Quality Payment
PROGRAM

Emergency Medicine

Measure Justification Form

September 2022

Table of Contents

1.0 Introduction

This Measure Justification Form (MJF) provides results for the testing and evaluation of the Emergency Medicine episode-based cost measure. The form is intended to provide detailed information about the testing conducted on this measure, and accompanies the Measure Methodology and Measure Codes List file, which together, comprise the specifications for this cost measure. [1]

1.1 Project Title

Physician Cost Measure and Patient Relationship Codes

1.2 Date

Information included is current on September 27, 2022.

1.3 Project Overview

The Centers for Medicare & Medicaid Services (CMS) has contracted with Acumen, LLC to develop care episode and patient condition groups for use in cost measures to meet the requirements of the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). The contract name is "Physician Cost Measure and Patient Relationship Codes (PCMP)." The contract number is 75FCMC18D0015, Task Order 75FCMC19F0004.

1.4 Measure Name

Emergency Medicine Episode-Based Cost Measure

1.5 Type of Measure

Cost/Resource Use

1.6 Measure Description

The Emergency Medicine episode-based cost measure evaluates a clinician's or clinician group's risk-adjusted cost to Medicare for patients who present to the emergency department (ED) with abdominal pain or chest pain during the performance period. This setting-based measure includes the costs of Part A and B services that are clinically related to the attributed clinician's role in managing care during an emergency medicine episode of care.

[1]CMS, "Emergency Medicine Measure Methodology" and "Emergency Medicine Measure Codes List" *MACRA Feedback Page*, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback

2.0 Importance

2.1 Evidence to Support the Measure Focus

The Emergency Medicine measure was developed for use in the Merit-based Incentive Payment System (MIPS) to meet the requirements of the Social Security Act section 1848(r), added by MACRA. MIPS aims to reward high-value care by measuring clinician performance through 4 areas:

- quality
- improvement activities
- Promoting Interoperability
- cost

Each category assesses different aspects of care, and the categories are weighted such that they're combined into one composite score. CMS is introducing MIPS Value Pathways (MVPs) as a way to align and connect quality measures, cost measures, and improvement activities across performance categories of MIPS for different specialties or conditions. MVPs aim to provide a holistic assessment of clinician value for a specific type of care to achieve better healthcare outcomes and lower costs for patients.

The use of cost measures is required by statute, and their purpose is to assess resource use. To be effective, they should capture costs related to a clinician's care decisions and account for factors outside of their influence. This measure provides clinicians with information about their costs of care that they can use to understand the costs associated with their decision-making. Clinicians play an important role in variation in healthcare expenditures due to their ability to affect costs.[2] A cost measure offers opportunity for improvement if clinicians can exercise influence on the intensity or frequency of a significant share of costs during the episode, or if clinicians can achieve lower spending and better quality of care through changes in clinical practice.

Emergency medicine care is costly, with Medicare outpatient hospital program spending on ED visits increasing from $2.3 billion in 2011 to $4.1 billion in 2017.[3] Medicare patients are more likely to visit the emergency department (ED) than physician offices for unscheduled care.[4] In 2018, there were an estimated 130 million ED visits in the United States and patients aged 65 and over represented 17.9% of these visits.[5] Each year, there are approximately 9 million hospital admissions resulting from ED visits for people aged 65 years and older, which represents greater than 70% of hospital admissions for this age group.[6] A study from 2018

[2]David Cutler et al., "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending," American Economic Journal: Economic Policy 11, no. 1 (February 1, 2019): 192–221. https://doi.org/10.1257/pol.20150421.

[3] Medicare Payment Advisory Commission. Options for slowing the growth of Medicare fee-for-service spending for emergency department services. (June 2019). https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-jun19_ch11_medpac_reporttocongress_sec-pdf/

[4] Venkatesh, A. K., H. Mei, L. Shuling, G. D'Onofrio, C. Rothenberg, Z. Lin, and H. M. Krumholz. "Cross-Sectional Analysis of Emergency Department and Acute Care Utilization among Medicare Beneficiaries." Acad Emerg Med 27, no. 7 (Jul 2020): 570-79. https://doi.org/10.1111/acem.13971.

[5] Cairns C, Kang K, Santo L. National Hospital Ambulatory Medical Care Survey: 2018 emergency department summary tables. Available from: https://www.cdc.gov/nchs/data/nhamcs/web_tables/2018_ed_web_tables-508.pdf.

[6] Smulowitz, P. B., A. J. O'Malley, L. Zaborski, J. M. McWilliams, and B. E. Landon. "Variation in Emergency Department Admission Rates among Medicare Patients: Does the Physician Matter?". Health Aff (Millwood) 40, no. 2 (Feb 2021): 251-57. https://doi.org/10.1377/hlthaff.2020.00670.

found that for Medicare patients seen in the ED, 41.5% were admitted to the hospital and 10.5% were placed into observation care.[7]

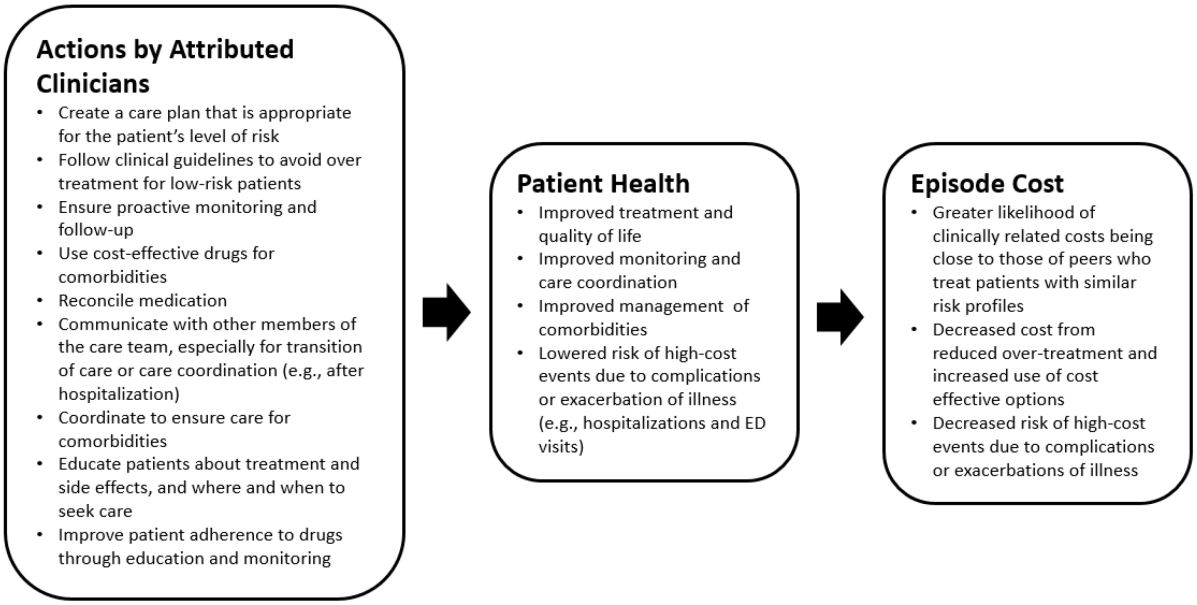
In terms of an estimated cost for an Emergency Medicine episode of care, in 2017, aggregate costs for ED visits among patients aged 65 years and older totaled $20.2 billion, accounting for 26.4% of the $76.3 billion total ED visits among patients of all ages in the United States.[8] The average cost of an ED visit, not adjusted for complexity of care, was $530 in 2017 in the United States. Patients with Medicare as the primary payer had substantially higher costs per ED visit when compared to patients with private insurance or Medicaid ($660 versus $560 and $420, respectively).[9, 10] One study evaluating ED visits for Medicare patients by disposition found that the costs for patients that were admitted to the hospital after an ED visit have increased since 2011 (i.e., $7,639 in 2011 to $7,726 in 2016) as have ED visits resulting in discharge to the community (i.e., $443 in 2011 to $456 in 2016).[11]

The Emergency Medicine episode-based cost measure was recommended for development through feedback gathered during a public comment period. The public recommended this measure because of its high impact in terms of patient population and Medicare spending, and the opportunity for incentivizing cost-effective, high quality clinical care in this area. A measure-specific Clinician Expert Workgroup was then convened with clinicians, healthcare experts, and patient representatives who have appropriate experience to provide extensive, detailed input on this measure throughout its development.

[7] Gabayan, Gelareh Z., Li-Jung Liang, Brian Doyle, David Yu-Chuang Huang, and Catherine A. Sarkisian. "Emergency Department Increased Use of Observation Care for Elderly Medicare Patients." [In eng]. Journal of hospital administration 7, no. 3 (2018): 9-16. https://doi.org/10.5430/jha.v7n3p9.
[8] Moore BJ (IBM Watson Health), Liang L (AHRQ). Costs of Emergency Department Visits in the United States, 2017. HCUP Statistical Brief #268. December 2020. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/reports/statbriefs/sb268-ED-Costs-2017.pdf.
[9] Moore and Liang (2017)
[10] Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Trends in the Utilization of Emergency Department Services, 2009-2018. 2021. https://aspe.hhs.gov/pdf-report/utilization-emergency-department-services
[11] Burke, Laura G., Ryan C. Burke, Stephen K. Epstein, E. John Orav, and Ashish K. Jha. "Trends in Costs of Care for Medicare Beneficiaries Treated in the Emergency Department from 2011 to 2016." JAMA Network Open 3, no. 8 (2020): e208229-e29. https://doi.org/10.1001/jamanetworkopen.2020.8229 %J JAMA Network Open.

2.1.1 Logic Model

Figure 1: Logic Model of Steps between Actions by Attributed Clinicians and Episode Cost



2.2 Performance Gap

2.2.1 Rationale

According to the literature and feedback received through stakeholder input activities, this measure's focus represents an area where there are opportunities for improvement. As discussed in the rest of this section, primary opportunities for improving emergency medicine cost outcomes include (i) enhancing decision making strategies for ED clinicians to reduce variation in inpatient admission rates, (ii) improving transitional care strategies around follow-up and discharge care to prevent unscheduled ED revisits and adverse outcomes, and lastly, (iii) improving efficiency in resource utilization.

One area that clinicians can reduce costs associated with ED visits is enhancing decision strategies to reduce variation in inpatient admission rates. Admission rates vary widely across the country and across hospitals, with one study reporting that between 31-57% of all ED visits are "gray-area" admissions deemed as "intermediate complex," (e.g., asthma exacerbation), where one physician may admit and the other might discharge.[12] Inpatient ED admissions rates vary significantly at the county, hospital, and physician-level, even within the same hospital system[13]. Studies evaluating physician and hospital-level ED admission rates across all diagnoses have demonstrated large variation, with one study in 2013 reporting hospital

[12] Sukayna Z. Alfaraj and Jesse M. Pines, "What We Can Learn from Medicare Data on Early Deaths after Emergency Department Discharge," *Journal of thoracic disease* 9, no. 7 (2017). https://doi.org/10.21037/jtd.2017.06.44

[13] Caines et al., "County-Level Variation in Emergency Department Admission Rates Among US Medicare Beneficiaries" in Annals of Emergency Medicine 68, no.4 (2016): 456-460. https://doi.org/10.1016/j.annemergmed.2016.03.019

admission rates from 9.8% to 25.8% at the 10th and 90th percentiles for all diagnoses.[14] A second study found that reporting admission rates varied from 27% to 41% at the hospital level and 21% to 49% at the physician level, and substantial variation in admission rates among physicians within a hospital is likely due to factors other than patient characteristics.[15] Reducing hospital-level variation in admission rates is estimated to save up to $3.3 billion out of $10.3 billion total expenditures on chest pain ED visits[16]. Additionally, the variation in physician and hospital-level admission rates suggests an opportunity to improve the standards of care, via care pathways or feedback of physician admission metrics, to support ED physicians' decision-making by providing more information about the patient.

Improving transitional care strategies around follow-up and discharge care to prevent unscheduled ED readmissions and adverse outcomes presents another opportunity for containing costs. 22% of all abdominal pain ED discharges generate 30-day revisits with 72-hour returns as high as 23%.[17] Older patients (aged >65 years), who make up over 25% of ED visits with substantially greater risk for severe illness, are 3 times more likely to return to the ED and be hospitalized within 72-hours of an ED visit compared to their younger counterparts under the age of 30.[18] Due to the presence of multiple comorbidities, including cognitive disorders, evaluating older patients tends to be time-consuming with potential for delayed diagnosis and other missteps during an episode of care. In addition to increasing age, previous studies have identified several other patient-level risk factors for readmission or adverse outcomes including dual eligibility, inadequate health literacy, psychiatric comorbidities, ethnicity, and sex.[19,20] While, abdominal pain and chest pain are associated with the highest rates of readmission,[21] for profit status and teaching status of the hospital were also found to be positively associated with short-term recidivism.[22]

Lastly, avoiding the overuse of resources, especially imaging services, during an ED visit can also help reduce health spending. Pines et al. highlighted this inefficiency in the ED setting by demonstrating significant increases in utilization across several types of services between 2000 and 2009, from 36% for radiography to 167% for computed tomography (CT) scans.[23] Another

---

[14] J. M. Pines, R. L. Mutter, and M. S. Zocchi, "Variation in Emergency Department Admission Rates across the United States," *Med Care Res Rev* 70, no. 2 (Apr 2013). https://doi.org/10.1177/1077558712470565

[15] Jameel Abualenain et al., "Emergency Department Physician-Level and Hospital-Level Variation in Admission Rates," *Annals of emergency medicine* 61, no. 6 (2013/06// 2013). https://doi.org/10.1016/j.annemergmed.2013.01.016

[16] Sabbatini, Amber, Brahmajee Nallamothu, and Keith Kocher, "Reducing Variation in Hospital Admissions from The Emergency Department for Low-Mortality Conditions May Produce Savings" in Health Affairs 33, no.9 (2014). https://doi.org/10.1377/hlthaff.2013.1318

[17] Pitts et al., "Where Americans Get Acute Care: Increasingly, It's Not at Their Doctor's Office", Health Affairs 29, no.9 (2010). https://doi.org/10.1377/hlthaff.2009.1026

[18] Sophia Sheikh, "Risk Factors Associated with Emergency Department Recidivism in the Older Adult," *The western journal of emergency medicine* 20, no. 6 (2019). https://doi.org/10.5811/westjem.2019.7.43073.

