, ether by age or by gender (< 0.05).

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data*

collection and analysis. Note: ***If patient preference is an exclusion****, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

The above exclusions are necessary to prevent events from being measured when, logistically and logically, it would be difficult or impossible for a patient to seek outpatient care or to identify outpatient follow up in claims data.

While a large percentage of QHP patients were excluded due to non-acute inpatient care following their acute event, a patient's status as an inpatient prevents them from receiving outpatient follow-up and outpatient follow-up may be inappropriate if non-acute inpatient care is still required. A large percentage of QHP patients were also exclude due to the fact that they switch plans during the follow-up window. However, the measured entities (health products/issuers) cannot coordinate follow-up for patients who are no longer ensured through them and cannot reasonably be measured on these patients. While many acute events are excluded due to these two exclusions, it would be inappropriate to include these events in the denominator.

Notably, these exclusions do not require additional data collection beyond what is already collected in administrative claims data.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4**.**

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

N/A, as this measure is a process measure, not an outcomes measure.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

N/A

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

N/A

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

N/A

2b3.4a. What were the statistical results of the analyses used to select risk factors?

N/A

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

N/A

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

N/A

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

N/A

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

N/A

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

N/A

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

N/A

2b3.9. Results of Risk Stratification Analysis:

N/A

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

N/A

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

N/A

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

Besides the signal-to-noise ratio, which shows that there are identifiable differences in performance scores between products (signal), we also calculated the distribution of measure scores across products, excluding products with fewer than 30 events in the denominator. We calculated confidence intervals for each product to determine whether it contained the mean (is not different from average), was above the mean (better than average), or below the mean (worse than average).

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g.,

number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined)

Medicare Advantage

Exhibit 14: Measure Scores Distribution

Year	N Products	Mean	StdDev	IQR	Min	P10	1 QTL	Median	3 QTL	P90	Max
2014	506	59%	13%	13%	0%	45%	54%	62%	67%	72%	83%
2015	523	58%	12%	14%	6%	44%	52%	59%	66%	70%	87%
2016	498	60%	11%	11%	3%	48%	56%	63%	67%	71%	90%

Exhibit 15: Significant Differences from Mean

Year	N Products	UNDER	%_UNDER	ABOVE	%_ABOVE
2014	506	125	25%	213	42%
2015	523	144	28%	206	39%
2016	498	119	24%	196	39%

Qualified Health Plan/Commercial (2015)

Exhibit 16: Measure Scores Distribution

Year	N Products	Mean	StdDev	IQR	Min	P10	1 QTL	Median	3 QTL	P90	Max
2015	6	58%	7%	10%	48%	48%	52%	58%	63%	66%	66%

Exhibit 17: Significant Differences from Mean

Year	N Products	UNDER	%_UNDER	ABOVE	%_ABOVE
2015	6	1	17%	1	17%

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

The MA measure scores had a wide distribution. Measure scores ranged from 3%-90% in 2016, though ninety percent of products scored between 48%-90%, meaning that there were only a small handful of outliers on the low end of the range. The median score for 2016 was 63%.

In addition, 24% of MA products scored statistically significantly under the average measure score and 39% of products scored statistically significantly above the average measure score.

The QHP/Commercial data had a narrower rance of measure scores, however only 6 products could be included for the analysis due to small denominator sizes.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the*

numerator). ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

All data sources use only one set of specifications.

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

N/A

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

N/A

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

While bias cannot truly be completely eliminated from measurement, we have developed a measure that we believe keeps these to a minimum.

Information Bias: There should be little concern for information bias since the care process is objective and there is a low likelihood of misreporting the given care process.
Missing Data Bias: There is little concern for missing data bias because CMS mandates that all Medicare Advantage claims be submitted and processed via the Integrated Data Repository. Similarly, there is evidence that claims submitted to QHP plans aren't substantively different from claims submitted to other health plans, in which case the measure is tested in a broader claims set. It is important that the data set include some claims from QHPs.
Selection Bias: There should be little concern that the target population is not representative of the population. Based on our understanding of the literature, the need for follow-up spans across a wide range of individuals and there is a gap in care.
Confounding Bias: No empirical testing was performed since this metric is neither an outcome or resource use measure.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Missing data in the Medicare Advantage population is rare. While there is not a large volume of literature on this, there is a report from the Office of Inspector General on the data. In short, it does note that there were significant missing values for NPIs and provider last names, but asserts that the NPI issue was subsequently fixed.[10] The nature of the Health Insurance Exchange is less well-known. The instability of the Marketplace enrollment contributes to data inconsistencies and at times missingness of data. As this is a relatively new

