

Parkinson's Syndromes, Multiple Sclerosis (MS), and Amyotrophic Lateral Sclerosis (ALS)

Measure Justification Form

July 2024



Table of Contents

1.0	Introduction	4
1.1	Project Title	4
1.2	Date	4
1.3	Project Overview	4
1.4	Measure Name	4
1.5	Type of Measure	4
1.6	Measure Description	4
2.0	Importance	5
2.1	Evidence to Support the Measure Focus	5
2.1.1	Logic Model	7
2.2	Performance Gap	7
2.2.1	Rationale	7
2.2.2	Performance Scores	7
2.2.3	Disparities	8
3.0	Scientific Acceptability	9
3.1	Data Sample Description	9
3.1.1	Type of Data Used for Testing	9
3.1.2	Specific Dataset Used for Testing	9
3.1.3	Dates of the Data Used in Testing	9
3.1.4	Levels of Analysis Tested	9
3.1.5	Entities Included in the Testing and Analysis	9
3.1.6	Patient Cohort Included in the Testing and Analysis	10
3.1.7	Social Risk Factors Included in Analysis	11
3.2	Reliability Testing	12
3.2.1	Level of Reliability Testing	12
3.2.2	Method of Reliability Testing	12
3.2.3	Statistical Results from Reliability Testing	13
3.2.4	Interpretation	13
3.3	Validity Testing	14
3.3.1	Level of Validity Testing	14
3.3.2	Method of Validity Testing	14
3.3.3	Statistical Results from Validity Testing	15
3.3.4	Interpretation	16
3.4	Exclusions Analysis	17
3.4.1	Method of Testing Exclusions	17
3.4.2	Statistical Results from Testing Exclusions	17
3.4.3	Interpretation	18
3.5	Risk Adjustment or Stratification	19
3.5.1	Method of Controlling for Differences	19
3.5.2	Conceptual, Clinical, and Statistical Methods	20
3.5.3	Conceptual Model of Impact of Social Risks	20
3.5.4	Statistical Results	21
3.5.5	Analyses and Interpretation in Selection of Social Risk Factors	21
3.5.6	Method for Statistical Model or Stratification Development	24
3.5.7	Statistical Risk Model Discrimination Statistics	24
3.5.8	Statistical Risk Model Calibration Statistics	24
3.5.9	Statistical Risk Model Calibration – Risk Decile	24
3.5.10	Interpretation	25
3.6	Identification of Meaningful Differences in Performance	25
3.6.1	Method	25
3.6.2	Statistical Results	25
3.6.3	Interpretation	25
3.7	Missing Data Analysis and Minimizing Bias	26

3.7.1	Method	26
3.7.2	Missing Data Analysis.....	26
3.7.3	Interpretation	26
4.0	Feasibility	28
4.1	Data Elements Generated as Byproduct of Care Processes	28
4.2	Electronic Sources	28
4.3	Data Collection Strategy.....	28
4.3.1	Data Collection Strategy Difficulties	28
5.0	Usability and Use	29
5.1	Use	29
5.1.1	Current and Planned Use	29
5.1.2	Feedback on the Measure by Those being Measured or Others.....	29
5.2	Usability	33
5.2.1	Improvement	33
5.2.2	Unexpected Findings.....	33
5.2.3	Unexpected Benefits.....	33
6.0	Related and Competing Measures	34
6.1	Relation to Other Measures.....	34
6.2	Harmonization	35
6.3	Competing Measures	35
	Additional Information.....	36

1.0 Introduction

This Measure Justification Form (MJF) provides results for the testing and evaluation of the Parkinson's Syndromes, Multiple Sclerosis (MS), and Amyotrophic Lateral Sclerosis (ALS) measure. The form is intended to provide detailed information about the testing conducted on this measure, and accompanies the Measure Information Form¹ and Measure Codes List² file, which together, comprise the specifications for this cost measure.

1.1 Project Title

Physician Cost Measure and Patient Relationship Codes

1.2 Date

Information included is current on July 29, 2024.

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 Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) requirements. The contract name is "Physician Cost Measure and Patient Relationship Codes (PCMP)." The contract number is 75FCMC18D0015, Task Order 75FCMC19F0004.

1.4 Measure Name

Parkinson's Syndromes, Multiple Sclerosis (MS), and Amyotrophic Lateral Sclerosis (ALS)
Episode-Based Cost Measure

1.5 Type of Measure

Cost/Resource Use

1.6 Measure Description

The Parkinson's Syndromes, MS, and ALS episode-based cost measure evaluates a clinician's or clinician group's risk-adjusted and specialty-adjusted cost to Medicare for patients who receive medical care to manage and treat Parkinson's and related conditions, MS, or ALS. This chronic condition measure includes the costs of services that are clinically related to the attributed clinician's role in managing care during a Parkinson's Syndromes, MS, or ALS episode.

¹CMS, "Parkinson's Syndromes, Multiple Sclerosis (MS), and Amyotrophic Lateral Sclerosis (ALS)" Measure Methodology," *QPP Cost Measure Information Page*, <https://www.cms.gov/medicare/quality/value-based-programs/cost-measures>

²CMS, "Parkinson's Syndromes, Multiple Sclerosis (MS), and Amyotrophic Lateral Sclerosis (ALS)" Measure Codes List" *QPP Cost Measure Information Page*, <https://www.cms.gov/medicare/quality/value-based-programs/cost-measures>

2.0 Importance

2.1 Evidence to Support the Measure Focus

The Parkinson's Syndromes, MS, and ALS 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 the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). MIPS aims to reward high-value care by measuring clinician performance through four areas: quality, improvement activities, Promoting Interoperability, and cost. Each category assesses different aspects of care, and the categories are weighted to combine into one composite score. CMS introduced MIPS Value Pathways (MVPs) 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 patient costs.

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 their influence. This measure provides clinicians with information about their care costs that they can use to understand the costs associated with their decision-making. Clinicians play an important role in variation in health care expenditures due to their ability to affect costs.³ A cost measure offers an 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 quality through changes in clinical practice.

According to the literature and feedback received through stakeholder input activities, this measure's focus represents an area with opportunities for improvement. As discussed in the rest of this section, primary opportunities for improving Parkinson's syndromes, MS, and ALS cost outcomes include:

- I. Improving physical activity and fall-related education and treatment
- II. Screening patients for additional comorbidities not related to physical complications
- III. Mitigating drug interactions or use of inappropriate medications

In a survey of Parkinson's patients at 10 years of the disease, 39.8% indicated they were not exercising.⁴ Increased activity improves both physical health and mental acuity in both Parkinson's and MS patients⁵ and, in fact, meaningfully improves fall-related outcomes in Parkinson's patients.⁶ Educating patients on the benefits of exercise and/or appropriate physical activity is thus imperative to the improvement of fall-related outcomes and reducing any costs of subsequent hospitalizations.

³ 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>.