[19] Kim et al., "Depression is Associated with Recurrent Chest Pain with or without Coronary Artery Disease: A prospective Cohort Study", American Heart Journal 191 (2017): 47-54. https://doi.org/10.1016/j.ahj.2017.06.003

[20] Magidson et al., "Prompt Outpatient Care for Older Adults Discharged from the Emergency Department Reduces Recidivism", Western Journal of Emergency Medicine 12, no.16 (2020): 198-204. https://doi.org/10.5811/westjem.2020.8.47276

[21] Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Trends in the Utilization of Emergency Department Services, 2009-2018. 2021. https://aspe.hhs.gov/pdf-report/utilization-emergency-department-services

[22] Abualenain et al., "Emergency department physician-level and hospital-level variation in admission rates" in Annals of Emergency Medicine 61, no.6 (2013): 638-43. https://doi.org/10.1016/j.annemergmed.2013.01.016

[23] Pines et al., "National Trends in Emergency Department Use, Care Patterns, and Quality of Care of Older Adults in the United States" in *Journal of the American Geriatric Society* 61(1) (2013): 12-17. https://doi.org/10.1111/jgs.12072

study conducted in 2020 found radiography services to be overutilized in undifferentiated abdominal pain in the ED, and imaging performed by non-radiologists was significantly associated with downstream imaging costs.[24] Despite this dramatic rise in imaging services in ED care, rates of admission and detection remained largely unchanged.[25] This overuse of services in the ED setting presents substantial opportunities for improving efficiency, thereby reducing costs of care.

2.2.2 Performance Scores

Table 1 shows the distribution of the measure score for clinician groups identified by a Tax Identification Number (TIN) and individual clinicians identified by a combination of a Tax Identification Number and National Provider Identifier (TIN-NPI).

There are variations in cost performance observed in the measure score for both TIN and TIN-NPI as evidenced by the interquartile ranges and score standard deviations. For both TINs and TIN-NPIs, the 90th percentile score is about 1.4 times larger than the 10th percentile score. The variation in the measure score is in the thousands of dollars, which highlights an opportunity for improvement in the costs of care for an emergency medicine episode by closing the gap between the most and least efficient providers.

Table 1. Distribution of the Measure Score

Metric	TIN	TIN-NPI
Count	4,080	79,787
Mean Score	$5,334	$5,058
Score Standard Deviation	$1,018	$831
Minimum Score	$1,076	$1,135
Maximum Score	$15,063	$84,903
Score Interquartile Range (IQR)	$811	$837
Score Percentile		
10th	$4,399	$4,101
20th	$4,722	$4,511
30th	$4,916	$4,760
40th	$5,051	$4,945
50th	$5,172	$5,101
60th	$5,312	$5,247
70th	$5,499	$5,397
80th	$5,852	$5,578
90th	$6,517	$5,868

[24] Denham et al., "Exploring the evidence-practice gap in the use of plain radiography for acute abdominal pain and intestinal obstruction: a systematic review and meta-analysis" in International Journal of Evidence-Based Healthcare 18(2) (2020): 159-169. doi:10.1097/XEB.0000000000000218

[25] Rao et al., "Trends in Utilization Rates of the Various Imaging Modalities in Emergency Departments: Nationwide Medicare Data From 2000 to 2008" in *Journal of the American College of Radiology* 8(10) (2011): 706-709. https://doi.org/10.1016/j.jacr.2011.04.004

2.2.3 Disparities

Data on how the measure, as specified, addresses disparities is described in Sections 3.1.7 and 3.5.5.

3.0 Scientific Acceptability

3.1 Data Sample Description

Testing is based on the full population of measured entities with a minimum of 20 episodes and patients meeting inclusion and exclusion criteria for the measure, not based on a sample.

3.1.1 Type of Data Used for Testing

Medicare administrative claims, Long-Term Minimum Data Set (MDS), Medicare Enrollment Database (EDB), and Common Medicare Environment (CME).

3.1.2 Specific Dataset Used for Testing

The Emergency Medicine measure uses Medicare Part A and Part B claims data maintained by CMS. Part A and B claims data are used to build episodes of care, calculate episode costs, and construct risk adjustors. Episode costs are payment standardized and risk adjusted to ensure accurate comparison of cost across clinicians. Payment standardization adjusts the allowed amount for a Medicare service to limit observed differences in costs to those that may result from healthcare delivery choices. Data from the EDB are used to determine beneficiary-level exclusions and secondary risk adjustors, specifically Medicare Parts A, B, and C enrollment, primary payer, disability status, end-stage renal disease (ESRD), patient birth dates, and patient death dates. The risk adjustment model also accounts for expected differences in payment for services provided to patients in long-term care based on data from the MDS. Specifically, the MDS is used to create the long-term care indicator variable in risk adjustment.

3.1.3 Dates of the Data Used in Testing

Emergency Medicine episodes ending from January 1, 2019, through December 31, 2019.

3.1.4 Levels of Analysis Tested

The measure was tested at group/practice (TIN) and individual clinician (TIN-NPI) levels.

3.1.5 Entities Included in the Testing and Analysis

Table 2 shows the individual clinician (identified by combination of TIN and NPI) and clinician group/practice (identified by TIN) included in the testing of the emergency medicine measure.

Table 2: Characteristics of Measured Entities with 20 Cases or More

Metric	TIN		TIN-NPI	
	Count	%	Count	%
Count	4,080	100%	79,787	100%
Number of Episodes Attributed	-	-	-	-
10-19 Episodes	0	0%	0	0%
20-39 Episodes	690	16.91%	12,400	15.54%
40-59 Episodes	256	6.27%	8,034	10.07%
60-79 Episodes	165	4.04%	5,987	7.50%
80-99 Episodes	97	2.38%	5,066	6.35%
100-199 Episodes	288	7.06%	17,889	22.42%
200-299 Episodes	147	3.60%	11,873	14.88%
300+ Episodes	2,437	59.73%	18,538	23.23%
Census Region	-	-	-	-
Northeast	769	18.85%	14,491	18.16%

Metric	TIN		TIN-NPI	
	Count	%	Count	%
Midwest	910	22.30%	17,599	22.06%
South	1,547	37.92%	32,162	40.31%
West	755	18.50%	15,238	19.10%
Unknown	99	2.43%	297	0.37%

3.1.6 Patient Cohort Included in the Testing and Analysis

Table 3 shows the patient population for the emergency medicine measure used for testing. It consists of Medicare beneficiaries enrolled in Medicare Parts A and B who have an ED visit that triggers an emergency medicine episode.

Table 3: Beneficiary Demographics

Metric	Value
Count	8,376,135
Mean Age	71.8
Female %	57.1%

3.1.7 Social Risk Factors Included in Analysis

The analysis on social risk factors (SRFs) focused on examining the impact of Dual Medicare and Medicaid enrollment status on the measure. Table 4 outlines variables that may indicate SRFs and their advantages and disadvantages as indicators of individual-level SRFs. On balance, the analysis used dual Medicare and Medicaid enrollment status as the proxy of SRFs due to their broad availability in claims data, accurate measurement at the individual level, and wide acceptance of being a powerful indicator of health outcomes.[26]

Table 4: Social Risk Factors Available for Analysis

Variable	Advantages	Disadvantages	Used in Testing
Dual Medicare and Medicaid enrollment status	• Available for all beneficiaries • Most powerful predictor of poor outcomes[27]	• Variation in Medicaid eligibility across states	Yes

[26] Office of the Assistant Secretary for Planning and Evaluation. "Second report to Congress on social risk and Medicare's value-based purchasing programs." (2020) https://aspe.hhs.gov/pdf-report/second-impact-report-to-congress

[27] See footnote 4.

Variable	Advantages	Disadvantages	Used in Testing
Race/Ethnicity	• Available for most beneficiaries, except for ambiguous categories of "Unknown" or "Other"	• Social risk driven by someone's race is often correlated with and partially captured by dual status[28] • Only 5 categories available, which may lack granularity to fully capture disparities[29,30]	No
ICD-10 Z codes for social determinants of health	• Reflects individual-level factors that influence health status and contact with health services	• Not routinely and consistently coded on claims, only available for 0.1% of all fee-for-service claims in 2019[31]	No
American Community Survey	• Can link beneficiary's ZIP code to socioeconomic (SES) measurement of their neighborhood • Many SES indices can be derived from the survey data (e.g. Agency for Healthcare Research and Quality [AHRQ] index, deprivation index)	• Only a proxy measure, not always accurate at individual-level	No

3.2 Reliability Testing

3.2.1 Level of Reliability Testing

The following levels of reliability were tested: critical data elements used in the measure, clinician group/practice (TIN) and individual clinician (TIN-NPI) levels.

3.2.2 Method of Reliability Testing

Data Element Reliability

The Emergency Medicine measure is constructed using CMS claims data, as described in Section 3.1.2. CMS has implemented several auditing programs to assess overall claims code accuracy, ensure appropriate billing, and recoup any overpayments.

- First, CMS routinely conducts data analyses to identify potential problem areas and detect fraud, and audits important data fields used in this measure, including diagnosis and procedure codes and other elements that are consequential to payment.

[28] See footnote 4.

[29] Nguyen, Kevin H., Kaitlyn P. Lew, and Amal N. Trivedi. "Trends in Collection of Disaggregated Asian American, Native Hawaiian, and Pacific Islander Data: Opportunities in Federal Health Surveys." *American Journal of Public Health* (2022).

[30] Kader, Farah, Lan N. Doan, Matthew Lee, Matthew K. Chin, Simona C. Kwon, and Stella S. Yi. "Disaggregating Race/Ethnicity Data Categories: Criticisms, Dangers, And Opposing Viewpoints", *Health Affairs Forefront* (2022).

[31] Centers for Medicare & Medicaid, Office of Minority Health. "Utilization of Z Codes for Social Determinants of Health among Medicare Fee-for-Service Beneficiaries." (2019) https://www.cms.gov/files/document/z-codes-data-highlight.pdf

Specifically, CMS works with Zone Program Integrity Contractors, and formerly Program Safeguard Contractors, to ensure program integrity; the agency also uses Recovery Audit Contractors to identify and correct for underpayments and overpayments.

- Second, CMS uses the Comprehensive Error Rate Testing (CERT) Program to ensure that Medicare payments are correct in accordance with coverage, coding, and billing rules. CMS continues to perform corrective actions and give providers additional education to ensure accurate billing.
- Lastly, to ensure claims completeness and inclusion of any corrections, the measure was developed and tested using data with a three-month claims run-out from the end of the measurement period.

Clinician-level Reliability

Measure reliability is the degree to which repeated measurements of the same entity agree with each other. For measures of clinician performance, the measured entity is the TIN or TIN-NPI, and reliability is the extent to which repeated measurements of the TIN or TIN-NPI give similar results. To estimate measure reliability, we used a signal-to-noise analysis.

This approach seeks to determine the extent to which variation in the measure is due to true, underlying clinician performance, rather than random variation (i.e., statistical noise) within clinicians due to the sample of cases observed. To achieve this, we calculate reliability scores as:

$$R_j = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_{w_j}^2}$$

Where:

$\sigma_{w_j}^2$ is the within-group variance of the mean measure score of clinician *j*

σ_b^2 is the between-group variance of clinicians within the episode group

That is, reliability is calculated as the ratio of between-group variance to the sum of between-group variance and within-group variance. Reliability closer to a value of one indicates that the between-group variance is relatively large compared to the within-group variance, which suggests that the measure is effectively capturing the systematic differences between the clinician and their peer cohort.

3.2.3 Statistical Results from Reliability Testing

Data Element Reliability

Between 2005 and 2019, CERT estimates that proper payment, which includes payments that met Medicare coverage, coding, and billing rules, ranged from 87.3% to 96.4% of total payments each year.[32] The fiscal year 2020 Medicare fee-for-service program proper payment rate was 93.7%.[33]

[32]Comprehensive Error Rate Testing (CERT) Program. "Appendices Medicare Fee-for-Service 2020 Improper Payments Report". Table A6. https://www.cms.gov/files/document/2020-medicare-fee-service-supplemental-improper-payment-data.pdf-1.
[33]Ibid.

Clinician-level Reliability

Table 5: Reliability at the Accountability Entity Level

Reporting Level	Entities Meeting Case Minimum	Mean Reliability	Median Reliability	% Above 0.4	% Above 0.7
TIN	4,080	0.908	0.987	100%	87.60%
TIN-NPI	79,787	0.783	0.836	100%	72.34%

3.2.4 Interpretation

The results of the data element testing show very high reliability of the critical data elements used by the measure. The measure is highly reliable for both the TIN and TIN-NPI reporting levels, at 0.908 and 0.783 respectively. For reference, CMS generally considers 0.4 as the threshold indicating 'moderate' reliability and 0.7 as high reliability.[34] Additionally, at each testing volume threshold, all TINs and TIN-NPI meet or exceed the moderate reliability threshold of 0.4 and the vast majority are above the high reliability threshold of 0.7.

3.3 Validity Testing

3.3.1 Level of Validity Testing

The validity of the measure was tested using face validity and empirical validity at the clinician group/practice (TIN) and individual clinician (TIN-NPI) levels.

3.3.2 Method of Validity Testing

Face Validity

The Emergency Medicine measure was developed through a structured, iterative process for gathering detailed input from recognized clinician experts on the measure. Experts in this clinical area evaluated specifications to ensure that each aspect of the measure (e.g., assigned services) was intentionally capturing only the costs of care within the reasonable influence of the attributed clinician for a defined patient population (i.e., the ability of the measure score to differentiate good from poor performance).

In developing this measure, Acumen incorporated input from:

(i) an Emergency Medicine Clinician Expert Workgroup;
(ii) a Technical Expert Panel (TEP); and
(iii) the Person and Family Partners.