[10] https://oig.hhs.gov/oei/reports/oei-03-15-00060.pdf. Last accessed on July 17, 2018.

population and new program that has not yet reached national implementation, weare continuing to find ways to address missing data. CMS is also continuing to monitor the prevalence of missing data in this population and continues to seek input from expert stakeholders to enhance the scientific vigor of measures. We anticipate that in future years more data will become available.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

N/A

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in electronic claims

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

N/A – All the data elements needed to compute the performance measure are from electronic sources.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

N/A – This measure is not instrument-based and is not being submitted for maintenance of endorsement. As this measure has not yet been implemented, difficulties with data collection have not been experienced.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

N/A – We do not anticipate any fees, licensing, or other requirements to use the measure as specified.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Payment Program	

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

N/A- This measure is not currently in use.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

This measure has been developed under contract by CMS. The goal of the contract is to develop and test the de novo measure and then integrate it into the Medicare Part C and D Star Ratings program and the Quality Rating System program. As part of the contract, CMS requested that NQF-endorsement be sought for this measure before the measure was adopted into either of the aforementioned programs. Therefore, this measure has not yet been implemented in practice or in any other programs.

Program Links:

Quality Rating System (https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/ACA-MQI/Quality-Rating-System/About-the-QRS.html)

Medicare Part C and D Star Rating System (https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/PerformanceData.html)

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

As noted in 4a1.2, this measure has been developed under a CMS contract. The goal of the contract is to integrate this issuer-product level quality measure into the Medicare Part C and D Star Ratings program and the Quality Rating System program. CMS has requested that we seek NQF endorsement of this measure first before implementation considerations into current CMS initiatives.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

Although this measure has not yet been implemented, and we therefore do not have any usability data to date, our team solicited and obtained input from multiple stakeholder groups throughout the measure development process. We believe in a transparent measure development process; highly value the feedback received on the importance/relevance, scientific acceptability, feasibility, and usability of the draft measure; and have taken it into consideration to create a more robust measure. Our team worked closely with a technical expert panel (TEP) composed of clinicians, payers, patient representatives, researchers, and other experts in the field (see section 4a.2.2.3) and held a public comment period to solicit feedback on the measure before finalizing the specifications.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

N/A-This measure has not yet been implemented and we do not have any usability data. We do have limited feedback from the public comment period.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

N/A

4a2.2.2. Summarize the feedback obtained from those being measured.

N/A

4a2.2.3. Summarize the feedback obtained from other users

PUBLIC COMMENT

The public comment survey included a link to the full specifications of the measure—as well as a document outlining evidence related to importance to measure and clinical appropriateness, usability, and use, and related and competing measures. The development team also reached out to individuals involved in the development of prior iterations of the measure to solicit their feedback during the comment period. These individuals included physicians, clinical claims experts, care coordination specialists, and academic researchers.

Despite outreach efforts, the public comment results were limited by a very low number of respondents. These respondents focused on a few key themes:

- Evidential strength concerns: Commenters noted that the evidence supporting the time periods for follow-ups is relatively weak and is variable in certain cases.

- Risk Adjustment: Commenters noted that risk adjustment is needed to account for sociodemographics, clinical severity, and other factors.
- Communication Barriers: Commenters noted significant communication barriers between acute facilities and primary care clinicians that make coordinating follow-up difficult.
- Link to Improvement: Commenters requested additional, clearer information on the link between increasing health plan adherence to follow-up periods and better outcomes, decreased expenditures, increased patient safety, and other quality improvements.

IMPAQ's responses to these comments are outlined in 4a2.3.