⁴ da Silva, Franciele Cascaes et al. "Effects of physical exercise programs on cognitive function in Parkinson's disease patients: A systematic review of randomized controlled trials of the last 10 years." *PloS one* vol. 13,2 e0193113. 27 Feb. 2018, doi:10.1371/journal.pone.0193113

⁵ Döring, Andrea et al. "Exercise in multiple sclerosis -- an integral component of disease management." *The EPMA journal* vol. 3,1 2. 24 Dec. 2011, doi:10.1007/s13167-011-0136-4

⁶ Shen, Xia et al. "Effects of Exercise on Falls, Balance, and Gait Ability in Parkinson's Disease: A Meta-analysis." *Neurorehabilitation and neural repair* vol. 30,6 (2016): 512-27. doi:10.1177/1545968315613447

Beyond experiencing physical constraints, between 43 to 70% of MS patients report cognitive impairment, which requires regular assessment to detect.⁷ Other studies have also found that clinically significant depressive disturbances affect 40 to 50% of Parkinson's patients, whereas only 36.9% of applicable providers completed a comprehensive annual review of psychiatric disorders.^{8,9} As such, the screening of patients for both cognitive impairment and mental/behavioral health intervention represents a relevant opportunity to improve their quality of life. Studies focused on ALS have provided evidence these screenings can take place in both multidisciplinary and specialized clinics.^{10,11}

The simultaneous use of multiple drugs is very prevalent amongst elderly people with various comorbidities and those with severe chronic diseases, such as MS patients, making them more prone to suffer from potential drug-drug interactions (pDDIs).¹² Drug interactions can affect the efficacy of one or more medications and lead to treatment failure and/or serious side effects.¹³ Additionally, for patients with Parkinson's Disease, contra-indicated dopamine blocking agents are often used as antipsychotics, which can cause severe adverse drug reactions and worsen Parkinson's-related motor symptoms.^{14,15}

⁷ Langdon, D W et al. "Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS)." *Multiple sclerosis (Houndmills, Basingstoke, England)* vol. 18,6 (2012): 891-8. doi:10.1177/1352458511431076

⁸ Reijnders, Jennifer S A M et al. "A systematic review of prevalence studies of depression in Parkinson's disease." *Movement disorders : official journal of the Movement Disorder Society* vol. 23,2 (2008): 183-9; quiz 313. doi:10.1002/mds.21803

⁹ Baek, William S et al. "Quality care assessment of Parkinson's disease at a tertiary medical center." *The International journal of neuroscience* vol. 123,4 (2013): 221-5. doi:10.3109/00207454.2012.751024

¹⁰ Woolley, Susan C et al. "Detecting frontotemporal dysfunction in ALS: utility of the ALS Cognitive Behavioral Screen (ALS-CBS)." *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases* vol. 11,3 (2010): 303-11. doi:10.3109/17482961003727954

¹¹ Gordon, Paul H et al. "A screening assessment of cognitive impairment in patients with ALS." *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases* vol. 8,6 (2007): 362-5. doi:10.1080/17482960701500817

¹² Bachmann, Paula et al. 2022. "Prevalence and Severity of Potential Drug–Drug Interactions in Patients with Multiple Sclerosis with and without Polypharmacy" *Pharmaceutics* 14, no. 3: 592. <https://doi.org/10.3390/pharmaceutics14030592>

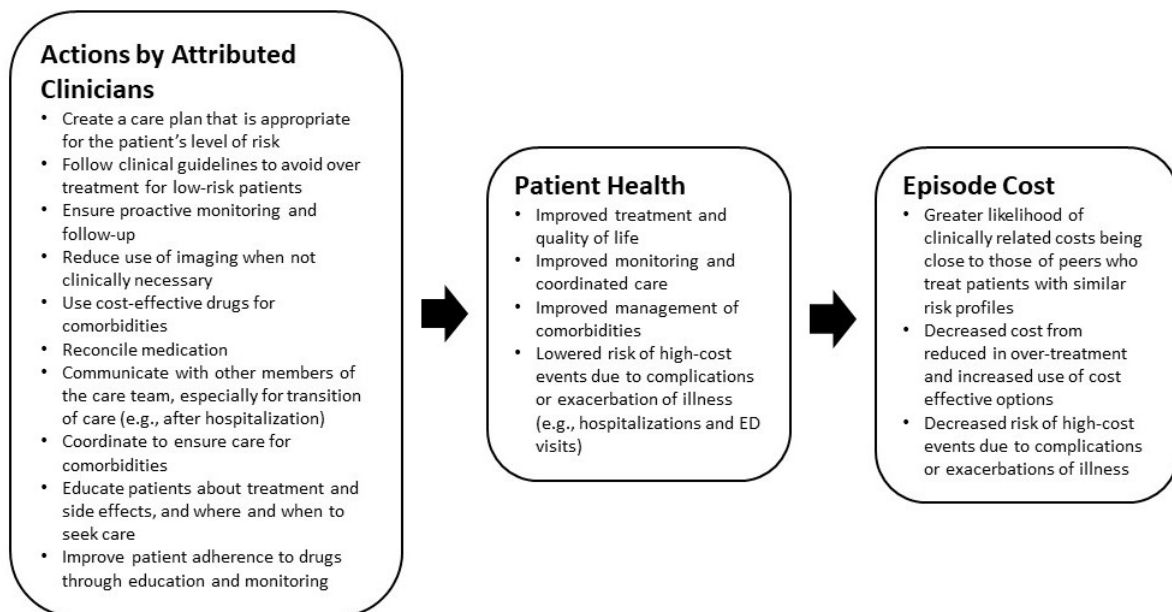
¹³ Ibid.

¹⁴ Lertxundi, Unax et al. "Adverse reactions to antipsychotics in Parkinson disease: an analysis of the Spanish pharmacovigilance database." *Clinical neuropharmacology* vol. 38,3 (2015): 69-84. doi:10.1097/WNF.0000000000000080

¹⁵ Weintraub, Daniel et al. "Patterns and trends in antipsychotic prescribing for Parkinson disease psychosis." *Archives of neurology* vol. 68,7 (2011): 899-904. doi:10.1001/archneurol.2011.139

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

Given the impact of Parkinson's Syndromes, MS, and ALS on the older adult population, the high costs to Medicare for managing the condition and its complications, and the performance gaps identified in the literature, a cost measure represents an opportunity for improving overall cost performance.

The Parkinson's Syndromes, MS, and ALS episode-based cost measure was recommended for development because of its high impact in terms of patient population, clinician coverage, and Medicare spending, and the opportunity to build a chronic condition measure that would address a condition not captured by other episode-based cost measures in the MIPS cost performance category. A measure-specific Clinician Expert Workgroup was then convened with clinicians, health care experts, and patient representatives who have appropriate experience to provide extensive, detailed input on this measure throughout its development.

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). Substantial variation is observed in the measure, indicated by the interquartile range, standard deviation, and coefficient of variation. The 90th percentile score is more than double the 10th percentile at the TIN and TIN-NPI levels. These results highlight an opportunity for improvement by closing the gap between the most and least efficient providers.

Table 1. Distribution of the Measure Score

Metric	TIN	TIN-NPI
Count	2,930	2,930
Mean Score	\$14,646	\$14,425
Score Standard Deviation	\$4,130	\$4,801
Minimum Score	\$2,014	\$2,014
Maximum Score	\$39,314	\$40,684
Score Interquartile Range (IQR)	\$4,658	\$5,731
Score Percentile		
10 th	\$9,897	\$9,168
20 th	\$11,390	\$10,615
30 th	\$12,533	\$11,686
40 th	\$13,463	\$12,669
50 th	\$14,276	\$13,840
60 th	\$15,149	\$14,907
70 th	\$16,124	\$16,143
80 th	\$17,389	\$17,790
90 th	\$19,665	\$20,537

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 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 data from the Common Working File (CWF), Long-Term Care Minimum Data Set (LTC MDS), and Common Medicare Environment (CME).