This process is detailed in the Episode-Based Cost Measures Development Process document posted on the MACRA Feedback Page.[35]

[34] CMS, "Medicare Program; CY 2022 Payment Policies Under the Physician Fee Schedule and Other Changes to Part B Payment Policies; Medicare Shared Savings Program Requirements; Provider Enrollment Regulation Updates; and Provider and Supplier Prepayment and Post-Payment Medical Review Requirements," 86 FR 64996-66031.
[35] CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback.

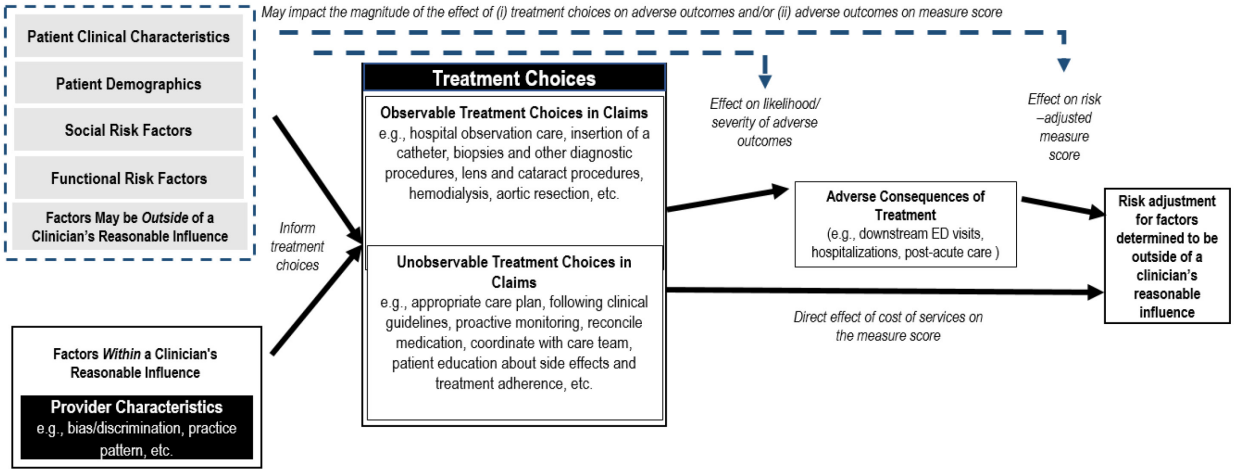
One of the key roles of the measure-specific Clinician Expert Workgroup is to develop service assignment rules for the cost measure. These service assignment rules are intended to ensure clinicians are evaluated on services and costs that are clinically related to the attributed clinician's role in treating and managing the condition, thus limiting cost variation unrelated to clinician care this measure. Therefore, assigned services are services that the Clinical Expert Workgroup believed an attributed clinician can influence their occurrence, frequency, or intensity.

Prior to submitting the measure for the Measure Under Consideration list, members of the Clinician Expert Workgroup were asked to consider the measure as specified and rate the degree to which the actions outlined in the logic model are within the reasonable influence of an attributed clinician, and by extension, can affect patient health outcomes and downstream costs.

Empirical Validity Testing

We evaluated the empirical validity of the Emergency Medicine measure by estimating the effect of relevant treatment choices on the measure score using multiple regression, based on the conceptual model outlined in Figure 2. For more information on the conceptual model, see Section 3.5.3.

Figure 2: Conceptual Model of Treatment Choices on the Measure Score



The cost measure is designed to reflect cost directly related to treatment choices, as well as cost of adverse outcomes as a result of care. Therefore, treatment choices, either observable in claims or otherwise, by an attributed clinician can directly impact the measure score or indirectly when they're mediated through the cost of adverse outcomes. The cost of adverse outcome, in turn, contributes to the total cost that are captured by the measure score.

To demonstrate that the measure score is reflective of both the direct and indirect effects of treatment choices, this analysis first estimates the association between treatment choices and the measure score while controlling for the cost of adverse outcomes. Then, the association between treatment choices and cost of adverse outcomes is estimated to demonstrate the indirect effect.

Generally, adverse outcomes are non-trigger inpatient hospitalizations, non-trigger emergency room visits, and post-acute care. The remaining cost categories are generally considered treatment. For each of these categories, the regression models use the mean cost across

episodes that were attributed to an individual clinician. The measure score is represented by a clinician's mean observed cost over expected cost ratio across their attributed episodes.

3.3.3 Statistical Results from Validity Testing

Face Validity

Figures 3 to 7 show the responses of the Clinical Expert Workgroup Members, when asked to consider the measure as specified and rate the degree to which the actions by an attributed clinician outlined in the logic model is within their reasonable influence and can affect patient health outcomes and downstream costs.

Figure 3: Reponses of Clinical Expert Workgroup Members when Asked to Rate the Degree of Influence of Attributed Clinicians over Actions Outlined in the Logic Model

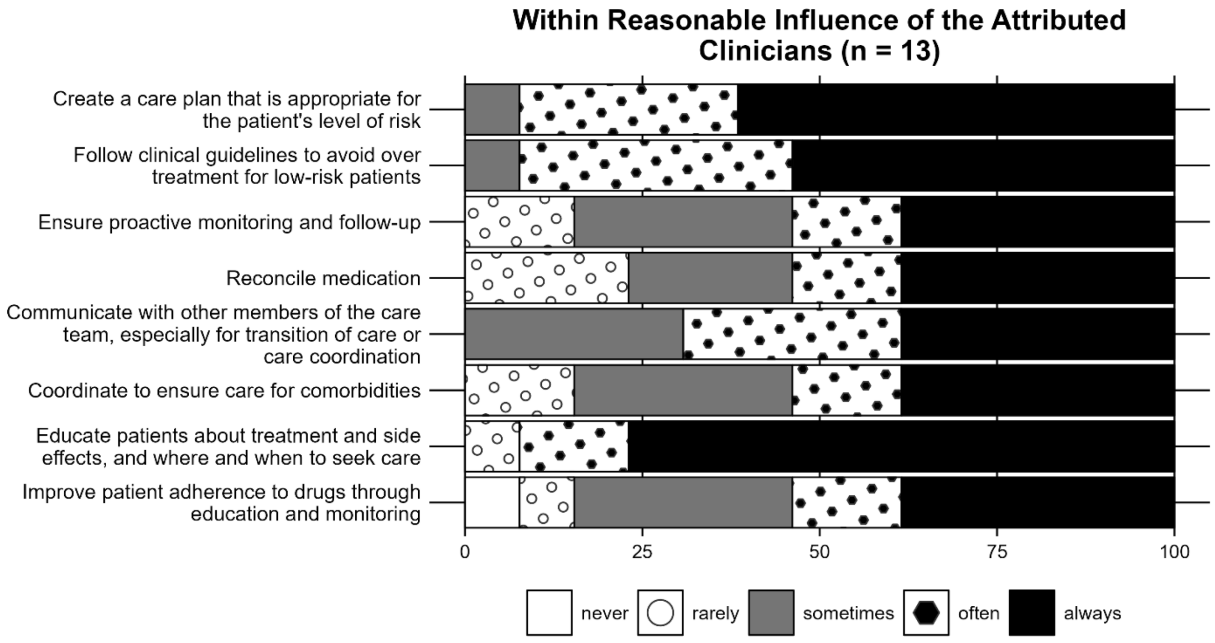


Figure 4: Reponses of Clinical Expert Workgroup Members when Asked to Rate the Likelihood of Impact on Risk of High-Cost Events for Actions Outlined in the Logic Model

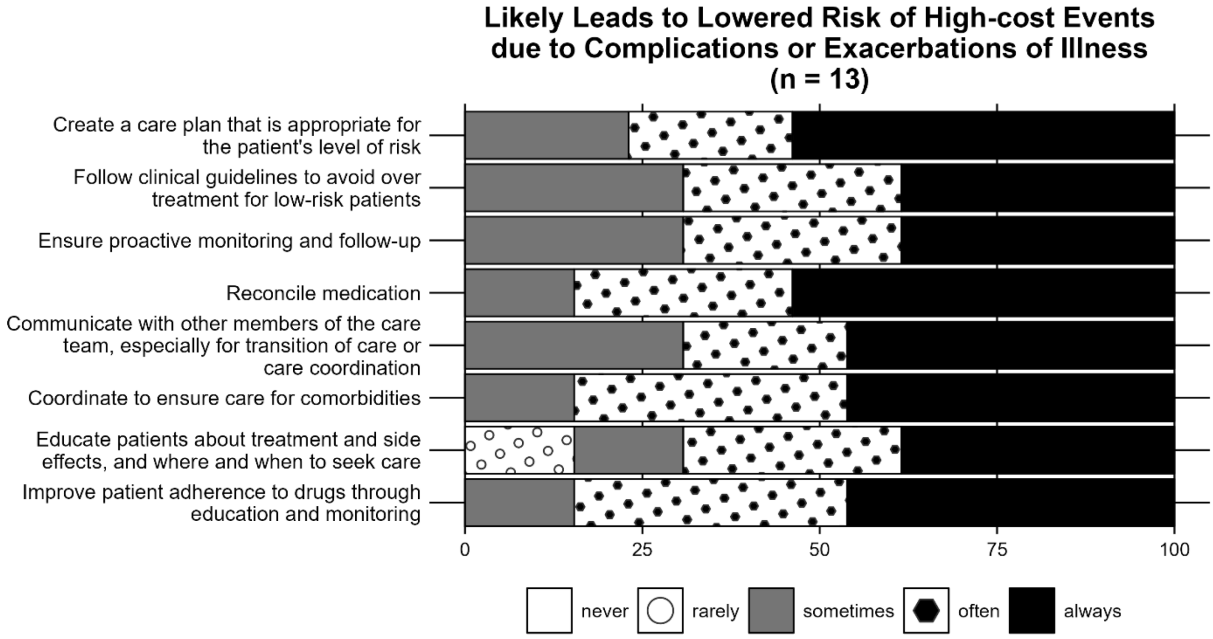


Figure 5: Reponses of Clinical Expert Workgroup Members when Asked to Rate the Likelihood of Improving Patient Treatment and Quality of Life for Actions Outlined in the Logic Model

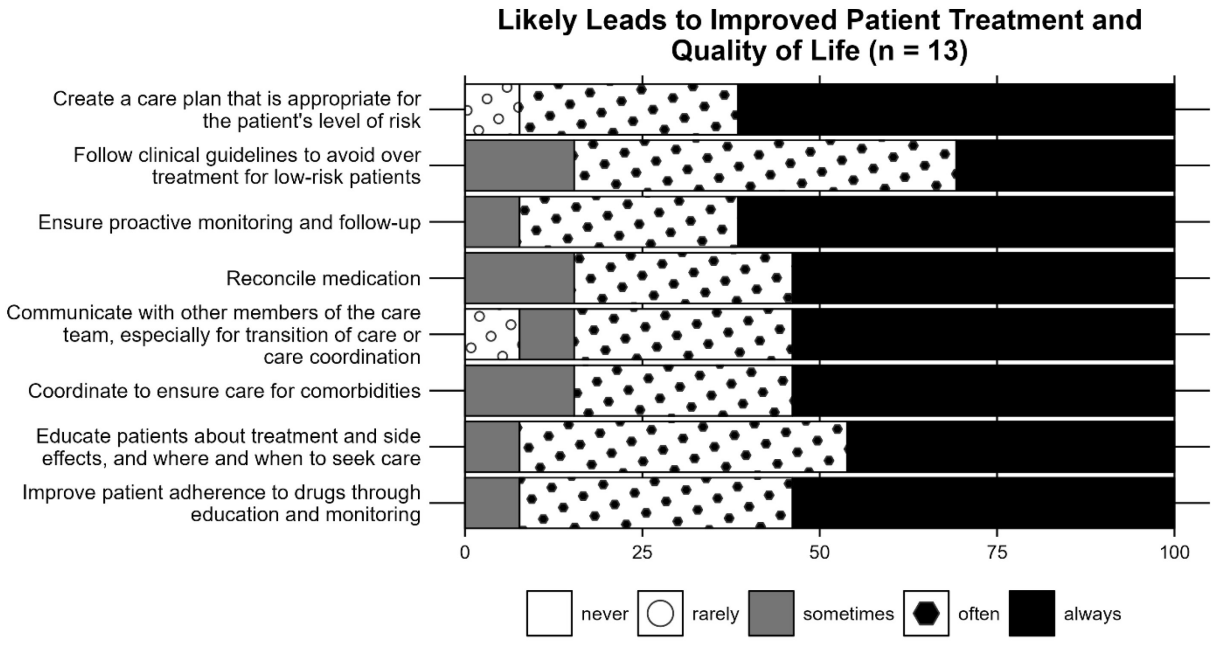


Figure 6: Reponses of Clinical Expert Workgroup Members when Asked to Rate the Likelihood of Improving Monitoring and Care Coordination for Actions Outlined in the Logic Model