TEP

The TEP had met previously to provide feedback on the measure at an earlier stage in development, and so were already intimately familiar with the measure. The development team provided the TEP members with additional information on the measure and all the updates that had occurred since the previous TEP meeting—including providing full and detailed specifications, an explanation of the scoring methodology, and answers to any other questions the TEP may have had. Following the discussion and presentation, the TEP was asked to judge whether the tool is a credible measure of accountability and federal payment programs, and its potential to drive process improvement and judged favorably as detailed below. Workgroup members were requested to state (yes or no) their agreement with the following statement:

"The performance scores resulting from the measure Timely Follow-Up After Acute Exacerbations of Chronic Conditions, as specified, can be used to distinguish good from poor plan-level quality." (92% yes, 8% no)

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

We provide responses to each comment theme below:

• Evidential strength concerns: The follow-up periods used in this measure are based on the most recent available clinical practice guidelines and the extant literature. In many cases, this evidence is from observational studies, smaller RCTs, or expert consensus. This is because, generally, the 'Level A' or 'Level B' evidence preferred by NQF does not currently exist. Despite this, the clinical guidelines and literature universally recognize the importance of follow-up, and we believe that, in addition to improving outcomes, this issuer-product-level measure may provide additional empirical evidence on the effectiveness of optimal follow-up.

• Risk Adjustment: The TEP has taken into consideration risk adjustment and will review again before finalizing the measure. However, NQF typically does not consider risk adjustment appropriate in non-outcome measures. Additionally, while some comments centered on providers being penalized due to factors outside their control (e.g., socioeconomic determinants), .IMPAQ emphasizes that this measure is designed to incentivize improvement at the issuer-product level, rather than at the individual provider level. Still, IMPAQ recognizes that if a health plan (insurance product) provides financial incentives to providers based on its quality performance, an unintended consequence of holding plans accountable could be that providers are financially penalized if the plan performs poorly. However, providers, in this case, have every incentive to help patients receive timely follow-up care, which could render the net effects being positive. In the end, the inter-relationship between health plans, care providers, and hospitals makes our measure even more important as it works to encourage an effective collaboration across these parties in pursuit of a better patient health.

• Communication Barriers: IMPAQ agrees that systemic communication barriers impede follow-up and care coordination. The goal of this measure is to incentivize health plans (insurance products) to overcome these barriers in order to coordinate care and provide timely follow-up and improve outcomes.

• Link to Improvement: IMPAQ has conducted additional literature reviews and provided additional empirical evidence demonstrating the link between the reception of timely follow-up care and improved outcomes such as lower readmission rates. This literature, which is summarized in 1b.1, is bolstered by our validity testing,

which finds that higher follow-up rates are convergent with improved performance on a wide range of other quality measures within the testing population. Please see the testing form for more details.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

This measure has not been implemented yet. However, there is strong reason to believe that this measure, by incentivizing health plans (insurance products) to ensure patients receive appropriate follow-up care, will promote development and implementation of innovative quality improvement activities to improve high performance on this measure. Evidence has shown that quality measures are an effective method in driving quality improvement (1).

Ensuring patients receive appropriate follow-ups is a complex process and may involve many different quality improvement (QI) activities and process improvements. Although each institution, health system, or other measured entity has its own unique established care processes, the QI activities can be expected to include the domain of care coordination, with a focus on improving communication between care team members and with patients pre- and post-discharge. Broadly, there is evidence to support a number of QI activities that may improve rates of follow-up, including conducting telephone outreaches to patients to remind them of appointments and further assist in scheduling (2, 3), scheduling follow-up care/visit(s) at discharge (4), including scheduling follow-up care/visit(s) using innovative referral systems to minimize barriers (5), and designating specific providers to design plans for follow-up care (such as within the advanced practice nurse Transition of Care model) (6). To improve their scores, measured entities should examine their own care processes and potential barriers for patients to receive follow-up care to determine if these or other evidence-driven activities could be implemented or improved if currently in place.

As this measure is intended to be used in payment programs, the reported measure scores will influence public reporting. The measure will be used to distinguish high-performing from low-performing health plans (insurance products) and other measured entities. In addition, as this measure is being put forth for the Medicare Star Ratings program, the data from this measure may be made available publicly via Public Use Files each year—enabling researchers to conduct longitudinal quality analysis on the measured entities.

CITATIONS:

1) Chassin M, Loeb J, Schmaltz S, Wachter R. Accountability Measures – Using Measurement to Promote Quality Improvement. N Engl J Med 2010; 363:683-688. doi: 10.1056/NEJMsb1002320.