3.1.2 Specific Dataset Used for Testing

The Parkinson's Syndromes, MS, and ALS measure uses Medicare Part A and Part B claims and Part D prescription drug event data maintained by CMS. Part A, B, and D claims data are used to build episodes of care, calculate episode costs, and construct risk adjusters. 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 health care delivery choices. Data from the EDB are used to determine beneficiary-level exclusions and secondary risk adjusters, 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

Parkinson's Syndromes, MS, and ALS episodes ending from January 1, 2023, through December 31, 2023.

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 demographics of individual clinicians (identified by combination of TIN and NPI) and clinician groups (identified by TIN) included in the testing of the Parkinson's Syndromes, MS, and ALS measure.

Table 2: Measured Entities Demographics

Metric	TIN		TIN-NPI	
	Count	%	Count	%
Count	2,930	100%	2,930	100%
Number of Episodes Attributed	-	-	-	-
20-39 Episodes	1,262	43.07%	1,774	60.55%
40-59 Episodes	472	16.11%	563	19.22%
60-79 Episodes	267	9.11%	262	8.94%
80-99 Episodes	170	5.80%	138	4.71%
100-199 Episodes	405	13.82%	176	6.01%
200-299 Episodes	145	4.95%	13	0.44%
300+ Episodes	209	7.13%	4	0.14%
Census Region	-	-	-	-

Metric	TIN		TIN-NPI	
	Count	%	Count	%
Northeast	610	20.82%	602	20.55%
Midwest	594	20.27%	577	19.69%
South	1,149	39.22%	1,113	37.99%
West	575	19.62%	636	21.71%
Unknown	2	0.07%	2	0.07%

Table 3 shows the top 10 attributed specialties for the Parkinson's Syndromes, MS, and ALS measure at the 20-episode testing volume threshold. The most frequently attributed specialties reflect the intent of the measure to capture costs of the management of Parkinson's Syndromes, MS, and ALS, including neurologists, nurse practitioners, and physician assistants. These clinicians are also consistent with input provided by stakeholders, including patient and family partners (PFPs), during the measure development process. PFPs identified neurologists, physical/occupational therapists, and psychiatrists, amongst others, as being part of their care team.

Table 3: Count of the Top 10 Attributed Specialties

Specialty	Number of TIN-NPIs Attributed
Neurology	2,456
Nurse Practitioner	229
Physician Assistant	105
Internal Medicine	49
Physical Medicine and Rehabilitation	24
Family Practice	21
Psychiatry	12
Neuropsychiatry	8
Geriatric Medicine	5
General Practice	5

3.1.6 Patient Cohort Included in the Testing and Analysis

Table 4 shows the patient population for the Parkinson's Syndromes, MS, and ALS measure testing. It consists of Medicare beneficiaries enrolled in Medicare Parts A and B who are receiving care for the management and treatment of Parkinson's syndromes, MS, or ALS that triggers a Parkinson's Syndromes, MS, and ALS episode and do not meet the measure's exclusion criteria, as outlined in 3.4.1.

Table 4: Beneficiary Demographics

Metric	Value
Count	283,806
Mean Age	73.90
Female %	50.66%
Part D Enrollment %	78.45%

3.1.7 Social Risk Factors Included in Analysis

The analysis of social risk factors (SRFs) focused on examining the impact of Dual Medicare and Medicaid enrollment status on the measure. Table 5 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.¹⁶

Table 5: Social Risk Factors Available for Analysis

Variable	Advantages	Disadvantages	Used in Testing
Dual Medicare and Medicaid enrollment status	<ul style="list-style-type: none"> Available for all beneficiaries Most powerful predictor of poor outcomes¹⁶ 	<ul style="list-style-type: none"> Variation in Medicaid eligibility across states 	Yes
Race/Ethnicity	<ul style="list-style-type: none"> Available for most beneficiaries, except for ambiguous categories of “Unknown” or “Other” 	<ul style="list-style-type: none"> Social risk driven by someone’s race is often correlated with and partially captured by dual status¹⁶ Only 5 categories available, which may lack granularity to fully capture disparities^{17,18} 	No
ICD-10 Z codes for social determinants of health	<ul style="list-style-type: none"> Reflects individual-level factors that influence health status and contact with health services 	<ul style="list-style-type: none"> Not routinely and consistently coded on claims, only available for 0.1% of all fee-for-service claims in 2019¹⁹ 	No
American Community Survey	<ul style="list-style-type: none"> 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., AHRQ index, deprivation index) 	<ul style="list-style-type: none"> Only a proxy measure, not always accurate at individual-level 	No

¹⁶ 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>

¹⁷ 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).

¹⁸ 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).

¹⁹ Centers for Medicare and 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>

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, group/practice (TIN) and individual clinician (TIN-NPI) levels.

3.2.2 Method of Reliability Testing

Data Element Reliability

The Parkinson's Syndromes, MS, and ALS 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 necessary data fields used in this measure, including diagnosis and procedure codes and other elements consequential to payment. Specifically, CMS works with Zone Program Integrity Contractors, 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 also uses the Comprehensive Error Rate Testing (CERT) Program to ensure that Medicare payments are correct under 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 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 how much of the variation in the measure score is explained by differences among clinician performance (i.e., signal) rather than random variation (i.e., statistical noise) among 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 2022, CMS Comprehensive Error Rate Testing (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.²⁰ The fiscal year 2023 Medicare fee-for-service program proper payment rate was 92.62%.²¹

Clinician-level Reliability

The table below shows reliability metrics at the 20-episode testing volume thresholds. While higher thresholds generally yield higher reliability results, these increases must be considered against decreasing the number of clinicians and clinician groups eligible for the measure, which would limit the applicability of measures to larger group practices and potentially limit the impact of the measure in encouraging performance improvement. For testing purposes, we used a 20-episode volume threshold. If the measure is implemented in MIPS in the future, CMS will establish a case minimum through notice-and-comment rulemaking.

Table 6: Reliability at the Accountability Entity Level

Reporting Level	Entities Meeting Case Minimum	Mean Reliability	Median Reliability	% Above 0.4	% Above 0.7
TIN	2,930	0.611	0.628	82.70%	38.87%
TIN-NPI	2,930	0.571	0.581	78.94%	29.28%

3.2.4 Interpretation

The results of the data element testing show high reliability of the critical data elements used by the measure. Based on existing scientific evidence on the different interpretations and methods of estimating reliability, CMS finalized in the CY 2022 Physician Fee Schedule (86 FR 64996) rule that the 0.4 threshold for mean reliability continues to be appropriate for indicating moderate reliability for performance measures in the Cost category in the MIPS program. Mean reliability levels above 0.7 continue to demonstrate high reliability for cost measures, as previously established in the CY 2017 Quality Payment Program final rule (81 FR 77169 through 77171).²² At the entity level, the Parkinson's Syndromes, MS, and ALS measure is moderately reliable at the TIN and TIN-NPI reporting levels, at 0.61 and 0.57, respectively. Additionally, the overwhelming majority of TINs and TIN-NPIs meet or exceed the MIPS reliability threshold of 0.4 at 82.70% and 78.94%, respectively.

²⁰Comprehensive 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>.

²¹ Fiscal Year 2023 Agency Final Report, Department of Health and Human Services Agency Financial Report <https://www.hhs.gov/sites/default/files/fy-2023-hhs-agency-financial-report.pdf>

²² 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](https://www.federalregister.gov/documents/2021/12/16/2021-24686).

3.3 Validity Testing

3.3.1 Level of Validity Testing

The validity of the measure was tested using empirical validity at the accountable entity level (TIN and TIN-NPI).

3.3.2 Method of Validity Testing

Face Validity

The Parkinson's Syndromes, MS, and ALS measure was developed through a structured, iterative process for gathering detailed input on the measure from recognized clinician experts. 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 between good from poor performance).