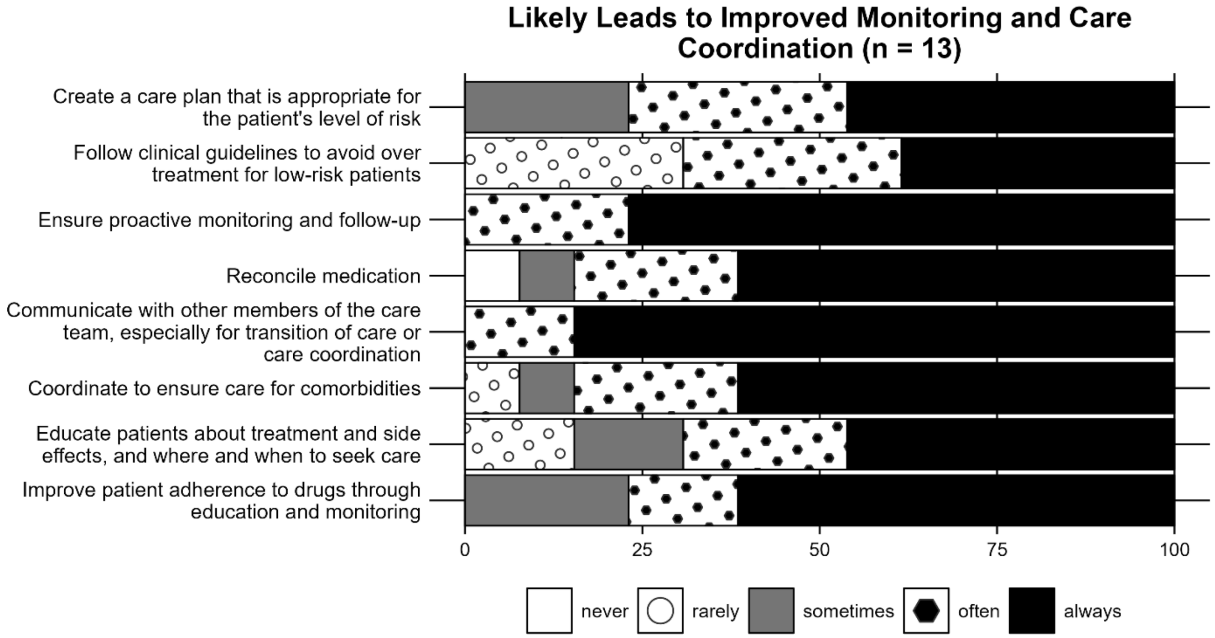
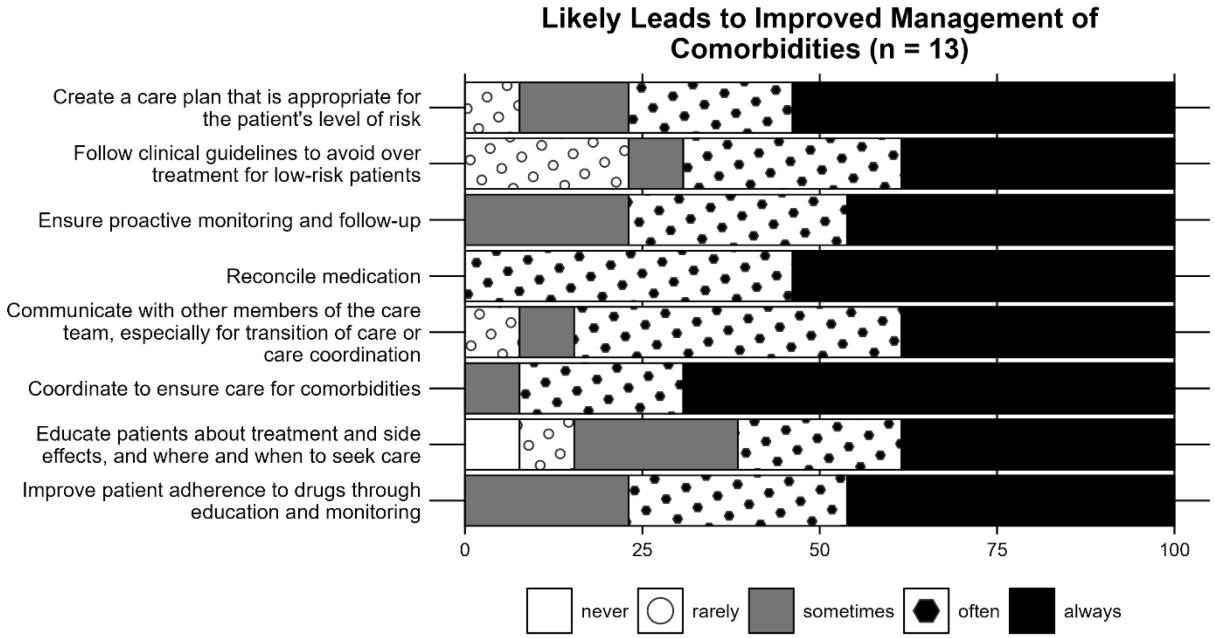


Figure 7: Reponses of Clinical Expert Workgroup Members when Asked to Rate the Likelihood of Improving Management of Comorbidities for Actions Outlined in the Logic Model



Empirical Validity Testing

Table 6 shows two regression models for each reporting level. Model 1 shows the effect on the clinicians' mean observed cost to expected cost ratio (O/E) for each additional one thousand dollars of a service category that is assigned to an episode, on average, while holding the remaining categories of cost constant. Model 2 shows the effect on the mean cost of adverse events for each additional one thousand dollars of a cost category that is assigned to an episode, on average, while holding the remaining categories of services constant.

Table 6: Estimated Effect of Treatment Choices

Categories of Service	Coefficient in Thousands [95% Confidence Interval] (p-value)			
	TIN		TIN-NPI	
	Model 1: Mean O/E = Mean Cost of Treatment Choices + Mean Cost of Adverse Events	Model 2: Mean Cost of Adverse Events = Mean Cost of Treatment Choices	Model 1: Mean O/E = Mean Cost of Treatment Choices + Mean Cost of Adverse Events	Model 2: Mean Cost of Adverse Events = Mean Cost of Treatment Choices
Adverse Outcomes	0.25 [0.24, 0.27] (p <0.01)	-	0.29 [0.29, 0.29] (p <0.01)	-
Inpatient Hospital: Trigger	0.00 [-0.00, 0.01] (p = 0.76)	0.06 [0.05, 0.07] (p <0.01)	-0.01 [-0.01, -0.00] (p <0.01)	0.09 [0.08, 0.09] (p <0.01)
Outpatient Evaluation & Management Services	0.09 [0.06, 0.12] (p <0.01)	0.04 [-0.03, 0.11] (p = 0.25)	0.23 [0.22, 0.24] (p <0.01)	0.04 [0.02, 0.06] (p <0.01)
Ambulatory/Minor Procedures	0.43 [0.38, 0.49] (p <0.01)	0.04 [-0.07, 0.16] (p = 0.45)	0.40 [0.38, 0.42] (p <0.01)	-0.05 [-0.09, -0.01] (p = 0.03)
Laboratory, Pathology, and Other Tests	0.31 [0.23, 0.39] (p <0.01)	-0.25 [-0.43, -0.07] (p <0.01)	0.23 [0.21, 0.25] (p <0.01)	-0.09 [-0.14, -0.04] (p <0.01)
Imaging Services	0.33 [0.26, 0.40] (p <0.01)	-0.14 [-0.30, 0.01] (p = 0.06)	0.35 [0.33, 0.37] (p <0.01)	-0.04 [-0.08, 0.01] (p = 0.11)
Anesthesia Services	1.41 [1.26, 1.56] (p <0.01)	-1.11 [-1.43, -0.78] (p <0.01)	0.41 [0.34, 0.48] (p <0.01)	-0.01 [-0.16, 0.14] (p = 0.91)
Chemotherapy and Other Part B-Covered Drugs	0.03 [-0.01, 0.07] (p = 0.15)	0.18 [0.09, 0.28] (p <0.01)	0.14 [0.13, 0.15] (p <0.01)	0.14 [0.12, 0.16] (p <0.01)
Physician Services During Hospitalization	0.01 [-0.02, 0.03] (p = 0.68)	0.16 [0.10, 0.22] (p <0.01)	0.07 [0.07, 0.08] (p <0.01)	0.12 [0.10, 0.14] (p <0.01)
Emergency Evaluation & Management Services	0.14 [0.12, 0.16] (p <0.01)	0.17 [0.12, 0.23] (p <0.01)	0.15 [0.15, 0.16] (p <0.01)	0.25 [0.24, 0.27] (p <0.01)
Emergency Room Procedures	0.92 [0.85, 0.99] (p <0.01)	-0.20 [-0.35, -0.04] (p = 0.01)	0.82 [0.77, 0.87] (p <0.01)	-0.93 [-1.04, -0.81] (p <0.01)
Emergency Room Laboratory, Pathology, and Other Tests	-2.44 [-4.78, -0.10] (p = 0.04)	-5.91 [-11.06, -0.76] (p = 0.02)	-0.58 [-1.24, 0.09] (p = 0.09)	6.49 [5.02, 7.96] (p <0.01)
Emergency Imaging Services	1.67 [1.41, 1.93] (p <0.01)	-0.55 [-1.13, 0.03] (p = 0.06)	1.13 [1.05, 1.22] (p <0.01)	-2.16 [-2.35, -1.96] (p <0.01)

3.3.4 Interpretation

Face Validity

Overall, there's overwhelmingly strong consensus among the members that the actions outlined in the logic model are often or always within the reasonable influence of an attributed clinician, lead to lowered risk of high-cost events due to complications or exacerbation of illness, lead to improving patient treatment and quality of care, lead to improving monitoring and care coordination, lead to improving management of comorbidities (Figures 3 – 7).

Empirical Validity Testing

Overall, the results demonstrate that the cost measure is reflective of both the cost directly related to treatment choices, as well as cost of adverse outcomes as a result of care (Table 6). Therefore, there's evidence that the measure is capturing what it purports to measure.

The results are also consistent with performance gaps identified from the literature review in section 2.2.1, such as preventing unscheduled ED revisits and adverse outcomes. Model 1 shoes that the cost of adverse events is associated with a worse measure score. We also see that while imaging services during the initial ER visit may result in an increase in the measure score by, it greatly reduces the cost of adverse effects ($2,160 at the TIN-NPI level). We observe similar trends for anesthesia services, emergency room procedures, laboratory, pathology and other tests. This pattern also suggests that, while these treatment choices are able to reduce the risk of adverse events and help improve the measure score, they may also be prone to overuse.

ER laboratory, pathology, and other tests are associated with worse score and higher cost of adverse outcomes, which may reflect higher service intensity that are linked to adverse outcomes and overall higher usage among sicker patients.

3.4 Exclusions Analysis

3.4.1 Method of Testing Exclusions

Exclusions are used in the Emergency Medicine measure to ensure a comparable patient population within the scope of the measure's focus on Medicare A/B beneficiaries who present to the ED with abdominal pain or chest pain and that episodes provide meaningful information to attributed clinicians. Exclusions are also used as part of data processing so that sufficient data are available to accurately determine episode spending and calculate risk adjustment for each episode.

For the exclusions analysis discussed in this section, we focused on exclusion criteria intended to ensure a comparable patient population.

- Standard exclusions to ensure data completeness
    - The patient has a primary payer other than Medicare for anytime overlapping the episode window or 120-day lookback period prior to the episode start date
    - The patient wasn't enrolled in Medicare Parts A and B for the entirety of the lookback period.
    - The patient's date of birth is missing.
    - The patient's death date occurred before the episode end date. These episodes were excluded as they may not accurately reflect a clinician's performance as the truncated episode window doesn't capture the full length of care intended by the measure.

This analysis also included exclusion criteria specific to the Emergency Medicine episode-based cost measure. The following episodes were excluded because they've different care pathways or characteristics that make them incomparable to rest of the episodes.

- Episodes with hospital-to-hospital transfers
- Episodes with ED-to-ED transfers
- Episodes with medical complications or that can't be classified into a visit type

Given the rationales for these exclusions, we would expect these excluded episodes to have a different profile than the included episodes, such as a higher mean cost, or a different distribution of costs (e.g., a long tail of high-cost episodes). For each exclusion, we examined the number of episodes and beneficiaries affected, as well as the distributions of observed cost. We then compared the cost characteristics of the excluded episodes to those of episodes included in measure calculation to assess the distinctness between the 2 patient cohorts. A full list of the exclusions used for the Emergency Medicine measure is provided in the Measure Codes List available on the MACRA Feedback Page.[36]

3.4.2 Statistical Results from Testing Exclusions

Table 6 below presents descriptive statistics of all episodes meeting the measure's triggering logic, excluded episodes, and final reportable episodes at both TIN and TIN-NPI levels. These exclusion criteria ensure that the reportable episode populations are more homogenous and comparable than all episodes meeting triggering logic.

[36]CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback.

Table 6: Cost Statistics for Measure Exclusions

Exclusion Criteria	Episodes		Observed Episode Cost					
				Percentile				
	Count	Percent of All Episodes Meeting Trigger Logic	Mean	10th	25th	50th	75th	90th
All Episodes Meeting Triggering Logic	17,736,867	100.00%	$5,891	$496	$876	$2,328	$8,351	$14,283
Beneficiary Death in Episode	434,964	2.45%	$13,301	$1,505	$6,956	$11,786	$14,911	$25,886
Outlier Cases	318,250	1.79%	$22,490	$633	$1,441	$16,765	$32,680	$51,087
Episodes with Medical Complications or Cannot Be Classified into a Visit Type	856,822	4.83%	$4,423	$391	$725	$1,414	$3,579	$12,964
Emergency Department to Emergency Department Transfer	135,281	0.76%	$4,746	$1,134	$1,764	$3,136	$5,147	$9,804
Hospital to Hospital Transfer	432,603	2.44%	$22,865	$8,037	$10,750	$16,074	$26,052	$43,913
TIN does not Meet Testing Volume Threshold	7,430	0.04%	$5,559	$290	$686	$1,898	$6,969	$14,031
TIN-NPI does not Meet Testing Volume Threshold	235,043	1.33%	$4,446	$333	$594	$1,217	$5,656	$12,195
Reportable Episodes (if all clinicians reported as TIN at the testing volume threshold)	15,590,546	87.90%	$4,997	$487	$849	$2,038	$7,599	$12,886
Reportable Episodes (if all clinicians reported as TIN-NPI at the testing volume threshold)	15,390,204	86.77%	$5,014	$490	$853	$2,060	$7,625	$12,907

3.4.3 Interpretation

Overall, the exclusion criteria decrease the mean episode cost slightly, from $5,891 to $4,997 at the TIN and $5,014 at the TIN-NPI level (Table 6).

Episodes with beneficiary death during an episode also have higher mean observed cost than all episodes meeting triggering logic, at $13,301 (Table 6). Excluding these episodes ensures that episodes are comparable and the observation window isn't truncated.

Episodes that aren't reliably predicted by the risk adjustment model are excluded because they deviate substantially from the projected cost for a given patient risk profile than the all episodes meeting triggering logic, with mean observed cost of $22,490 (Table 6).