2) Smith SR, Jaffe DM, Fisher EB, Trinkaus KM, Highstein G, Strunk RC. Improving follow-up for children with asthma after an acute Emergency Department visit. J Pediatr. 2004 Dec;145(6):772-7. doi: 10.1016/j.jpeds.2004.08.029.

3) Turner D. Can telephone follow-up improve post-discharge outcomes? Br J Nurs. 1996 Dec 12-1997 Jan 8;5(22):1361-5. doi: 10.12968/bjon.1996.5.22.1361.

4) Zorc JJ, Scarfone RJ, Li Y, Hong T, Harmelin M, Grunstein L, Andre JB. Scheduled follow-up after a pediatric emergency department visit for asthma: a randomized trial. Pediatrics. 2003 Mar;111(3):495-502. doi: 10.1542/peds.111.3.495.

5) Messina FC, McDaniel MA, Trammel AC, Ervin DR, Kozak MA, Weaver CS. Improving specialty care follow-up after an ED visit using a unique referral system. Am J Emerg Med. 2013 Oct;31(10):1495-500. doi: 10.1016/j.ajem.2013.08.007.

6) Naylor M, Keating S. Transitional Care: Moving patients from one care setting to another. Am J Nurs. 2008 Sep; 108(9 Suppl): 58–63. doi: 10.1097/01.NAJ.0000336420.34946.3a.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

N/A-This measure has not yet been implemented.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

NA-This measure has not been implemented yet.

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0229 : Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following heart failure (HF) hospitalization

1789 : Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)

1891 : Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

1. Follow-Up After Emergency Department Visit for People with High-Risk Multiple Chronic Conditions. (Steward: NCQA)

2. Proportion of persons with a chronic condition that have a potentially avoidable complication during a calendar year. ("PAC") NQF Measure #0709*

3. Post-Discharge Appointment for Heart Failure Patients. NQF Measure #2439*

4. Heart Failure (HF): Detailed discharge instructions. NQF Measure #0136*

*NQF measures that are no endorsed

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

Yes

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

There are several measures that may be considered distally related to our measure, listed below. Our team has evaluated these measures and believes that harmonization has been appropriately addressed and is not an issue. 1. Hospital-Wide All-Cause Unplanned Readmission Measure. NQF Measure #1789. (Currently Endorsed)-Related Denominator Our measure is distally related to NQF Measure #1789. Because NQF Measure #1789's denominator is composed of discharges from inpatient care across all conditions, our measure's denominator overlaps in the domain of discharge from inpatient care for our six chronic conditions. However, the quality construct captured by our measure is fundamentally different than NQF Measure #1789: our measure captures timely follow-up after acute events, while NQF Measure #1789 captures readmissions. Additionally, our measure is captured at the issuer-product level, while NQF Measure #1789 is a hospital/patient-level measure. 2. Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following chronic obstructive pulmonary disease (COPD) hospitalization. NQF Measure #1891. (Currently Endorsed)-Related Denominator Our measure is distally related to NQF Measure #1891, as both measures utilize COPD discharges from inpatient care as part of the denominator specifications. However, the quality construct captured by our measure is fundamentally different than NQF Measure #1891: our measure captures timely follow-up after acute events, while NQF Measure #1891 captures readmissions. Additionally, our measure is captured at the issuer-product level, while NQF Measure #1891 is a hospital/patient-level measure. 3. Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following heart failure (HF) hospitalization. NQF Measure #0229. (Currently Endorsed)-Related Denominator Our measure is distally related to NQF Measure #0229, as both measures utilize heart failure discharges from inpatient care as part of the denominator specifications. However, the quality construct captured by our measure is substantively different than NQF Measure #0229: our measure captures timely follow-up after acute events, while NQF Measure #0229 captures readmissions. Additionally, our measure is captured at the issuer-product level, while NQF Measure #0229 is an inpatient/hospital-level measure. 4. Follow-Up After Emergency Department Visit for People with High-Risk Multiple Chronic Conditions. Non NQF-Endorsed Measure – Related Denominator and Numerator An environmental scan noted that NCQA's non-NQF-endorsed measure, 'Follow-Up After Emergency Department Visit for People with High-Risk Multiple Chronic Conditions,' captures a similar quality construct as our own measure. However, we believe our measure differs in meaningful and beneficial ways. Our follow-up measure improves on the NCQA measure by including ED, observation, and inpatient events (NCQA only includes ED events), which helps our measure capture quality across more care settings. Our measure also tailors the follow-up window, based on clinical practice guidelines, for each condition; the NCQA measure offers a blanket follow-up window of 7 days for all included conditions. The NCQA measure only captures follow-up for patients with at least 2 chronic conditions, while patients with 1 condition or multiple conditions are included in our measure. Finally, the conditions included in the 2 measures differ—as the NCQA measure includes asthma, COPD, Alzheimer's, chronic kidney disease, depression, HF, acute myocardial infarction, atrial fibrillation, and stroke. 5. Proportion of persons with a chronic condition that have a potentially avoidable complication during a calendar year. ("PAC") NQF Measure #0709. (NOT Currently Endorsed)-Related Denominator. Our team worked throughout development to ensure our measure is closely harmonized with NQF Measure #0709 ('PAC'). While NQF Measure #0709 focuses on rates of potentially avoidable complications for 6 chronic conditions, our proposed measure captures timely follow-up after ED or hospital (observation and inpatient) visits for beneficiaries with the same 6 chronic conditions. Our intention was to align our measure with PAC denominator to ease reporting burden, while capturing valuable information about quality of care. At the time of the development of this measure, PAC had been endorsed by NQF since 2011. However, it has come to our attention that PAC has recently been de-endorsed, as the steward was unable to continue upholding their measure maintenance responsibilities. Regardless, PAC