In developing this measure, Acumen incorporated input from:

- (i) a Parkinson's Syndromes, MS, and ALS 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 [QPP Cost Measure Information Page](#).²³

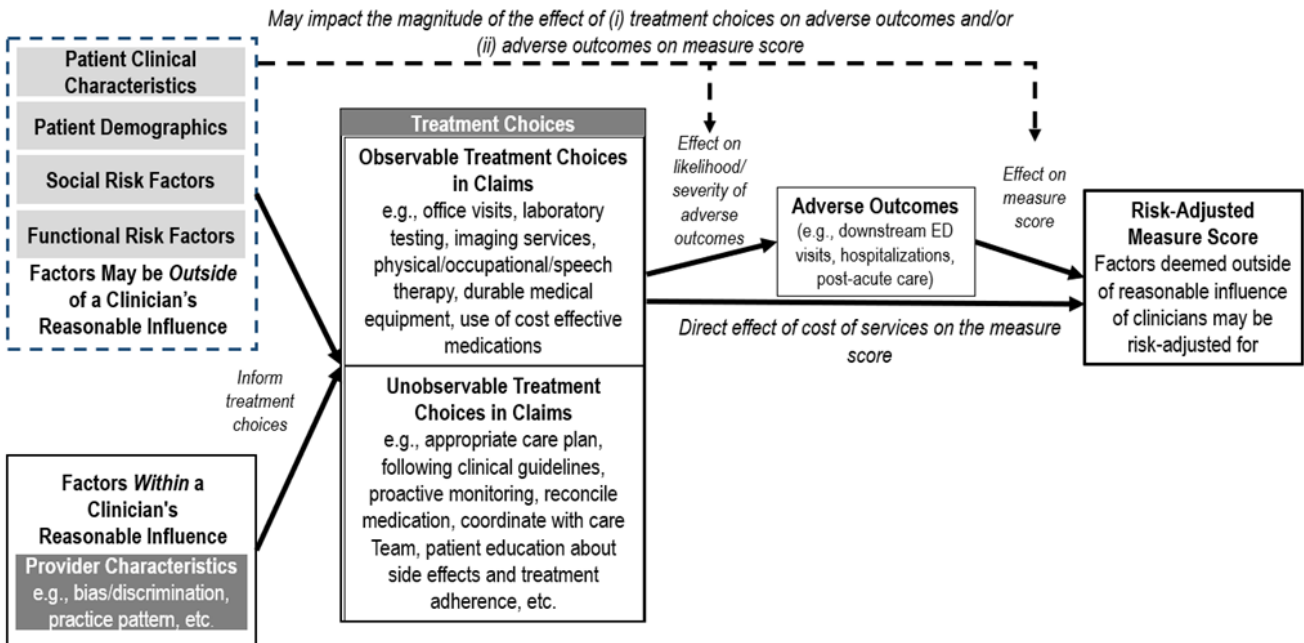
One of the primary roles of the Clinician Expert Workgroup is to develop service assignment rules for the cost measure. These service assignment rules seek to ensure clinicians are evaluated on services and costs that are clinically related to the attributed clinician's role in treating and managing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

Empirical Validity Testing

Validity is a criterion used to assess whether the cost measure can quantify the construct it aims to measure, which is the cost directly related to treatment choices and the cost of adverse outcomes resulting from care. We evaluated the empirical validity of the Parkinson's Syndromes, MS, and ALS 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.

²³ CMS, QPP Cost Measure Information Page, <https://www.cms.gov/medicare/quality/value-based-programs/cost-measures>.

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



The cost measure is designed to reflect costs directly related to treatment choices, and the cost of adverse outcomes resulting from care. Therefore, treatment choices, either observable in claims or otherwise, by an attributed clinician can directly impact the measure score or indirectly when they are mediated through the cost of adverse outcomes. In turn, the cost of adverse effects is related to the total cost captured by the measure score.

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

Empirical Validity Testing

Table 7 shows two regression models for each reporting level. Model 1 shows the effect on the clinicians' mean observed cost to expected cost ratio for each additional one thousand dollar of a cost 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 dollar of a cost category that is assigned to an episode, on average, while holding the remaining categories of cost constant.

Table 7: Estimated Effect on Treatment Choices on the Measure Score

Service Categories	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 Events	0.04 [0.04,0.05] (p < 0.01)	-	0.06 [0.05,0.06] (p < 0.01)	-
Outpatient Evaluation & Management Services	-0.05 [-0.07,- 0.03] (p < 0.01)	4.34 [4.13,4.54] (p < 0.01)	-0.03 [-0.07,0.00] (p = 0.07)	3.87 [3.61,4.12] (p < 0.01)
Major Procedures	0.34 [0.25,0.44] (p < 0.01)	-1.38 [-2.40,- 0.35] (p < 0.01)	0.14 [0.06,0.22] (p < 0.01)	-0.14 [-0.79,0.50] (p = 0.67)
Outpatient Physical, Occupational, or Speech and Language Pathology Therapy	0.06 [0.05,0.08] (p < 0.01)	0.08 [-0.09,0.26] (p = 0.35)	0.09 [0.07,0.11] (p < 0.01)	-0.26 [-0.40,- 0.11] (p < 0.01)
Laboratory, Pathology, and Other Tests	0.14 [-0.02,0.30] (p = 0.08)	-5.17 [-6.90,- 3.44] (p < 0.01)	0.32 [0.18,0.46] (p < 0.01)	-3.55 [-4.69,- 2.42] (p < 0.01)
Imaging Services	0.25 [0.11,0.39] (p < 0.01)	-4.34 [-5.87,- 2.82] (p < 0.01)	-0.02 [-0.12,0.09] (p = 0.75)	-0.92 [-1.76,- 0.08] (p = 0.03)
Durable Medical Equipment and Supplies	0.00 [-0.01,0.00] (p = 0.59)	0.20 [0.15,0.25] (p < 0.01)	-0.01 [-0.02,0.01] (p = 0.33)	0.28 [0.19,0.38] (p < 0.01)
Chemotherapy and Other Part B-Covered Drugs	0.03 [0.02,0.03] (p < 0.01)	-0.04 [-0.06,- 0.01] (p = 0.02)	0.03 [0.03,0.03] (p < 0.01)	-0.02 [-0.04,0.00] (p = 0.01)
Part-D Drugs	0.02 [0.02,0.02] (p < 0.01)	0.02 [-0.02,0.06] (p = 0.30)	0.02 [0.02,0.02] (p < 0.01)	0.03 [0.01,0.06] (p = 0.01)

3.3.4 Interpretation

The testing results in Table 7 demonstrate that the Parkinson's Syndromes, MS, and ALS measure reflects the cost directly related to treatment choices and the cost of related adverse outcomes. Therefore, there is evidence that the measure captures what it purports to measure.

Model 1 shows that the cost of adverse events is associated with a worse measure score. Physical/occupational/speech pathology therapy, major procedures, imaging services, laboratory tests, and Part B medications are all associated with worse measure scores (Model 1) but also with lower costs of adverse events at the TIN-NPI level (Model 2), suggesting that

these services are beneficial to patient outcomes but may be prone to overuse. Costs of outpatient evaluation and management (E/M) services and durable medical equipment are marginally associated with better measure scores at the TIN level (Model 1), which suggests that they are beneficial to the overall episode spending. Finally, Part D costs are associated with a worse measure score at both reporting levels but do not have a statistically significant association with the cost of adverse events, which could indicate the potential to reduce medication costs without increasing the occurrence of adverse events.