Based on the input of stakeholders during the development process, the following episodes are excluded because they've different care pathways or characteristics that make them incomparable to the rest of the episodes: episodes with medical complications or can't be classified into a visit type defined by the clinical expert workgroup, emergency department to emergency department transfers, and hospital to hospital transfers. More information on the definitions of these criteria can be found in the Cost Measure Methodology. Except for episodes

with hospital to hospital transfers, the mean observed costs for episodes with medical complications or can't be classified into a visit type and ED to ED transfers are marginal lower than that of all episode meeting triggering logic (Table 6).

The last exclusion criteria come from applying the testing volume threshold to ensure that there's sufficient sample size to calculate the measure for providers, which have similar mean observed cost than the rest of the population.

3.5 Risk Adjustment or Stratification

3.5.1 Method of Controlling for Differences

Differences in case mix are controlled for using a statistical risk model with 104 risk factors, 29 visit types that are also stratified into: (i) subsequent observation care or inpatient admission, or (ii) discharged without subsequent observation care or inpatient admission.

The risk adjustment model for the Emergency Medicine measure adjusts for comorbidities based on the CMS Hierarchical Condition Category (HCC) model, count of HCCs, end-stage renal disease (ESRD) status, disability status, recent use of institutional long-term care, age, and dual eligibility status.

The model also includes measure-specific factors:

- Medicare Severity-Diagnosis Related Group (MS DRG) of inpatient stay for episodes that end in inpatient admission
- Episodes triggered in a critical access hospital (CAH)
- Episodes with transfers from an inpatient rehabilitation facility (IRF), long-term care hospital (LTCH), or skilled nursing facility (SNF)

A separate linear regression is run for each combination of visit type and whether the episode ended in subsequent observation care or inpatient admission to ensure fair comparison:

- Subsequent observation care or inpatient admission
- Discharged without subsequent observation care or inpatient admission

Full details of the risk adjustment model are in the Measure Codes List File available on the MACRA Feedback page.[37]

3.5.2 Conceptual, Clinical, and Statistical Methods

We selected the CMS-HCC model based on previous studies evaluating its appropriateness for use in risk adjusting Medicare claims data. This model was developed specifically for use in the Medicare population, meaning that it accounts for conditions found in the Medicare population. In addition, the CMS-HCC model is routinely updated for changes in coding practices (e.g., the transition from ICD-9 to ICD-10 codes). Because the CMS-HCC model has already been extensively tested, we focus our testing on the adaptation of the CMS-HCC model to the Emergency Medicine measure's patient population.

The workgroup provided input on measure-specific risk adjustors after reviewing empirical analyses on subpopulations of interest to assess whether and if so, how, particular factors should be accounted for in the model. These could include patient characteristics, factors outside of the reasonable influence of the clinician, or any other factors that would help prevent unintended consequences. These additional risk adjustors are listed in the section above.

[37]CMS, MACRA Feedback Page, https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback.

As previously noted, the risk adjustment model is run on episodes stratified into episode sub-groups, which may qualify as "ordering" of risk factors. Episode sub-groups were also determined based on the workgroup's input, with the goal of ensuring clinical comparability among episodes so that the cost measure fairly compares clinicians with similar patient case-mix.

3.5.3 Conceptual Model of Impact of Social Risks

Figure 2 in Section 3.3.2 shows the conceptual model that outlines how SRFs can influence the measure score, which is informed by both published external research and our own data analysis.[38,39,40,41,42] The conceptual model outlines risk factors that are either known by the literature or informed by the Clinical Expert Workgroup to be within or outside of influence of the attributed clinician. Risk factors, including SRFs, can both influence the treatment choices and impact the size of the effect of treatment choices on mitigating risk of adverse outcomes and the cost of adverse outcomes.

A systematic approach then guides the decision of which factors to include in the risk adjustment model. First, we reviewed the literature to gather known risk factors and drivers of resource use. These factors are usually diagnoses, therefore the first set of risk adjustors are commonly the HCCs. Then, we consulted our clinical expert panels on additional factors that are known to be associated with resource use. Together with our clinical expert panel, we reviewed the stratified results on episode cost across many different patient characteristics. We arrived at the final list of risk adjustors based on those discussions and consensus among the clinical experts. Additionally, during our testing phases, we also follow a structured and systematic approach to decide whether SRFs should be adjusted for, which is further described in Section 3.5.5.

3.5.4 Statistical Results

The literature has extensively tested the use of the HCC model as applied to Medicare claims data. Although the variables in the HCC model were chosen to predict annual cost, CMS has also used this risk adjustment model in a number of other settings (e.g., Accountable Care Organizations, previous physician Quality and Resource Use Report programs, and other administrative claims-based measures such as the Knee Arthroplasty episode-based cost measure, Total Per Capita Cost (TPCC) cost measure, Medicare Spending Per Beneficiary (MSPB)-PAC cost measure and MSPB-Hospital cost measure). Recalling that the risk model relies on the existing CMS-HCC model, testing results for factors included in the CMS-HCC V22 2016 model can be found in the Evaluation of the CMS-HCC Risk-Adjustment Model report[43] and the Report to Congress: Risk Adjustment in Medicare Advantage[44]. For measure-specific factors not included in the CMS-HCC model, we sought expert clinician input through the

---

[38] See footnote 15.

[39] Assistant Secretary of Health and Human Services for Planning and Evaluation. Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs. Washington, D.C. December 2016.

[40] Chen LM, Epstein AM, Orav EJ, Filice CE, Samson LW, Joynt Maddox KE. Association of Practice-Level Social and Medical Risk With Performance in the Medicare Physician Value-Based Payment Modifier Program. JAMA. 2017;318(5):453-461

[41] Medicare Payment Advisory Commission. Beneficiaries Dually Eligible for Medicare and Medicaid. 2018; https://www.macpac.gov/publication/data-book-beneficiaries-dually-eligible-for-medicare-and-medicaid-3/.

[42] Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health & Human Services. Second Report to Congress on Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program. 2020. https://aspe.hhs.gov/social-risk-factors-and-medicares-value-based-purchasing-programs

[43] Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

[44] CMS, "Report to Congress: Risk Adjustment in Medicare Advantage," https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf.

workgroup, which provided recommendations on additional risk adjustors and measure sub-groups.

3.5.5 Analyses and Interpretation in Selection of Social Risk Factors

To determine whether it's appropriate to risk adjust for SRFs, the following criteria are considered:

(i) whether there's an association between social risk and performance by examining the coefficient of patient-level dual status when added into the risk model,

(ii) whether the observed association is most influenced by patient-level factors or clinician-level factors by examining the stability of the patient-level dual status coefficient after adding clinician's dual share variable, as well as including clinician's fixed effects,

(iii) whether patient's need or complexity rather than poor quality is driving the observed performance differences by examining the differences in performance on dual patients versus non-dual patients and if there are many clinicians who are able to perform similarly or better on their dual patients than their non-dual patients, and

(iv) the impact of risk adjusting for SRFs by examining the performance shift of clinicians compared to a risk adjustment model that doesn't risk adjust for SRFs.

Table 7: Coefficient of Patient-level Dual Status under Different Models at TIN Level

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit with Observation Care or Ending in IP Stay: Abdominal Pain, Nausea, And Vomiting	2.08%	1.46 (p: 0.95)	-3.88 (p: 0.87)	-29.22 (p: 0.22)
ED Visit with Observation Care or Ending in IP Stay: Altered Mental State	0.05%	-104.31 (p: 0.34)	-125.26 (p: 0.26)	-163.85 (p: 0.19)
ED Visit with Observation Care or Ending in IP Stay: Behavioral Health	1.49%	134.75 (p: <0.01)	78.97 (p: 0.05)	50.49 (p: 0.2)
ED Visit with Observation Care or Ending in IP Stay: Cancer	0.85%	-377.1 (p: <0.01)	-346.63 (p: <0.01)	-435.8 (p: <0.01)
ED Visit with Observation Care or Ending in IP Stay: Diabetes	0.46%	29.1 (p: 0.33)	8.6 (p: 0.78)	-5.87 (p: 0.85)
ED Visit with Observation Care or Ending in IP Stay: ENT and Eye Disorders	0.19%	-204.25 (p: <0.01)	-216.21 (p: <0.01)	-172.79 (p: 0.02)
ED Visit with Observation Care or Ending in IP Stay: Fracture	0.16%	-170.06 (p: <0.01)	-158.51 (p: 0.01)	-119.5 (p: 0.06)
ED Visit with Observation Care or Ending in IP Stay: Gastrointestinal or Liver Conditions	5.2%	75.16 (p: <0.01)	63.58 (p: <0.01)	41.54 (p: <0.01)
ED Visit with Observation Care or Ending in IP Stay: General Infection	1.28%	272.94 (p: 0.01)	190.79 (p: 0.06)	139.81 (p: 0.18)

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit with Observation Care or Ending in IP Stay: Gynecological Disorders	0.04%	-3.79 (p: 0.97)	-29.73 (p: 0.76)	-3.93 (p: 0.97)
ED Visit with Observation Care or Ending in IP Stay: Healthcare Maintenance	0.01%	36.8 (p: 0.67)	7.3 (p: 0.93)	4.57 (p: 0.97)
ED Visit with Observation Care or Ending in IP Stay: Hematologic and Immunologic	1.46%	-28.89 (p: 0.73)	-48.29 (p: 0.57)	-67.63 (p: 0.43)
ED Visit with Observation Care or Ending in IP Stay: Kidney and Urinary	3.74%	18.43 (p: 0.1)	20.28 (p: 0.08)	-4.53 (p: 0.7)
ED Visit with Observation Care or Ending in IP Stay: Neurologic	1.45%	37.39 (p: 0.23)	13.51 (p: 0.67)	3.24 (p: 0.92)
ED Visit with Observation Care or Ending in IP Stay: Non-Diabetic Endocrine	0.01%	17.38 (p: 0.91)	19.08 (p: 0.9)	-167.21 (p: 0.45)
ED Visit with Observation Care or Ending in IP Stay: Non-Fracture Musculoskeletal	1.26%	42.15 (p: 0.2)	14.31 (p: 0.67)	-10.4 (p: 0.76)
ED Visit with Observation Care or Ending in IP Stay: Non-Respiratory Chest Pain	1.75%	116.12 (p: <0.01)	69.86 (p: 0.02)	22.32 (p: 0.44)
ED Visit with Observation Care or Ending in IP Stay: Other Cardiovascular	5.93%	10.45 (p: 0.51)	0.92 (p: 0.95)	-45.81 (p: <0.01)
ED Visit with Observation Care or Ending in IP Stay: Peripheral Vasc	0.74%	39.58 (p: 0.36)	21.29 (p: 0.63)	11.95 (p: 0.79)
ED Visit with Observation Care or Ending in IP Stay: Poisoning	0.01%	-702.66 (p: 0.02)	-743.89 (p: 0.02)	-504.35 (p: 0.25)
ED Visit with Observation Care or Ending in IP Stay: Pregnancy	0.01%	-604.73 (p: 0.09)	-592.8 (p: 0.09)	-355.87 (p: 0.51)
ED Visit with Observation Care or Ending in IP Stay: Respiratory	4.7%	77.93 (p: <0.01)	52.99 (p: <0.01)	28.23 (p: 0.07)
ED Visit with Observation Care or Ending in IP Stay: Sepsis	3.18%	25.28 (p: 0.16)	-10.92 (p: 0.56)	-56.16 (p: <0.01)
ED Visit with Observation Care or Ending in IP Stay: Skin Conditions, Rashes, and Abscesses	0.28%	-26.49 (p: 0.67)	-51.53 (p: 0.41)	-92.4 (p: 0.15)
ED Visit with Observation Care or Ending in IP Stay: Stroke	1.59%	17.89 (p: 0.43)	-9.75 (p: 0.67)	-20.4 (p: 0.38)
ED Visit with Observation Care or Ending in IP Stay: Syncope	0.38%	-17.33 (p: 0.63)	4.83 (p: 0.9)	-18.62 (p: 0.63)

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit with Observation Care or Ending in IP Stay: Trauma: Major or Head	1.21%	96.13 (p: <0.01)	37.51 (p: 0.27)	2.9 (p: 0.93)
ED Visit with Observation Care or Ending in IP Stay: Trauma: Minor or Unclear Severity	0.18%	-294.79 (p: 0.13)	-372.94 (p: 0.06)	-427.64 (p: 0.04)
ED Visit without Observation Care nor Ending in IP Stay: Abdominal Pain, Nausea, And Vomiting	5.4%	-81.6 (p: <0.01)	-72.64 (p: <0.01)	-78.18 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Altered Mental State	0.01%	315.45 (p: 0.67)	394.6 (p: 0.6)	367.02 (p: 0.73)
ED Visit without Observation Care nor Ending in IP Stay: Behavioral Health	1.94%	-82.81 (p: <0.01)	-93.8 (p: <0.01)	-71.69 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Cancer	0.39%	-488.63 (p: <0.01)	-526.48 (p: <0.01)	-493.36 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Diabetes	0.75%	12.76 (p: 0.51)	36.98 (p: 0.06)	25.73 (p: 0.21)
ED Visit without Observation Care nor Ending in IP Stay: ENT and Eye Disorders	1.8%	-27.57 (p: 0.01)	-24.1 (p: 0.02)	-26.38 (p: 0.01)
ED Visit without Observation Care nor Ending in IP Stay: Fracture	1.23%	-377.41 (p: <0.01)	-338.49 (p: <0.01)	-341.58 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Gastrointestinal or Liver Conditions	3.42%	-1.37 (p: 0.91)	4.91 (p: 0.69)	-0.95 (p: 0.94)
ED Visit without Observation Care nor Ending in IP Stay: General Infection	1.2%	1.53 (p: 0.95)	6.63 (p: 0.77)	-3.02 (p: 0.9)
ED Visit without Observation Care nor Ending in IP Stay: Gynecological Disorders	0.15%	-33.81 (p: 0.23)	-33.61 (p: 0.24)	-39.32 (p: 0.19)
ED Visit without Observation Care nor Ending in IP Stay: Healthcare Maintenance	0.11%	-11.57 (p: 0.56)	-15.15 (p: 0.45)	-20.97 (p: 0.32)
ED Visit without Observation Care nor Ending in IP Stay: Hematologic and Immunologic	1.59%	-191.29 (p: <0.01)	-188.51 (p: <0.01)	-208.94 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Kidney and Urinary	6.54%	-98.68 (p: <0.01)	-89.06 (p: <0.01)	-91.24 (p: <0.01)