remains in use for accountability and reporting in a variety of programs across the country, and we are confident that aligning our measure will ultimately reduce response burden. 6. Post-Discharge Appointment for Heart Failure Patients. NQF Measure #2439 (NOT Currently Endorsed)-Related Denominator/Numerator Our measure is distally related to NQF Measure #2439, which is no longer endorsed, as both measures utilize heart failure discharges from inpatient care as part of the denominator specifications. However, our measure is captured at the issuer-product level, while NQF Measure #2439 is a hospital/patient-level measure. Furthermore, we believe our measure captures the quality construct more comprehensively by also integrating timely follow-up for patients discharged from the ED, and by measuring whether follow-up actually occurred (NQF Measure #2439 only captures whether or not follow-up was scheduled within 7 days of the appointment, not when [or whether] the appointment actually occurred). Additionally, our measure is captured at the issuer-product level, while NQF Measure #2439 is a hospital/patient-level measure. 7. Heart Failure (HF): Detailed discharge instructions. NQF Measure #0136 (NOT Currently Endorsed)-Related Denominator Our measure is distally related to NQF Measure #0136, which is no longer endorsed, as both measures utilize heart failure discharges from inpatient care as part of the denominator specifications. However, NQF Measure #0136 only measures whether or not follow-up was scheduled as part of discharge materials for patients discharge from the hospital with heart failure; our measure captures whether the follow-up visit actually occurred and captures follow-up for five additional conditions, and also includes patients discharged from the ED. Finally, our measure is captured at the issuer-product level, while NQF Measure #0136 is a facility-level measure.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

A few NQF-endorsed measures may be considered distally related to this measure by NQF harmonization criteria, as indicated above. However, our team has not identified any NQF-endorsed or non-NQF endorsed measures that are competing measures to our measure. Our measure's numerator and denominator (measure focus and target population) are conceptually and substantively different from any other existing NQF-endorsed or other measures identified.

While some of the measures may be focused on some of the same target population (denominator, e.g., patients with diabetes, heart failure, COPD, etc.), our measure's denominator is specific to patients with certain chronic conditions who experience acute exacerbations and is, thus, importantly different because of (1) the risk of costly readmissions and other adverse events this population may face, and (2) the focus on whether the patient received the follow-up appointment/care, not just the simple recommendation or order for follow-up appointment/care.