3.4 Exclusions Analysis

3.4.1 Method of Testing Exclusions

Exclusions are used in the Parkinson's Syndromes, MS, and ALS measure to ensure a comparable patient population within the scope of the measure's focus on the management and treatment of Parkinson's syndromes, MS, and ALS 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. These are standard exclusions applied to chronic condition episode-based cost measures. Other exclusions are due to outlier data or providers not meeting a minimum amount of cases for measurement (20 episodes).

- Episodes where patient 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 does not capture the full length of care intended by the measure.
- Episode that is less than one year in length
 - These episodes were excluded as they are not sufficiently long to indicate an ongoing care relationship for a chronic condition.

Given the rationales for these exclusions, we 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 the measure calculation to assess the distinctness between the two patient cohorts. A full list of the exclusions used for the Parkinson's Syndromes, MS, and ALS measure is provided in the Measure Codes List available on the [QPP Cost Measure Information Page](https://www.cms.gov/medicare/quality/value-based-programs/cost-measures).²⁴

3.4.2 Statistical Results from Testing Exclusions

Table 8 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.

²⁴CMS, QPP Cost Measure Information Page, <https://www.cms.gov/medicare/quality/value-based-programs/cost-measures>.

Table 8: Cost Statistics for Measure Exclusions

Exclusion	Episodes		Mean	Observed Cost Percentile				
	#	% of All Episodes Meeting Triggering Logic		10 th	25 th	50 th	75 th	90 th
All Episodes Meeting Triggering Logic	467,336	100.00%	\$16,373	\$727	\$1,712	\$5,987	\$19,637	\$48,381
Episode Length Less Than One Attribution Window	27,944	5.98%	\$37,580	\$1,759	\$5,042	\$16,223	\$44,276	\$95,421
Beneficiary Death in Episode	69,794	14.93%	\$29,039	\$1,565	\$4,788	\$14,755	\$35,386	\$69,424
Outlier	7,896	1.69%	\$45,465	\$4,522	\$15,209	\$51,845	\$61,050	\$93,812
TIN Does not Meet Case Minimum	102,494	21.93%	\$17,624	\$717	\$1,815	\$6,911	\$21,479	\$50,025
No Attributed NPI	56,873	12.17%	\$18,891	\$940	\$2,363	\$7,919	\$23,796	\$53,446
TIN-NPI Does not Meet Case Minimum	261,424	55.94%	\$16,976	\$691	\$1,695	\$6,273	\$20,484	\$49,179
Microvascular Decompression	13	0.00%	\$14,617	\$832	\$1,081	\$4,430	\$14,002	\$50,299
Spinal Cord Injury	12	0.00%	\$7,502	\$673	\$1,561	\$3,272	\$9,706	\$15,481
Stereotactic Radiosurgery	64	0.01%	\$20,610	\$1,204	\$2,730	\$7,776	\$29,140	\$66,829
Reportable Episodes - Group Reporting	305,938	65.46%	\$13,310	\$672	\$1,465	\$4,715	\$15,341	\$40,084
Reportable Episodes - Individual Reporting	131,723	28.19%	\$12,426	\$689	\$1,417	\$4,244	\$13,572	\$37,411

3.4.3 Interpretation

Table 8 displays descriptive statistics of all episodes meeting the measure's triggering logic, excluded episodes, and the final reportable episodes at the group- and individual level. The statistical results show that the exclusion criteria decrease the distribution of episode costs for reportable episodes. Costs of all episodes meeting the triggering logic range from \$727 at the 10th percentile to \$48,381 at the 90th percentile. After exclusions are applied, group reporting episode costs vary from \$672 at the 10th percentile to \$40,084 at the 90th percentile, whereas individual reporting ranges from \$689 to \$37,411. .

Most excluded episodes have higher mean observed costs than the episodes meeting the triggering logic, with the largest exclusions owing to removing episodes with no attributed clinician and applying the 20-episode testing volume threshold to ensure a sufficient sample size for the measure. Episodes where a beneficiary died before the episode end date are excluded because they do not provide sufficient data in the episode window period. These episodes also have a higher mean observed cost than all episodes meeting triggering logic, at \$29,039. Episodes classified as outlier cases have a mean observed episode cost of \$45,465 compared to \$16,373 for all episodes meeting triggering logic. The wide variability of observed episode costs for outlier cases also supports their exclusion. At the 10th percentile the outlier cases observed cost is \$4,522 and at the 90th percentile the observed cost is \$93,812. Based on testing results and input from the Parkinson's Syndromes, MS, and ALS Clinician Expert Workgroup, episodes with spinal cord injury, stereotactic radiosurgery, and microvascular

decompression are excluded given the small sample size of each variable and their impacts on the risk adjustment model.

Therefore, without substantially changing the composition of attributed episodes, excluding episodes in these categories will ensure a comparable and clinically coherent patient cohort that will yield a clinically coherent measure and meaningful information to attributed clinicians.

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 122 risk factors and stratification by 6 risk categories.

The risk adjustment model for the Parkinson's Syndromes, MS, and ALS measure adjusts for comorbidities based on the CMS Hierarchical Condition Category (HCC) model, count of HCCs, end-stage renal disease (ESRD) status, disability status, number and types of clinician specialties from which the patient has received care, recent use of institutional long-term care, age, and dual eligibility status.

The model also includes measure-specific factors:

- Frailty
- Wheelchair dependence
- History of falling
- Difficulty swallowing
- Cognitive status impairment, decline, or deficit
- Deep brain stimulation during episode
- Intrathecal pump during episode
- Dysphonia
- Dysarthria and anarthria
- Other degenerative diseases of basal ganglia
- Past contracture diagnoses
- Sleep apnea
- Bowel or bladder incontinence
- Dependence on respirator

A separate linear regression is run for episodes with and without Medicare Part D enrollment status combination to ensure fair comparison:

- Parkinson's and Related Conditions
- Multiple Sclerosis
- Amyotrophic Lateral Sclerosis

The episode's scaled (i.e., annualized) observed costs are winsorized at the 98th percentile prior to the regression for each model to handle extreme observations. Full details of the risk adjustment model are in the Measure Codes List File available on the [QPP Cost Measure Information Page](#).²⁵

²⁵CMS, QPP Cost Measure Information Page, <https://www.cms.gov/medicare/quality/value-based-programs/cost-measures>.

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 Parkinson's Syndromes, MS, and ALS measure's patient population.

The workgroup provided input on measure-specific risk adjusters 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 adjusters are listed in the section above.

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 3 shows the conceptual model that outlines how SRFs can influence the measure score, which is informed by published external research and Acumen's data analysis.^{16,26,27,28,29} 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 the influence of the attributed clinician. Risk factors, including SRFs, can influence the treatment choices and impact the size of the effect of treatment choices on mitigating the risk and cost of adverse outcomes.

A systematic approach then guides the decision of which factors to include in the risk adjustment model:

1. 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 adjusters are commonly the HCCs.
2. 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 patient characteristics. We arrived at the final list of risk adjusters based on those discussions and consensus among the clinical experts.
3. During our testing phases, we also follow a structured and systematic approach to deciding whether SRFs should be adjusted for, further described in Section 3.5.5.

²⁶Assistant 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.

²⁷Chen 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

²⁸Medicare 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/>.