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit without Observation Care nor Ending in IP Stay: Neurologic	2.3%	-95.18 (p: <0.01)	-88.52 (p: <0.01)	-81.98 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Non-Diabetic Endocrine	0.1%	-106.66 (p: 0.08)	-79.16 (p: 0.21)	-101.87 (p: 0.13)
ED Visit without Observation Care nor Ending in IP Stay: Non-Fracture Musculoskeletal	5.15%	-135.81 (p: <0.01)	-117.85 (p: <0.01)	-125.28 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Non-Respiratory Chest Pain	2.34%	-114.49 (p: <0.01)	-105.68 (p: <0.01)	-113.02 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Other Cardiovascular	7.29%	-100.55 (p: <0.01)	-91.41 (p: <0.01)	-96.64 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Peripheral Vasc	0.74%	-4.92 (p: 0.89)	6.6 (p: 0.85)	16.4 (p: 0.66)
ED Visit without Observation Care nor Ending in IP Stay: Poisoning	0.23%	0.14 (p: 1)	2.14 (p: 0.94)	5.66 (p: 0.86)
ED Visit without Observation Care nor Ending in IP Stay: Pregnancy	0.02%	70.94 (p: 0.52)	-61.62 (p: 0.57)	122.54 (p: 0.37)
ED Visit without Observation Care nor Ending in IP Stay: Respiratory	5.76%	-67.1 (p: <0.01)	-69.9 (p: <0.01)	-77.48 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Sepsis	0.22%	283.68 (p: 0.07)	447.39 (p: <0.01)	292.15 (p: 0.07)
ED Visit without Observation Care nor Ending in IP Stay: Skin Conditions, Rashes, and Abscesses	1.97%	-39.6 (p: <0.01)	-31.3 (p: <0.01)	-38.64 (p: <0.01)
ED Visit without Observation Care nor Ending in IP Stay: Stroke	0.35%	111.04 (p: 0.07)	144.16 (p: 0.02)	135.54 (p: 0.04)
ED Visit without Observation Care nor Ending in IP Stay: Syncope	1.55%	17.33 (p: 0.32)	22.66 (p: 0.2)	11.12 (p: 0.54)
ED Visit without Observation Care nor Ending in IP Stay: Trauma: Major or Head	3.38%	-20.04 (p: 0.15)	-18.17 (p: 0.2)	-13.12 (p: 0.35)
ED Visit without Observation Care nor Ending in IP Stay: Trauma: Minor or Unclear Severity	4.41%	16.46 (p: 0.07)	18.62 (p: 0.04)	17.23 (p: 0.06)

Please see Appendix A for the TIN-NPI level results.

Table 8: Mean Ratio of Episode Observed Cost to Expected Cost (O/E) Stratified by Clinician's Dual Share and Patient's Dual Status

Dual Share	TIN			TIN-NPI		
	All Episode	Dual Episodes	Non-Dual Episodes	All Episodes	Dual Episodes	Non-Dual Episodes
All	1.04	1.04	1.05	0.98	0.98	0.98
0%	1.04	-	1.04	0.92	-	0.92
1-20%	1.08	1.08	1.08	0.99	0.99	0.99
21-40%	1.05	1.04	1.05	0.99	0.99	0.99
41-60%	1.01	1.01	1.01	0.97	0.97	0.97
61-80%	1.01	1.00	1.04	0.94	0.95	0.93
81-99%	1.08	1.08	1.04	0.97	0.97	0.96
100%	1.05	1.05	-	0.97	0.97	-

Table 9. Proportions of Clinicians Who Perform Significantly Worst, Equally Well, or Significantly Better on Their Dual Episodes than Non-Dual Episodes

Reporting Level	Significantly Worse	Equally Well	Significantly Better
TIN	6%	88%	7%
TIN-NPI	3%	95%	2%

Table 10. Clinicians' Performance Shift Measured by the Change in the Average Ratio of Observed-to-Expected Cost

TIN or TIN-NPI	Dual Share	Proportion of Clinicians Affected at Various Levels of Performance Shift	
		Ranking Shift by 1% or more	Ranking Shift by 5% or more
TIN	All	40.9%	0.4%
TIN	0%	25%	0%
TIN	1-20%	42.2%	0.9%
TIN	21-40%	33.4%	0.2%
TIN	41-60%	48.8%	0%
TIN	61-80%	60.6%	1.6%
TIN	81-99%	61.4%	2.3%
TIN	100%	66.7%	0%
TIN-NPI	All	46.9%	0.8%
TIN-NPI	0%	43.4%	1.4%
TIN-NPI	1-20%	53.4%	0.4%
TIN-NPI	21-40%	41.4%	0.5%
TIN-NPI	41-60%	47.2%	0.7%
TIN-NPI	61-80%	61.3%	2.2%
TIN-NPI	81-99%	65.7%	4.7%
TIN-NPI	100%	64.7%	0%

The results are mixed; however, there's evidence to suggest that it's appropriate to risk adjust for social risk factors. In Table 7 we see that patient level factors appear to be slightly more influential to the cost than provider level factors. Additionally, Table 8 shows that the degradation of episodes is elevated when there's a greater share of dual eligible patients (about 60% or more). In Table 9, we do see that the majority of clinicians perform equally well; however, at the TIN level, there are about 6-7% of clinicians whose performance becomes significantly impacted. In Table 10, we observe a similar pattern where the proportion of clinicians affected in their ranking increases greatly when there's a share of 60% or more dual eligible patients. Due to this reoccurring trend where clinicians who treat a greater share of dual eligible clinicians' performances are impacted more, it's appropriate to risk adjust for social risk factors.

3.5.6 Method for Statistical Model or Stratification Development

To analyze the validity of current risk adjustment model, we examined two criteria: discrimination and calibration.

1) Discrimination is a statistical criterion that evaluates the measure's ability to distinguish high-cost episodes from low-cost episodes, or the ability to explain the variance in cost of individual episodes. The amount of variance explained is estimated by the R-squared metric with the range between 0 and 1. These results are provided in Section 3.5.7.
2) Calibration evaluates the consistency of the measure in estimating episode cost across the full range of resource use patterns in the population. Calibration is estimated by the average predictive ratios across groups within the population, specifically groups are partitioned by deciles of expected episode cost. A well-calibrated measure should have predictive ratios close to 1.0 across all deciles. These are discussed in Sections 3.5.8 and 3.5.9.

3.5.7 Statistical Risk Model Discrimination Statistics

The overall R-squared for the Emergency Medicine cost measure, calculated by dividing explained sum of squares by total sum of squares, is 0.61. The adjusted R-squared is 0.61. More information on discrimination testing for the CMS-HCC model can be found at Pope et al. 2011.[45]

3.5.8 Statistical Risk Model Calibration Statistics

The predictive ratio is calculated using the formula of average expected cost / average observed cost for all episodes in each decile.

3.5.9 Statistical Risk Model Calibration – Risk Decile

Analysis of predictive ratios by risk decile for the measure shows minimal variation among risk deciles, as predictive ratios range from 0.99 to 1.01 across all risk deciles (with an overall average of 1.00). These results suggest that the Emergency Medicine measure is well-calibrated across the full range of resource use patterns observed in the population.

Table 11: Predictive Ratio by Decile of Predicted Episode Cost

Decile	Average Predictive Ratio
Decile 1	1.01
Decile 2	1.01
Decile 3	1.00
Decile 4	0.99

[45]Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

Decile	Average Predictive Ratio
Decile 5	1.00
Decile 6	0.99
Decile 7	1.00
Decile 8	1.00
Decile 9	1.00
Decile 10	1.00

3.5.10 Interpretation

The R-squared values for the model, which measure the percentage of variation in results predicted by the model, are similar to or higher than the values presented in similar analyses of risk adjustment models.[46] As noted in Section 3.5.6 and 3.5.7, these results should be interpreted alongside service assignment rules, which remove clinically unrelated services.

The remaining unexplained variance is due to variation in factors that aren't adjusted for by the measure, such as the clinician's performance. The objective of a cost measure is to evaluate and differentiate the performance of clinicians. Therefore, achieving high explained variance isn't essential because not all of the variation in cost of care should be adjusted. In collaboration with the experts from our clinical workgroup, this measure only adjusts for factors that are deemed to be outside of the influence of clinicians.

Table 11 shows that the risk adjustment model is consistent, with the average predictive ratios observed to be close to 1.00 across all deciles, with the range between 0.99 and 1.01. Overall, the risk adjustment model doesn't over- or under-predict cost across the full range of resource use patterns in the population.

3.6 Identification of Meaningful Differences in Performance

3.6.1 Method

To identify meaningful differences in performance, this analysis first examines the distribution of the measure score to highlight the performance gap between the most and least efficient clinicians. Then, this analysis examines the rate of high-cost events that may occur during an episode of care to highlight the variation in frequency and cost of those events.

3.6.2 Statistical Results

Table 1 shows the distribution of the measure score at the TIN and TIN-NPI levels. Additionally, the testing results show that 14.5% of episodes had an emergency department revisit with a mean risk-adjusted episode cost of $8,498, and 4% of episodes that had an inpatient admission after discharge from the emergency department (not including any admission directly from the emergency department) with a mean risk-adjusted episode cost of $17,541.

3.6.3 Interpretation

There are variations observed in the measure score in both TIN and TIN-NPI levels, indicated by the interquartile ranges, standard deviations, and coefficients of variation. The magnitude of the observed variation is in the thousands of dollars, which indicates that there are opportunities to close the gaps between the most and least efficient clinicians. There are also opportunities to

[46]Pope, Gregory C., John Kautter, Melvin J. Ingber, Sara Freeman, Rishi Sekar, and Cordon Newhart. "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

reduce costs associated with high-cost events, such as emergency department revisits and inpatient admissions. On average, episodes with an emergency department revisit cost Medicare approximately $16 billion more than an average episode, and $10 billion for episodes with an inpatient admission after discharge.

3.7 Missing Data Analysis and Minimizing Bias

3.7.1 Method

Since CMS uses Medicare claims data to calculate the Emergency Medicine measure, Acumen expects a high degree of data completeness. To further ensure that we've complete and accurate data for each patient, Acumen typically excludes episodes where patient date of birth information (an input to the risk adjustment model) can't be found in the enrollment database, the patient doesn't appear in the enrollment database, patient resides outside of the U.S., death occurred before the episode, the primary payer isn't Medicare, or episodes with invalid cost for the trigger claim.

The Emergency Medicine measure also excludes episodes where the patient is enrolled in Medicare Part C or has a primary payer other than Medicare in the 120-day lookback period and episode window. In such situations, Medicare Parts A and B claims data may not capture the complete clinical profile for the patient needed to capture the clinical risk of the patient in risk adjustment. Furthermore, Parts A and B claims data may not capture all Medicare resource use if some portion of the patient's care is covered under Medicare Part C.

3.7.2 Missing Data Analysis

Table 12 presents the frequency and observed episode cost for categories of missing data, which caused episodes to be excluded from the Emergency Medicine measure. Frequency is presented in terms of the number of episodes excluded due to missing data, as well as the cost profile. It's worth noting that only the observed cost is shown, which hasn't been risk adjusted for using our risk adjustment model. Therefore, the differences in cost may appear much smaller after risk adjustment than as-is.

As a note, the episode counts below reflect exclusion from the initial population of triggered episodes. After the missing data exclusions are applied, we then apply additional exclusions, as outlined in section 3.4, to this overall patient cohort to narrow the population to only applicable episodes.

Table 12: Cost Statistics for Missing Data Category

Missing Data Categories	Episode Count	Observed Episode Cost					
		Mean	Percentile				
			10th	25th	50th	75th	90th
All Episodes	20,799,002	$5,741	$481	$855	$2,160	$8,109	$14,084
Primary Payer Other than Medicare	2,099,160	$4,997	$390	$742	$1,555	$6,782	$12,974
Beneficiary Death before Trigger	78,297	$4,023	$554	$945	$1,385	$3,049	$12,276
No Continuous Enrollment in Medicare Parts A and B, and Any Enrollment in Part C	1,040,389	$4,613	$351	$719	$1,556	$5,818	$12,136
Episodes with Invalid Cost for the Trigger Claim	23	$4,785	$1,153	$1,947	$3,574	$7,534	$8,891

3.7.3 Interpretation

The results show that the missing data episodes don't appear to be substantially different than all episodes in the initial population in terms of cost (Table 12). Given their limited frequencies, the impact of removing these episodes on the overall measure should be minimal while ensuring that clinicians are fairly evaluated on episodes with complete data.

4.0 Feasibility

4.1 Data Elements Generated as Byproduct of Care Processes

The data elements used in this measure are pulled from Medicare claims. They can be based on information generated, collected and/or used by healthcare personnel during the provision of care (e.g., diagnoses), which are then translated into the appropriate coding system (e.g. ICD-10 diagnoses, Medicare Severity Diagnosis-Related Groups [MS-DRGs]) for use in Medicare claims by either the original healthcare personnel or another individual.

4.2 Electronic Sources

All data elements are in defined fields in electronic claims.

4.3 Data Collection Strategy

4.3.1 Data Collection Strategy Difficulties

Lessons and associated modifications may be categorized into three types: data collection procedures, handling of missing data, and sampling data associated with beneficiaries who died during an episode of care.