Additionally, this measure is defined at the issuer-product level and scored as an aggregate of all the 6 chronic conditions, as opposed to the other existing measures that focus on readmission or follow-up rates for separate chronic conditions. This distinguishes our measure from others in a meaningful sense. Taking into account the related measures noted in this application, we believe that the differences in measure focus, target population, and the unit of analysis lend credence to our claim that the follow-up measure makes an important contribution to the healthcare system.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Available at measure-specific web page URL identified in S.1 Attachment:

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): IMPAQ International

Co.2 Point of Contact: Jensen, Chiu, NQF@impaqint.com, 443-259-5194-

Co.3 Measure Developer if different from Measure Steward: IMPAQ International

Co.4 Point of Contact: Jensen, Chiu, NQF@impaqint.com, 443-259-5194-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

• Jon Mark Hirshon, MD, PhD, MPH (chair) - University of Maryland, School of Medicine

o Brings a strong skill set and perspective, including clinical and methodological expertise. Is a practicing emergency medicine physician and academic researcher. Emergency medicine is directly relevant to some of the measures being considered.

• Andy Amster, MSPH - Kaiser Permanente

o Has experience in epidemiology, public health, healthcare analytics, and performance measurement.

•Marybeth Farquhar, PhD, MSN, RN - URAC

o Has strong mix of experience including nursing, patient advocacy, and work on quality indicators. Also brings the perspective of URAC and the accreditation process, and worked previously on the TEP.

• Susan Fitzpatrick, RN, BSN - Cigna Healthcare

o Brings the perspective of the insurance industry, and has direct experience with the QRS. Has worked on the beta test of the current QRS measure set. Also leads quality measurement activities for physician and hospital performance.

• Aparna Higgins - Brandeis University

o Brings the perspective of AHIP, a major QRS stakeholder, having worked there before Brandeis. Is very well qualified in quality measurement activities, having served on MAP committees for patient safety, costs and resources, and the QRS. Perspective as a health economist also helps balance the TEP composition.

• Christine Hunter, MD - US Office of Personnel Management

o Uses systematically chosen measures to link profit to quality, and represents another major purchaser of care.

• Carol Keegan, MD - Patient Representative

o Brings the patient/consumer perspective to the TEP.

• Dana Mukamel, PhD - University of California, Irvine

o Has very strong academic research background that includes evaluation of quality rating systems and measure sets. Is a health economist with a focus on health disparities, performance measurement, and risk-adjustment. Also has prior experience on TEPs, and has worked with both AHRQ and CMS on projects around measurement and public reporting.

• Derek Robinson, MD, MBA, FACEP - Health Care Service Corporation

o Has executive-level experience with QHPs, including accreditation and the QRS. Has previous experience working with both CMS and NQF.

• Arlene Salamendra - Patient Representative

o Brings the patient/consumer perspective to the TEP.

• Ted von Glahn, MSPH - von Glahn Consulting

o Is well versed in performance measurement, not only from the technical side but from the end-user side. Has experience working with quality measures, which indicates good knowledge of how measures are used by consumers to inform health plan (including QHP) choice. Brings a consumer or patient perspective, though not directly representing patients.

• Chinwe Nwosu, MS - America's Health Insurance Plans

o Brings 5 years of CMS measure developer experience, represents a stakeholder, and works with other quality measurement organizations (e.g., NCQA).

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2019

Ad.3 Month and Year of most recent revision: 11, 2019

Ad.4 What is your frequency for review/update of this measure? Annual updates and potentially triannual endorsement maintenance cycles.

Ad.5 When is the next scheduled review/update for this measure?

Ad.6 Copyright statement: 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 these code sets. IMPAQ International, LLC disclaims all liability for use or accuracy of any Current Procedural Terminology (CPT- [R]) or other coding contained in the specifications.

CPT (R) contained in the Measure specifications is copyright 2004-2017 American Medical Association. ICD-10 copyright 2017 World Health Organization. 'Type of Bill' codes are copyright 2006-2015 American Hospital Association.

Ad.7 Disclaimers: The performance measure is not a clinical guideline and does not establish a standard of medical care, and has not been tested for all potential applications.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Due to technical limitations, registered trademarks are indicated by (R) or [R].

Ad.8 Additional Information/Comments: NOTE ON NEXT SCHEDULED REVIEW:

This is a de novo measure; we anticipate an annual review in Spring 2020 and a triannual review in Fall 2022.