²⁹Office 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-medicare-value-based-purchasing-programs>

3.5.4 Statistical Results

The literature has extensively tested using the HCC model for Medicare claims data. Although the variables in the HCC model were selected to predict annual cost, CMS has also used this risk adjustment model in several 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 V24 model can be found in the Evaluation of the CMS-HCC Risk-Adjustment Model report³⁰ and the Report to Congress: Risk Adjustment in Medicare Advantage³¹. For measure-specific factors not included in the CMS-HCC model, we sought expert clinician input through the workgroup, which provided recommendations on additional risk adjusters and sub-groups.

3.5.5 Analyses and Interpretation in Selection of Social Risk Factors

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

- (i) whether there is 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 does not risk adjust for SRFs.

³⁰Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

³¹CMS, "Report to Congress: Risk Adjustment in Medicare Advantage," <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf>.

Table 9: Coefficient of Patient-level Dual Status under Different Models

Level	Sub-Group Risk Model	% of All Episodes	Coefficient of Patient-level Dual Status (P-value)		
			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
TIN	Parkinson's and Related Conditions without Part D Enrollment	16.38%	0.19 (p: 0.00)	0.13 (p: 0.04)	0.22 (p: 0.00)
TIN	Parkinson's and Related Conditions with Part D Enrollment	59.37%	0.14 (p <.0001)	0.11 (p <.0001)	0.10 (p <.0001)
TIN	Multiple Sclerosis without Part D Enrollment	4.48%	0.13 (p: 0.25)	0.10 (p: 0.39)	0.15 (p: 0.32)
TIN	Multiple Sclerosis with Part D Enrollment	18.68%	0.47 (p <.0001)	0.47 (p <.0001)	0.47 (p <.0001)
TIN	Amyotrophic Lateral Sclerosis without Part D Enrollment	0.24%	-0.01 (p: 0.97)	-0.01 (p: 0.98)	-0.55 (p: 0.33)
TIN	Amyotrophic Lateral Sclerosis with Part D Enrollment	0.85%	-0.09 (p: 0.24)	-0.07 (p: 0.40)	-0.08 (p: 0.49)
TIN-NPI	Parkinson's and Related Conditions without Part D Enrollment	16.41%	0.17 (p: 0.01)	0.12 (p: 0.08)	0.21 (p: 0.04)
TIN-NPI	Parkinson's and Related Conditions with Part D Enrollment	59.78%	0.15 (p <.0001)	0.16 (p <.0001)	0.14 (p <.0001)
TIN-NPI	Multiple Sclerosis without Part D Enrollment	4.39%	0.09 (p: 0.47)	-0.04 (p: 0.78)	0.09 (p: 0.74)
TIN-NPI	Multiple Sclerosis with Part D Enrollment	18.40%	0.49 (p <.0001)	0.50 (p <.0001)	0.53 (p <.0001)
TIN-NPI	Amyotrophic Lateral Sclerosis without Part D Enrollment ³²	0.22%	N/A	N/A	N/A
TIN-NPI	Amyotrophic Lateral Sclerosis with Part D Enrollment	0.80%	-0.06 (p: 0.53)	0.07 (p: 0.51)	-0.02 (p: 0.92)

³² No Dual Status episodes were identified for this subgroup.

Table 10: 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 Episodes	Dual Episodes	Non-Dual Episodes	All Episodes	Dual Episodes	Non-Dual Episodes
All	1.09	1.16	1.09	1.08	1.27	1.06
0%-20%	1.05	1.26	1.04	1.02	1.32	1.02
21%-40%	1.08	1.15	1.07	1.04	1.32	1.03
41%-60%	1.08	1.20	1.06	1.03	1.21	1.02
61%-80%	1.11	1.15	1.10	1.10	1.26	1.07
81%-100%	1.13	1.09	1.16	1.20	1.27	1.16

Table 11: Proportions of Clinicians Who Perform Significantly Worse, Equally Well, or Significantly Better on Their Dual Episodes than Non-Dual Episodes

Reporting Level	Significantly Worse	Equally Well	Significantly Better
TIN	7.06%	90.69%	2.25%
TIN-NPI	8.58%	91.19%	0.23%

Table 12: Clinicians' Performance Shift after Adding a Dual Status Risk Adjustor

TIN or TIN-NPI	Proportion of Clinicians Affected at Various Levels of Performance Shift	
	Ranking Shift by 1% or more	Ranking Shift by 5% or more
TIN	75.46%	11.64%
TIN-NPI	69.17%	10.00%

The results suggest that it is appropriate to risk adjust for social risk factors in this measure. Table 9 shows there is a statistically significant association between the patient's dual status and episode cost for both TINs and TIN-NPIs in the largest sub-groups (i.e., Parkinson's and Related Conditions with Part D Enrollment and Multiple Sclerosis with Part D Enrollment). This association persists after adding variables to account for clinician-level factors, which suggests that the patient-level factors are more influential than clinician-level factors. For episodes without Part D enrollment, this association is not statistically significant across all models. Still, episodes with Part D enrollment are relatively more predominant at both reporting levels, which, in combination with the results in Table 9, suggests that it is appropriate to risk adjust for patient characteristics and that these are more influential than provider characteristics.

Further, Table 10 demonstrates that clinicians and group tend to perform worse on episodes with dual enrollment status, regardless of the share of episodes with dual enrollment status (Table 10). While many clinicians are able to perform equally well on their dual episodes and non-dual episodes, there are still a substantial number of clinicians performing significantly worse on their dual episodes than their non-dual episodes, which suggests that clinicians aren't able to fully mitigate the effect of SRFs (Table 11). Lastly, risk adjusting for dual status appears

to change the performance ranking for a subset of clinicians, with ten percent or more having a shift of five percent or more (Table 12).

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 Parkinson's Syndromes, MS, and ALS cost measure, calculated by dividing explained sum of squares by total sum of squares is 0.228. The adjusted R-squared is also 0.226. More information on discrimination testing for the CMS-HCC model can be found at Pope et al. 2011.³³

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 moderate variation among risk deciles, as predictive ratios range from 0.81 to 1.20 across all risk deciles (with an overall average of 1.00). All deciles are at least within 0.2 of 1.00.

Table 13: Predictive Ratio by Decile of Predicted Episode Cost

Decile	Average Predictive Ratio
Decile 1	0.81
Decile 2	0.86
Decile 3	0.96
Decile 4	0.96
Decile 5	0.95
Decile 6	0.92
Decile 7	0.88
Decile 8	0.88
Decile 9	0.96
Decile 10	1.20

³³Pope, Gregory C., John Kautter, et al., "Evaluation of the CMS-HCC Risk-Adjustment Model: Final Report." RTI International: March 2011.

3.5.10 Interpretation

The R-squared values for the model, which measure the percentage of variation in results predicted by the model, are higher than the values presented in similar analyses of risk adjustment models.³⁴ 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 are not 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 is optional because the measure should only adjust for some variations in the cost of care. In collaboration with the experts from our clinical workgroup, this measure only adjusts for factors that are deemed outside the reasonable influence of clinicians. The service assignment rules provide context for which costs are included in the measure and which are not.

Table 13 shows that the risk adjustment model is moderately consistent, with the average predictive ratios observed to be close to 1.00 across all deciles, with the range between 0.81 to 1.20. Overall, the risk adjustment model does not 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 adverse 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. There is a difference in mean score for TIN and TIN-NPI levels because each level has its own attribution rules, which resulted in slightly different populations of episodes used for measure score calculation (Table 2). However, clinicians are only compared to their peers at either the TIN or TIN-NPI level, therefore the differences in score across different levels can be ignored.

Episodes with certain clinical services or events have higher risk-adjusted episode costs compared to the average observed cost for all episodes (\$14,565). These include readmissions (\$37,769), inpatient rehabilitation/long-term care hospital stays (\$56,513), and skilled nursing facility services (e.g., post-acute care) (\$28,064).