4.3.1.1 Data Collection

Acumen receives claims data directly from the Common Working File (CWF) maintained at the CMS Baltimore Data Center. Medicare claims are submitted by healthcare providers to a Medicare Administrative Contractor (MAC), and are subsequently added to the CWF. However, these claims may be denied or disputed by the MAC, leading to changes to historical CWF data. In rare circumstances, finalizing claims may take many months, or even years. As a result, it isn't practical to wait until all claims for a given month are finalized before calculating this measure. As such, there's a trade-off between efficiency (accessing the data in a timely manner) and accuracy (waiting until most claims are finalized) when determining the length of the time (i.e., the "claims run-out" period) after which to pull claims data. To determine the appropriate claims run-out period, Acumen has performed testing on the delay between claim service dates and claims data finalization. Based on this analysis, Acumen uses a run-out period of three months after the end of the calendar year to collect data for development and testing purposes. If this measure is used in a CMS program, calculation and reporting would be done in line with that program's reporting practices.

4.3.1.2 Missing Data

This measure requires complete beneficiary information, and a small number of episodes with missing data are excluded to ensure completeness of data and accurate comparability across episodes. For example, episodes where the beneficiary wasn't enrolled in Medicare Parts A and B for the 120 days prior to the episode start date aren't included in this measure. This enables the risk adjustment model to accurately adjust for the beneficiary's comorbidities using data from the previous 120 days of Medicare claims. Additionally, the risk adjustment model includes a categorical variable for beneficiary age bracket, so episodes for which the beneficiary's date of birth can't be located aren't included in this measure.

4.3.1.3 Sampling

During measure testing, Acumen noted that episodes in which the beneficiary died prior to the episode end date exhibited different cost distributions compared to other episodes. To avoid this effect's potential impact on clinician scores, this measure doesn't include episodes for which the beneficiary's date of death occurs prior to the end of the episode window.

5.0 Usability and Use

5.1 Use

5.1.1 Current and Planned Use

The Emergency Medicine measure isn't currently in use, but is intended for use in a payment program and could eventually be publicly reported. The measure was specifically developed for potential use in the Cost performance category of MIPS to assess clinicians reporting as individuals or groups, under a contract with CMS.

For the measure to be used in MIPS, it must be reviewed by the Measure Application Partnership (MAP) and then undergo the notice-and-rulemaking process. Given these next steps, the earliest the measure could be in use in MIPS is CY 2024. If in use, CMS can then determine whether to publicly report the cost measure.

5.1.2 Feedback on the Measure by Those being Measured or Others

Throughout the Emergency Medicine measure development, we used an iterative and extensive process to gather feedback on the measure and its results to ensure that the measure can be used appropriately in the MIPS program by clinicians and clinician groups who practice in this clinical area. This process also aims to make sure that the measure performance results can be understood by the population that is being measured to help support decision making. A couple of the main ways that we gathered feedback was through i) reoccurring Clinician Expert Workgroup meetings, where members discussed the clinical perspective, the patient perspective, and empirical data, in order to recommend measure specifications, and ii) the national field testing of the measure.

5.1.2.1 Technical Assistance Provided During Development or Implementation

Clinician Expert Workgroup Meetings

For each Clinician Expert Workgroup meeting, Acumen provided empirical data (e.g., analyses on potentially relevant services to group and potential sub-populations to sub-group, risk adjust, or exclude). These analyses were conducted using all administrative claims data for Medicare Parts A and B. This data was shared with Workgroup members to help inform their feedback on the measure specifications throughout its development to ensure that the measure was appropriately assessing costs for the attributed clinicians.

Field Testing

Additionally, Acumen and CMS nationally field tested the draft Emergency Medicine measure, along with 4 other episode-based cost measures, for a 10-week comment period (January 10 to March 25, 2022). We provided a Field Test Report with performance data to all clinician groups and clinicians who were attributed 20 or more episodes.[47] This testing sample was selected to balance coverage and reliability, since a key goal of field testing was to test the measures with as many clinicians and other interested members of the public as possible. A total of 4,071 TIN reports and 79,540 TIN-NPI reports were developed for this measure. During this time, feedback was gathered on the usability of the performance data and the appropriateness of the measure.

[47]The field test reports were available for download from the Quality Payment Program website: https://qpp.cms.gov/login.

5.1.2.2 Technical Assistance with Results

Clinician Expert Workgroup Meetings

Acumen provided data in advance of or during each of the Clinician Expert Workgroup Meetings: Workgroup meeting, Service Assignment and Refinement Meeting, Post-Field Test Refinement Meeting. During the meetings, Acumen would guide Workgroup members through these analyses, providing clinical and programmatic context when needed. Using this iterative process, the Workgroup members discussed the testing results in depth during each meeting and allowed the data to inform their recommendations for measure specifications. The goal was to ensure that the measure was appropriately assessing clinicians cost of care within their reasonable influence, without creating potential unintended consequences so that it could be usable in the MIPS program.

Field Testing

During the field testing period, feedback on the appropriateness of the measures and the usability of the data was gathered from clinician and clinician groups who received a report as well as the general public. Comments from field testing were summarized in a public report, which was also shared with the Clinician Expert Workgroup to consider in recommending refinements to the measures based on the testing data and feedback.

The following sections offer more details on the contents of each report and describe the education and outreach efforts associated with the field testing feedback period.

5.1.2.2.1 Data Provided During Field Testing

Each Field Test Report contained:

- Detailed performance results for the attributed measure, including cost measure score and breakdown of episode cost compared to the national average and TIN/TIN-NPIs with a similar patient case mix (or risk profile).
- Drill-down detail for each measure, including more detailed information on potential cost drivers in the TIN/TIN-NPI's episodes. For example:
    - Analysis of utilization and cost for the measure by the Restructured BETOS Classification System (e.g., outpatient evaluation and management services, procedures, and therapy, hospital inpatient services, emergency room services, post-acute services)[48]
    - Breakdown of costs for Part B Physician/Supplier and inpatient claims (e.g., top 5 most billed services and by risk bracket)
    - Accompanying episode-level Comma Separated Value (CSV) file with detailed information for all episodes attributed to the TIN/TIN-NPI. This file provides detailed information on every episode used to calculate your measure score, which includes winsorized observed cost, risk-adjusted cost, facilities and clinicians rendering care, the share of cost by service setting, the patient relationship code (PRC) on the trigger/reaffirming claim line.

All interested members of the public, including those who did not qualify to receive a Field Test Report, could review a series of mock reports that were representative of each measure and reporting type. Other public documentation posted during field testing included: measure specifications for each measure (comprising a Draft Cost Measure Methodology document and a Draft Measure Codes List file), a Measure Development Process document, a Frequently Asked Questions document, a Measure Testing Form (including reliability and validity data), and

[48]CMS, "Restructured BETOS Classification System https://data.cms.gov/provider-summary-by-type-of-service/provider-service-classifications/restructured-betos-classification-system

a National Summary Data Report (including national level summary statistics on the measure).[49] During field testing, Acumen conducted education and outreach activities including multiple office hours sessions with specialty societies, a publicly posted field testing webinar recording, and Quality Payment Program Help Desk support.

5.1.2.2.2 Education and Outreach

Acumen directly conducted outreach via email to tens of thousands of outreach contacts using the contact list developed through previous education and outreach and clinician engagement efforts, as well as CMS, Quality Payment Program listservs. Acumen also sent emails directly to clinicians who received the field test reports via CMS's GovDelivery.

Acumen and CMS hosted two office hours sessions in January 2022 to provide an overview of field testing to specialty societies, discuss what information their members would be particularly interested in, and answer any questions. Across both office hours sessions, there were over 35 attendees from targeted specialty societies who are likely to have members who could be attributed the measure.

Acumen worked closely with Quality Payment Program Service Center to respond to inquiries during field testing and continued to answer questions after the feedback period ended.

Acumen and CMS posted the MACRA Wave 4 Cost Measures Field Testing Webinar to the Quality Payment Program Webinar Library at the start of the field testing period.[50] The webinar recording, slides, and transcript were publicly available for review throughout field testing. The webinar presentation outlined: (i) the cost measure field testing project (ii) the measure development and re-evaluation processes, and (iii) field testing activities.

5.1.2.3 Feedback on Measure Performance and Implementation

Clinician Expert Workgroup Meetings

Feedback from the Workgroup members was recorded throughout the meeting. More formal feedback was gathered using polls, typically requesting for votes on certain specifications or appropriateness of the measure. These polls were conducted following each meeting and on an ad hoc basis, as needed.

Field Testing

In total, Acumen received 64 survey responses and 19 comment letters, including from specialty societies representing large numbers of potentially attributed clinicians.

Survey responses and comment letters were collected via an online survey, which contained general and detailed questions on the reports themselves, questions on the supplemental documentation, and questions on the measure specifications.

5.1.2.4 Feedback from Measured Entities

Field Testing

The Field Testing Feedback Summary Report presents feedback gathered during the field testing period, including cross-measure feedback and measure-specific feedback.[51] The measure-specific feedback was used as the basis for the post-field testing refinements that

[49]The measure specifications, mock reports, Measure Development Process document, Frequently Asked Questions document, and testing documents are posted on the MACRA Feedback Page: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-Feedback.html.

[50]MACRA Wave 4 Cost Measures Field Testing Webinar materials are available on the Quality Payment Program Webinar Library: https://qpp.cms.gov/about/webinars.

[51]CMS, "2020 Field Testing Feedback Summary Report," MACRA Feedback Page, https://www.cms.gov/files/document/macra-2020-ft-feedback-summary-report.pdf.

were made to the measures. Overarching feedback about data that would be helpful for clinicians to receive was recorded and shared with CMS for future consideration. See Section 5.1.2.6 for post-field testing refinements made to the Emergency Medicine measure.

5.1.2.5 Feedback from Other Users

Person and Family Engagement

Acumen incorporated thoughtful input from patients and caregivers throughout the Emergency Medicine measure development process. Before each Clinical Expert Workgroup meeting, Person and Family Partners (PFPs) would provide input through focus groups and interviews to help inform the Workgroup's discussion. Attending PFPs would then present the findings for the Workgroup members, which would help shape the recommendations they made for the measure specifications. Some examples of feedback the PFP include shortening wait times before receiving care, improving communication between the ED clinicians, the primary care provider and the patient, and lastly, improving ED care coordination given the multiple specialties that are involved in ED care. With considerations of PFP findings, the Emergency Medicine measure includes HCFA Specialty codes.

5.1.2.6 Consideration of Feedback

Field Testing

Careful consideration was given to all feedback gathered during field testing, and several updates were made to the measure based on the recommendations of field testing commenters and the Clinician Expert Workgroup comprised of subject matter and measure-development experts. Acumen conducted analyses into potential adjustments that could be made to the measures to improve their ability to assess the intended clinician population.

After completing field testing, Acumen compiled the feedback provided through the survey and comment letters into a measure-specific report, which was then provided to the Clinician Expert Workgroup, along with the empirical analyses to inform their discussion and evaluation of any refinements needed to ensure that the measure is capturing what it was intended to capture.

The changes to the Emergency Medicine measure made after consideration of field testing analyses and feedback are:

- Measure Visit Types
    - Sub-group all Visit Types by whether the ED visit terminates in ED discharge or inpatient admission
    - Continue using the MS-DRG on inpatient claims to contribute to Visit Type mapping
    - Recategorize the 2 Visit Types Skin/Eye and Oral/Nasal/Skin into Skin Conditions/Rashes/Abscesses and Ear, Nose, Throat (ENT)/Eye Disorders, respectively
    - Rename Female Disorders to Gynecological Disorders
    - Create a Non-Diabetic Endocrine Visit Type
    - Include episodes with COVID-19 into an existing Visit Type (e.g., General Infection or Respiratory)
- Service assignment
    - Reduce the episode window from 30 days to 14 days
    - Exclude the following services:
        - Radiation oncology
        - Ambulance services
        - Part B drugs provided after the ED visit
- Risk adjustment

- Include and risk adjust Post-Acute Care to ED transfer episodes
- Risk adjust for episodes that occur in a CAH
- Risk adjust for patients' dual eligibility in Medicare/Medicaid

5.2 Usability

5.2.1 Improvement

The measure hasn't yet been implemented, and as such hasn't had influence over performance. Our testing suggests that there's a sufficiently large difference in measure scores among clinicians to meaningfully determine a difference in performance. The potential for this measure to distinguish between good and poor performance is promising in its ability to encourage improvement in cost efficient care.

Additionally, the face validity results suggest that the Clinician Expert Workgroup believes the measure assess care within the influence of the clinician and can positively impact care provision and coordination.

5.2.2 Unexpected Findings

There were no unexpected findings during the development and testing of this measure. The measure hasn't been implemented at this time, so we don't have data that confirms unexpected findings related to its implementation.

However, Acumen did consider potential unintended consequences of having a cost measure for this clinical area (e.g., potential stinting in care to receive a better cost score). For example, the empiric validity data previously presented in section 3.3 demonstrates that while providing additional E&M services is associated with a higher score, it's often mediated by an adverse event. If a provider attempted to stint on providing E&M services, this increases risk for high-cost issues downstream, so it isn't in providers' best interest to do so for their cost measure score.

Additionally, CMS monitors measures that are in use and has multiple processes in place to allow for changes to a measure if appropriate. These include i) annual maintenance for non-substantial changes and upkeep, ii) ad hoc maintenance if a specific issue occurs or a large change in clinical guidance takes place, and iii) measure reevaluation every three years where the suitability of a measure's specifications is comprehensively reassessed. If in the event the measure did have any unexpected findings, it would be identified and resolved through one of these methods.

5.2.3 Unexpected Benefits

Since the measure hasn't been implemented at this time, there are no testing results that identify unexpected benefits. However, many clinicians can only be assessed by the MSPB Clinician and TPCC measures in the cost performance category currently. This measure would provide a more tailored assessment of the care they've influence over, which many clinicians may prefer to be measured by compared to the population-based cost measures like MSPB Clinician or TPCC.

6.0 Related and Competing Measures

6.1 Relation to Other Measures

There are no competing measures with this measure. However, the following measures have been identified as potentially related.