3.6.3 Interpretation

There is substantial variation 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.

Since each episode with readmissions, rehabilitation, and/or post-acute care is very costly, every percentage reduction in such services represents substantial performance improvement for the attributed clinician or clinician group.

³⁴Pope, 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.

3.7 Missing Data Analysis and Minimizing Bias

3.7.1 Method

Since CMS uses Medicare claims data to calculate the Parkinson's Syndromes, MS, and ALS measure, Acumen expects a high degree of data completeness. To further ensure that we have complete and accurate data for each patient, Acumen excludes episodes where patient date of birth information (an input to the risk adjustment model) cannot be found in the EDB, the patient does not appear in the EDB, or the patient death date occurs before the episode trigger date.

The Parkinson's Syndromes, MS, and ALS 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

The table below presents the frequency of missing data across the categories of missing data which caused episodes to be excluded from the Parkinson's Syndromes, MS, and ALS measure. Frequency is presented in terms of the number of episodes excluded due to missing data, as well as the cost profile of episodes with missing data compared to episodes included in the measure reporting.

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

Table 14: Cost Statistics for Missing Data Category

Missing Data Categories	Episodes	Observed Cost					
		Mean	Percentile				
	#		10 th	25 th	50 th	75 th	90 th
All Episodes	687,973	\$16,608	\$695	\$1,695	\$5,925	\$19,543	\$48,629
Beneficiary Resides Outside of U.S. or Territories	219	\$8,426	\$183	\$457	\$1,108	\$6,644	\$25,433
Primary Payer Other than Medicare	49,928	\$16,134	\$487	\$1,209	\$4,427	\$17,875	\$48,903
No Continuous Enrollment in Medicare Parts A and B, and Any Enrollment in Part C	73,628	\$14,155	\$401	\$1,022	\$3,599	\$14,401	\$43,951

3.7.3 Interpretation

The results show that the missing data episodes, including episodes where individuals have a primary payer other than Medicare and where individuals lack continuous enrollment in Medicare Parts A and B but are enrolled in Part C, don't appear to be substantially different than all episodes in the initial population in terms of cost (Table 14). Medicare beneficiaries who live outside the U.S. or its territories have substantially smaller mean episode cost compared to all episodes in the initial population, however, the episode count is small.

It is appropriate to remove episodes in these categories as they are likely indicators of a discontinuation of the patient-clinician relationship or an absence of Medicare usage, and therefore do not provide sufficient data during the episode window. Furthermore, 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, 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 CWF maintained at the CMS Baltimore Data Center. Healthcare providers submit Medicare claims to a Medicare Administrative Contractor (MAC), which 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 such, it is not practical to wait until all claims for a given month are finalized before calculating the measure, resulting in a trade-off between efficiency (accessing the data on time) and accuracy (waiting until most claims are finalized) when determining the duration (i.e., the “claims run-out” period) after which to pull claims data. To determine the appropriate claims run-out period, Acumen has tested 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 CMS adopts this measure for use in a program, calculation and reporting would align with the program’s reporting practices.

4.3.1.2 Missing Data

This measure requires complete beneficiary information, therefore, a small number of episodes with missing data are excluded to ensure data completeness and accurate comparability across episodes. For example, episodes where the beneficiary was not enrolled in Medicare Parts A and B for the 120 days before the episode start date are excluded from this measure. Excluding these episodes 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 cannot be located are excluded from the measure.

4.3.1.3 Sampling

During measure testing, Acumen noted that episodes in which the beneficiary died before the episode end date exhibited different cost distributions than other episodes. As such, this measure excludes episodes to avoid negatively impacting clinician scores.

5.0 Usability and Use

5.1 Use

5.1.1 Current and Planned Use

The measure is not currently in use but is intended for use in a payment program and could eventually be publicly reported. It 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 CMS to approve this measure for use in MIPS, it must be reviewed by the Pre-Rulemaking Measure Review process (PRMR) and then undergo the notice-and-rulemaking process. Given these next steps, the earliest the measure could be used in MIPS is CY 2026. 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 Parkinson's Syndromes, MS, and ALS measure development, we used an iterative and extensive process to gather feedback on the measure and its results to ensure that it can be used appropriately in the MIPS program by clinicians and clinician groups who practice in this clinical area. This process also seeks to ensure that the measured entities can understand and interpret their performance results to help support decision-making. A couple of the main ways we gathered input was through reoccurring Clinician Expert Workgroup meetings, which incorporated feedback from the patient and caregiver perspective, empirical data, and discussion between clinician experts who recommend measure specifications, and through the national field testing of the measures.

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) to inform the Clinician Expert Workgroup members' recommendations. These analyses were conducted using all administrative claims data for Medicare Parts A, B, and D. This data was shared with Workgroup members to help inform their feedback on the measure specifications throughout its development to ensure that the measure is appropriately assessing costs for these clinicians.

Field Testing

Additionally, Acumen and CMS nationally field tested the draft Parkinson's Syndromes, MS, and ALS measure, along with 1 other episode-based cost measure, for a 6-week comment period (February 1 to March 14, 2024). We provided a Field Test Report with performance data to all clinician groups and clinicians who were attributed 20 or more episodes, which was the testing volume threshold.³⁵ 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 stakeholders as possible. A total of 5,947 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.

³⁵The 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 before or during each of the Clinician Expert Workgroup Meetings: the Workgroup Webinar, Service Assignment and Refinement Webinar, and Post-Field Test Refinement Webinar. 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 appropriately assessed 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, the measured entities (i.e., MIPS-eligible clinicians and clinician groups who received a report) and the general public provided feedback on the appropriateness of the measures and the usability of the data. The public comments were summarized in a report, which was shared with the Clinician Expert Workgroup for consideration when 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.

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 care services)³⁶
 - 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 stakeholders, 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,

³⁶CMS, "Restructured BETOS Classification System <https://data.cms.gov/provider-summary-by-type-of-service/provider-service-classifications/restructured-betos-classification-system>

and a Measure Testing Form (including reliability and validity data).³⁷ During field testing, Acumen conducted education and outreach activities for interested parties, including multiple office hours sessions with specialty societies, a publicly posted field testing webinar recording, and Quality Payment Program Help Desk support.

Education and Outreach

Acumen directly conducted outreach via email to tens of thousands of interested parties using a contact list developed through previous public engagement efforts, as well as CMS and Quality Payment Program (QPP) listservs. Acumen also emailed clinicians who received the field test reports via CMS's GovDelivery.

Acumen and CMS hosted two office hours sessions in February 2024 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 attendees from targeted specialty societies who are likely to have members who could be attributed the measure.

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

Acumen and CMS hosted the public 2024 MACRA Cost Measures Field Testing webinar in February 2024, where interested parties could learn more about field testing and the measures.³⁸ The webinar presentation outlined: (i) the cost measure field testing project (ii) the measure development and re-evaluation processes, and (iii) field testing activities. There was also an opportunity to ask questions during the Q&A portion of the webinar. The webinar recording, slides, and transcript were then made available for the public to review.

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 58 survey responses and 7 comment letters, including from specialty societies representing large numbers of potentially attributed clinicians and from persons with lived experience.

Survey responses and comment letters were collected via two online surveys, 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.³⁹ The

³⁷The measure specifications, mock reports, Measure Development Process document, Frequently Asked Questions document, and testing documents are posted on the Cost Measures Information Page:

<https://www.cms.gov/medicare/quality/value-based-programs/cost-measures>.