Table 13. Quality Measures Potentially Relevant for the Emergency Medicine Episode Group

Measure Title	Measure ID	Measure Description	Measure Type
Emergency Medicine: Emergency Department Utilization of CT for Minor Blunt Head Trauma for Patients Aged 18 Years and Older (Efficiency)	Q415	Percentage of emergency department visits for patients aged 18 years and older who presented with a minor blunt head trauma who had a head CT for trauma ordered by an emergency care clinician who have an indication for head CT.	Efficiency – High Priority
Hospital-Wide All-cause Readmission (HWR)	Q458	Attributes outcomes to MIPS participating clinician groups and assesses each group's readmission rate.	Outcome
ED Median Time from ED Arrival to ED Departure for Discharged ED Patients for Adult Patients (Process)	QACEP50	Median time from emergency department arrival to time of departure from the emergency room for patients discharged from the emergency department.	Process

The MIPS quality measures listed above are selected based on their clinical proximity to the Emergency Medicine measure such as assessing quality actions related to a similar patient cohort, and the number of clinicians with both cost and quality measures.

6.2 Harmonization

During the measure's development, the Clinician Expert Workgroup specifically considered how to align relevant cost and quality measures (e.g., episode window length). One such example included shortening the episode window to 30 days to allow for quality measure incentives to match up with the cost measure incentives.

6.3 Competing Measures

There are no measures that conceptually address both the same measure focus and the same target population as the Emergency Medicine measure.

Additional Information

Emergency Medicine Clinician Expert Workgroup Members:
As noted above, the following members provided detailed feedback on the measure specifications throughout its development based on public comments, clinical expertise, and empirical analyses.

Brandon Lewis, DO, MBA, FACOEP, FACEP, US Acute Care Solutions
Carleen Jogodka, PT, DPT, OCS, FAAOMPT, Carondelet Health Network
Carolyn Fruci, MD, PrimaCARE and St. Anne's Hospital
Dipali Ruby Sahoo, DO, MBA, FACP, SFHM, TeamHealth
John Lam, MD, MBA, Southern California Permanente Medical Group
Joshua Liao, MD, MSc, American College of Physicians
Michelle Lin, MD, MPH, MS, FACEP, Icahn School of Medicine at Mount Sinai
Mustafa Mark Hamed, MD, MPH, MBA, McKenzie Health System
Nabil Khoury, MD, Emergency Physician Medical Group
Nathan Ruch, MD, Sound Physicians
Nicholas Mohr, MD, MS, University of IOWA Carver College of Medicine
Patricia Bartzak, DNP, RN, CMSRN, TCRN, Lahey Hospital and Medical Center
Paula Tucker, DNP, FNP-BC, ENP-C, FAANP, Emory University School of Nursing
Rajeev Suri, MD, MBA, UT Health San Antonio
Sarah Eakin, MD, Pathology Associates of Erie
Stephen Epstein, MD, MPP, FACEP, Beth Israel Deaconess Medical Center
Susan Nedza, MD, MBA, SMN Health Policy Insights
Tyler Hill, DO, FACEP, Memorial Health System

Measure Developer Updates and Ongoing Maintenance

The measure isn't currently in use, but the earliest possible release of the measure in MIPS would be CY2025. If the measure becomes finalized for use in MIPS, it would undergo annual maintenance and a comprehensive re-evaluation every 3 years. This measure has been submitted to the 2022 Measures Under Consideration (MUC) List and may be reviewed by the MAP in winter of 2022. There are no further updates or reviews for this measure scheduled at this time.

Appendix A: TIN-NPI Level Results for SRF Risk Adjustment

This table presents the coefficient of patient-level dual status under three different models at the TIN-NPI level to help determine whether it's appropriate to risk adjust for SRFs.

Table A1: Coefficient of Patient-level Dual Status under Different Models at TIN-NPI Level

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit with Observation Care Or Ending in IP Stay: Abdominal Pain, Nausea, And Vomiting	2.09%	5.96 (p: 0.79)	6.12 (p: 0.79)	-18.81 (p: 0.48)
ED Visit with Observation Care Or Ending in IP Stay: Altered Mental State	0.05%	-71.97 (p: 0.51)	-103.35 (p: 0.36)	-279.86 (p: 0.34)
ED Visit with Observation Care Or Ending in IP Stay: Behavioral Health	1.53%	160.93 (p: <0.01)	80.79 (p: 0.04)	107.4 (p: 0.01)
ED Visit with Observation Care Or Ending in IP Stay: Cancer	0.85%	-380.69 (p: <0.01)	-343.1 (p: <0.01)	-374.41 (p: <0.01)
ED Visit with Observation Care Or Ending in IP Stay: Diabetes	0.47%	41.22 (p: 0.17)	22.64 (p: 0.46)	0.05 (p: 1)
ED Visit with Observation Care Or Ending in IP Stay: ENT and Eye Disorders	0.19%	-219.58 (p: <0.01)	-214.32 (p: <0.01)	-58.96 (SD: p: 0.63)
ED Visit with Observation Care Or Ending in IP Stay: Fracture	0.16%	-171.46 (p: <0.01)	-154.58 (p: 0.01)	-240.85 (p: 0.03)
ED Visit with Observation Care Or Ending in IP Stay: Gastrointestinal or Liver Conditions	5.21%	79.81 (SD: 13.34, p: 0)	64.7 (SD: 13.75, p: 0)	32.85 (p: 0.02)
ED Visit with Observation Care Or Ending in IP Stay: General Infection	1.29%	266.89 (p: 0.01)	140.91 (p: 0.16)	201.27 (p: 0.08)
ED Visit with Observation Care Or Ending in IP Stay: Gynecological Disorders	0.04%	-11.95 (p: 0.9)	-13.47 (p: 0.89)	446.84 (p: 0.21)
ED Visit with Observation Care Or Ending in IP Stay: Healthcare Maintenance	0.01%	70.73 (p: 0.42)	41.52 (p: 0.65)	545.35 (p: 0.43)
ED Visit with Observation Care Or Ending in IP Stay: Hematologic And Immunologic	1.47%	-14.14 (p: 0.86)	-26.3 (p: 0.75)	-28.43 (p: 0.77)
ED Visit with Observation Care Or Ending in IP Stay: Kidney And Urinary	3.76%	18.74 (p: 0.1)	18 (p: 0.12)	-6.35 (p: 0.61)
ED Visit with Observation Care Or Ending in IP Stay: Neurologic	1.47%	44.73 (p: 0.15)	6.31 (p: 0.84)	31.84 (p: 0.37)

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit with Observation Care Or Ending in IP Stay: Non-Diabetic Endocrine	0.01%	17.63 (p: 0.9)	23.92 (p: 0.87)	-1633.25 (p: 0.24)
ED Visit with Observation Care Or Ending in IP Stay: Non-Fracture Musculoskeletal	1.27%	39.58 (p: 0.22)	1.82 (p: 0.96)	-37.27 (p: 0.35)
ED Visit with Observation Care Or Ending in IP Stay: Non-Respiratory Chest Pain	1.74%	120.34 (p: <0.01)	68.41 (p: 0.02)	5.22 (p: 0.87)
ED Visit with Observation Care Or Ending in IP Stay: Other Cardiovascular	5.92%	16.44 (p: 0.29)	4.75 (p: 0.77)	-36.87 (p: 0.03)
ED Visit with Observation Care Or Ending in IP Stay: Peripheral Vasc	0.74%	49.14 (p: 0.26)	20.68 (p: 0.64)	51.74 (p: 0.36)
ED Visit with Observation Care Or Ending in IP Stay: Poisoning	0.01%	-717.7 (p: 0.02)	-720.58 (p: 0.02)	-173.88 (p: 0.92)
ED Visit with Observation Care Or Ending in IP Stay: Pregnancy	0.01%	-667.81 (p: 0.05)	-633.61 (p: 0.06)	-475.93 (p: 0.55)
ED Visit with Observation Care Or Ending in IP Stay: Respiratory	4.69%	81.9 (p: <0.01)	53.26 (p: <0.01)	35.99 (p: 0.03)
ED Visit with Observation Care Or Ending in IP Stay: Sepsis	3.18%	32.72 (p: 0.07)	-17.83 (p: 0.34)	-77.04 (p: <0.01)
ED Visit with Observation Care Or Ending in IP Stay: Skin Conditions, Rashes, and Abscesses	0.28%	-22.8 (p: 0.71)	-52.88 (p: 0.4)	-179.4 (p: 0.08)
ED Visit with Observation Care Or Ending in IP Stay: Stroke	1.6%	18.26 (p: 0.42)	-21.11 (p: 0.36)	-30.99 (p: 0.23)
ED Visit with Observation Care Or Ending in IP Stay: Syncope	0.38%	-22.98 (p: 0.52)	9.14 (p: 0.81)	-59.49 (p: 0.26)
ED Visit with Observation Care Or Ending in IP Stay: Trauma: Major Or Head	1.21%	90.4 (p: 0.01)	22.43 (p: 0.51)	-24.55 (p: 0.54)
ED Visit with Observation Care Or Ending in IP Stay: Trauma: Minor Or Unclear Severity	0.18%	-284.75 (p: 0.13)	-336.09 (p: 0.09)	-322.43 (p: 0.19)
ED Visit without Observation Care Nor Ending in IP Stay: Abdominal Pain, Nausea, And Vomiting	5.36%	-81.45 (p: <0.01)	-66.35 (p: <0.01)	-75.42 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Altered Mental State	0.01%	485.46 (p: 0.49)	646.89 (p: 0.38)	-1453.25 (p: 0.58)

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit without Observation Care Nor Ending in IP Stay: Behavioral Health	1.99%	-79.39 (p: <0.01)	-104.13 (p: <0.01)	-57.5 (p: 0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Cancer	0.39%	-481.17 (p: <0.01)	-516.42 (p: <0.01)	-626.36 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Diabetes	0.75%	9.76 (p: 0.62)	34.65 (p: 0.09)	63.12 (p: 0.01)
ED Visit without Observation Care Nor Ending in IP Stay: ENT and Eye Disorders	1.79%	-30.38 (p: <0.01)	-19.28 (p: 0.06)	-21.68 (p: 0.06)
ED Visit without Observation Care Nor Ending in IP Stay: Fracture	1.23%	-371.3 (p: <0.01)	-302.14 (p: <0.01)	-285.9 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Gastrointestinal or Liver Conditions	3.44%	0.86 (p: 0.94)	13.54 (p: 0.28)	1.7 (p: 0.9)
ED Visit without Observation Care Nor Ending in IP Stay: General Infection	1.21%	2.35 (p: 0.92)	14.75 (p: 0.52)	2.4 (p: 0.93)
ED Visit without Observation Care Nor Ending in IP Stay: Gynecological Disorders	0.15%	-32.71 (p: 0.25)	-22.65 (p: 0.44)	-16.4 (p: 0.79)
ED Visit without Observation Care Nor Ending in IP Stay: Healthcare Maintenance	0.11%	-7.95 (p: 0.69)	-7.02 (p: 0.73)	43.84 (p: 0.13)
ED Visit without Observation Care Nor Ending in IP Stay: Hematologic And Immunologic	1.61%	-173.48 (p: <0.01)	-158.24 (p: <0.01)	-198.75 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Kidney And Urinary	6.55%	-95.31 (p: <0.01)	-79.06 (p: <0.01)	-83.54 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Neurologic	2.32%	-92.18 (p: <0.01)	-92.34 (p: <0.01)	-68.35 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Non-Diabetic Endocrine	0.1%	-106.68 (p: 0.08)	-71.69 (p: 0.25)	-77.69 (p: 0.49)
ED Visit without Observation Care Nor Ending in IP Stay: Non-Fracture Musculoskeletal	5.11%	-134.95 (p: <0.01)	-100.62 (p: <0.01)	-118.48 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Non-Respiratory Chest Pain	2.32%	-116.52 (p: <0.01)	-101.75 (p: <0.01)	-118.23 (p: <0.01)

Subgroup Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status		
		Base Model + Patient-level Dual Status	Base Model + Patient-level Dual Status + Clinician's Dual Share	Base Model + Patient-level Dual Status + Clinician's Fixed Effect
ED Visit without Observation Care Nor Ending in IP Stay: Other Cardiovascular	7.28%	-98.88 (p: <0.01)	-77.3 (p: <0.01)	-86.89 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Peripheral Vasc	0.74%	-4.44 (p: 0.9)	8.15 (p: 0.82)	2.25 (p: 0.96)
ED Visit without Observation Care Nor Ending in IP Stay: Poisoning	0.23%	-4.83 (p: 0.87)	1.94 (p: 0.95)	88.26 (p: 0.12)
ED Visit without Observation Care Nor Ending in IP Stay: Pregnancy	0.02%	75.49 (p: 0.49)	-251.84 (p: 0.02)	-273.83 (p: 0.57)
ED Visit without Observation Care Nor Ending in IP Stay: Respiratory	5.73%	-68.31 (p: <0.01)	-70.37 (p: <0.01)	-78.35 (p: <0.01)
ED Visit without Observation Care Nor Ending in IP Stay: Sepsis	0.22%	267.61 (p: 0.08)	444.72 (p: <0.01)	356.27 (p: 0.15)
ED Visit without Observation Care Nor Ending in IP Stay: Skin Conditions, Rashes, and Abscesses	1.96%	-42.22 (p: <0.01)	-23.77 (p: 0.02)	-26.68 (p: 0.02)
ED Visit without Observation Care Nor Ending in IP Stay: Stroke	0.35%	112.94 (p: 0.06)	130.59 (p: 0.04)	144.5 (p: 0.11)
ED Visit without Observation Care Nor Ending in IP Stay: Syncope	1.53%	16.8 (p: 0.33)	22.48 (p: 0.21)	9 (p: 0.67)
ED Visit without Observation Care Nor Ending in IP Stay: Trauma: Major Or Head	3.36%	-18.11 (p: 0.19)	-9.03 (p: 0.52)	-12.18 (p: 0.43)
ED Visit without Observation Care Nor Ending in IP Stay: Trauma: Minor Or Unclear Severity	4.35%	17.86 (p: 0.05)	28.59 (p: <0.01)	15.64 (p: 0.1)