³⁸MACRA Wave 6 Cost Measures Field Testing Webinar materials are available on the Quality Payment Program Webinar Library: <https://qpp.cms.gov/about/webinars>.

³⁹CMS, "2024 Field Testing Feedback Summary Report," Cost Measures Information Page <https://www.cms.gov/medicare/quality/value-based-programs/cost-measures/current>.

measure-specific feedback was used as the basis for the post-field testing refinements that 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 Parkinson's Syndromes, MS, and ALS measure.

5.1.2.5 Feedback from Other Users

Person and Family Engagement

Acumen incorporated thoughtful input from patients and caregivers throughout the Parkinson's Syndromes, MS, and ALS measure development process. Before each Clinician 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 at webinars 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 from PFPs include the types of services that they typically received and what helped to improve their care (e.g., physical therapy, medication management, durable medical equipment) and noted the types of clinicians that contributed to their care team (e.g., neurologists, physical/occupational therapists). They also highlighted areas of concerns, such as complications and lack of care coordination that impacted the quality of their care.

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 field testing, Acumen compiled the feedback provided through the surveys 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 Parkinson's Syndromes, MS, and ALS measure made after consideration of field-testing analyses and stakeholder feedback are:

- Service Assignment
 - Added service codes for hospitalizations for metabolic nutritional status, functional performance caregiver training, and remote therapy
 - Removed craniotomy
- Exclusions
 - Added exclusion variables: spinal cord injury patients, stereotactic radiosurgery, microvascular decompression
- Risk Adjustment
 - Added risk adjustor variables: deep brain stimulation and intrathecal pump procedures occurring during the episode
- Measure Name
 - Change from "Movement Disorders" to "Parkinson's Syndromes, Multiple Sclerosis (MS), and Amyotrophic Lateral Sclerosis (ALS)"

5.2 Usability

5.2.1 Improvement

The measure has not yet been implemented, and as such has not had influence over performance. Our testing suggests that there is 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 assesses 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 has not been implemented at this time, so we do not have data that confirm 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 medications from Part B or D may be costly, they are not a major driver of the measure score, therefore, demonstrating the robustness of the risk adjustment model and the ability of the cost measure to differentiate performance that is most relevant to the treatment and management of patients with Parkinson's syndromes, MS, or ALS.

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 has not 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 have 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 Supportive Care for Neurodegenerative Conditions MVP includes quality measures that align with the measure's intent. Additionally, the Falls: Plan of Care quality measure aligns with the Parkinson's Syndromes, MS, and ALS measure's opportunities for improvements.

Table 15: MIPS Quality Measures Potentially Relevant for the Parkinson's Syndromes, MS, and ALS Episode Group

Measure Title	Measure ID	Measure Description	Measure Type
Assessment of Mood Disorders and Psychosis for Patients with Parkinson's Disease	497	Percentage of all patients with a diagnosis of Parkinson's Disease [PD] who were assessed for depression, anxiety, apathy, AND psychosis once during the measurement period.	Process
Assessment of Cognitive Impairment or Dysfunction for Patients with Parkinson's Disease	496	Percentage of all patients with a diagnosis of Parkinson's Disease [PD] who were assessed for cognitive impairment or dysfunction once during the measurement period.	Process
Rehabilitative Therapy Referral for Patients with Parkinson's Disease	498	Percentage of all patients with a diagnosis of Parkinson's Disease who were referred to physical, occupational, speech, or recreational therapy once during the measurement period.	Process
Amyotrophic Lateral Sclerosis (ALS) Patient Care Preferences	53	Percentage of patients diagnosed with Amyotrophic Lateral Sclerosis (ALS) who were offered assistance in planning for end of life issues (e.g., advance directives, invasive ventilation, hospice) at least once annually.	Process
Advance Care Plan	37	Percentage of patients aged 65 years and older who have an advance care plan or surrogate decision maker documented in the medical record or documentation in the medical record that an advance care plan was discussed but the patient did not wish or was not able to name a surrogate decision maker or provide an advance care plan.	Process
Use of High-Risk Medications in Older Adults	744	Percentage of patients 65 years of age and older who were ordered at least two high-risk medications from the same drug class.	Process
Screening for Social Drivers of Health	1664	Percent of patients 18 years and older screened for food insecurity, housing instability, transportation needs, utility difficulties, and interpersonal safety.	Process
Falls: Plan of Care	255	Percentage of patients aged 65 years and older with a history of falls that had a plan of care for falls documented within 12 months.	Process

The MIPS quality measures listed above are related to the Parkinson's Syndromes, MS, and ALS measure by assessing clinicians on the employment of certain processes in their care of patients with Parkinson's syndromes, MS, or ALS. As such, these quality measures (listed in Table 15 above) may include metrics that are focused on a similar patient cohort, or that are clinically related to the care provided for the episode group.

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). The Parkinson's Syndromes, MS, and ALS measure has the potential to be used in the Supportive Care for Neurodegenerative Conditions MVP. MVPs offer a participation framework meant to align cost and quality measures providing a degree of standardization to hold clinicians accountable for their clinical decisions in a consistent manner. MVPs also seek to connect measures with improvement activities to the relevant area of clinical practice. While there are no improvement activities in MIPS that are specific to the Parkinson's syndromes, MS, or ALS clinical area, there is an improvement activity related to chronic care—Chronic Care and Preventative Care Management for Empaneled Patients (IA_PM_13)—which may correlate with the Parkinson's Syndromes, MS, and ALS measure as it aims to improve outcomes for patients that have chronic conditions or diseases and care transition.

6.3 Competing Measures

There are no measures that conceptually address both the same measure focus and the same target population as the Parkinson's Syndromes, MS, and ALS measure.

Additional Information

Parkinson's Syndromes, MS, and ALS 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.

Deena Hassaballa, DO, FAAPMR, American Academy of Physical Medicine and Rehabilitation
Dheeraj Mahajan, MD, MBA, MPH, FACP, CIC, CMD, CHCQM, The Society for Post-Acute and Long-Term Care Medicine
Kathleen McCoy, DNSc, PMHNP-BC, PMHCNS-BC, FNP-BC, FAANP, American Association of Nurse Practitioners
Marisa McGinley, DO, MsC, American Academy of Neurology
Kelsey Peterson, OTD, OTR/L, Neuro-IFRAH Certified, American Occupational Therapy Association
Alexander Rae-Grant, MD, FRCPC, FAAN, American Academy of Neurology
Miriam Rafferty, PT, DPT, PhD, APTA Board Certified Neurologic Clinical Specialist, American Physical Therapy Association
Patricia Scheets, PT, DPT, American Physical Therapy Association
David Schultz, MD, American Academy of Family Physicians
Jason Schwalb, MD, American Association of Neurological Surgeons
David Seidenwurm, MD, American College of Radiology
Binit Shah, MD, American Academy of Neurology
Chloe Slocum, MD, MPH, American Academy of Physical Medicine and Rehabilitation
Laura Verdun, CCC-SLP, American Speech-Language-Hearing Association
Christine Williamitis, PhD, DNP, PMHNP, ACNP, FNP, American Academy of Nurse Practitioners

Measure Developer Updates and Ongoing Maintenance

The measure is not currently in use, but the earliest possible release of the measure in MIPS would be CY2026. If the measure becomes finalized for use in MIPS, it would undergo annual maintenance and a comprehensive re-evaluation every 3 years. This measure is included on the 2024 Measures Under Consideration (MUC) List and will be reviewed by PRMR in winter of 2024-2025. There are no further updates or reviews for this measure scheduled at this time.