

Environmental Scan on Developing and Implementing Performance Measures for Population-Based Total Cost of Care (PB-TCOC) Models

March 19, 2024

This environmental scan was prepared at the request of the Office of the Assistant Secretary for Planning and Evaluation (ASPE) as background information to assist the Physician-Focused Payment Model Technical Advisory Committee (PTAC) in preparing for a theme-based discussion on developing and implementing performance measures for population-based total cost of care (PB-TCOC) models. Topics that are addressed in this environmental scan include types of performance measures used in value-based payment models and pay-for-reporting programs; data sources used for constructing performance measures; features of PB-TCOC models; desired challenges related to developing and implementing performance measures; opportunities for Alternative Payment Models (APMs) and PB-TCOC models to address challenges in developing and implementing performance measures; trends in existing performance measures across several Center for Medicare and Medicaid Innovation (CMMI) models, Medicare value-based care programs, and submitted PTAC proposals; how performance measures are being linked to payment; risk adjustment methods used for PB-TCOC models; best practices for measuring performance; and opportunities for improving performance measures.ⁱ

ⁱ This analysis was prepared under contract #HHSP233201500048IHHS75P00123F37023 between the Department of Health and Human Services' Office of Health Policy of the Assistant Secretary for Planning and Evaluation (ASPE) and NORC at the University of Chicago. The opinions and views expressed in this analysis are those of the authors. They do not reflect the views of the Department of Health and Human Services, the contractor, or any other funding organizations. This analysis was completed on March 19, 2024.

Table of Contents

I. Introduction and Purpose	11
II. Key Highlights	12
II.A. Definitions	12
II.B. Key Findings	13
III. Research Approach	21
III.A. Research Questions	21
III.B. Research Methods	24
IV. Background.....	25
IV.A. Defining Performance Measures.....	25
IV.B. Measure Stewards, Measure Developers, and the Measure Lifecycle	29
IV.C. Goals of Performance Measurement	30
V. Challenges Related to Developing and Implementing Performance Measures.....	31
V.A. Technical Issues Related to Measuring Quality.....	31
V.B. Overlap Between Current Measures/Measure Definitions	34
V.C. Challenges in Developing New Performance Measures and Obtaining Endorsement of Performance Measures	35
V.D. Challenges in Implementing New Performance Measures.....	36
VI. Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures	37
VI.A. Discussion of Potential Opportunities.....	37
VI.B. Examples of Successful Models	39
VII. Trends in Existing Performance Measures.....	39
VII.A. High-Level Findings and Summary Statistics for Performance Measures in 24 Select Models or Programs.....	40
VII.B. Assessment of Performance Measures for Most Prevalent Chronic Conditions	55
VII.C. Measures Focused on Similar Aspects of Care	57
VIII. Performance Measures Used in CMMI Models.....	58
VIII.A. Types of Performance Measures	59
VIII.B. How Payment is Linked to Performance	61
VIII.C. Overlap in Current Performance Measures/Measure Definitions	62
VIII.D. Gaps in Current Measures.....	63
VIII.E. Lessons Learned Related to Developing and Implementing Performance Measures.....	63

IX. Performance Measures Used in the Medicare Shared Savings Program.....	63
IX.A. Types of Performance Measures.....	64
IX.B. How Payment is Linked to Performance	64
X. Performance Measures Used in the Medicare Advantage Star Ratings Program	65
X.A. Types of Performance Measures.....	65
X.B. How Payment is Linked to Performance	65
X.C. Gaps in Current Measures	66
X.D. Lessons Learned Related to Developing and Implementing Performance Measures	66
XI. Performance Measures Used in PTAC Proposals.....	67
XI.A. Types of Performance Measures.....	70
XI.B. How Payment is Linked to Performance	72
XI.C. Overlap in Current Performance Measures/Measure Definitions	75
XII. How Performance Measures are Linked with Payment in Other Programs	75
XII.A. Evidence on the Effectiveness of Performance-based Payment Programs	75
XII.C. Unintended Consequences of Pay-for-Performance Measures.....	77
XII.D Considerations for Future PBP Models	78
XIII. Attribution Methods Used for Models and Performance Measures.....	78
XIV. Assessment of the Different Types of Data Sources Used for Construction of Performance Measures.....	79
XIV.A. Key Considerations for Selecting Data Sources to Measure Outcomes.....	81
XIV.B. Best Practices/Trends in Addressing Quality Deficiencies in Data Sources	81
XV. Risk Adjustment Methods Used for PB-TCOC Models.....	82
XV.A. Alternatives to Risk Adjustment.....	83
XVI. Consumer Assessment of Healthcare Providers and Systems Survey.....	84
XVI.A. CAHPS Goals and Domains	84
XVI.B. Domains Not Captured in the CAHPS.....	88
XVI.C. What CAHPS Measures are Appropriate to Use in PB-TCOC Models	89
XVII. Best Practices for Measuring Performance.....	89
XVIII. Opportunities to Improve Performance Measures	94
XVIII.A. Areas Where New Measures May be Needed	95
XVIII.B. Opportunities to Modify Existing Measures	96
XIX. Areas Where Additional Information is Needed	96

Appendix A. Research Questions by Environmental Scan Section	98
Appendix B. Search Strategy	105
Appendix C. Definitions of Various Types of Performance Measures	115
Appendix D. Summary of Current Performance Measures for Selected CMMI Models and Medicare Payment Programs	118
Appendix E. Summary of Selected CMMI Models	120
Appendix F. Summary of Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP.....	144
Appendix G. Summary of Model and Performance Measurement Characteristics of Proposals Reviewed by PTAC as of September 2020.....	152
Appendix H. Summary of Responses from Discussion with Subject Matter Experts	185
Appendix I. Areas for Future Exploration and Research.....	189
Appendix J. Annotated Bibliography	191
Appendix K. References	252

List of Exhibits

Exhibit 1.	Types of Quality Measures.....	26
Exhibit 2.	Types of Outcome Measures	26
Exhibit 3.	Relationship Between Guiding Principles and the Types of Performance Measures for PB-TCOC Models	28
Exhibit 4.	Identifying Meaningful Performance Measures For PB-TCOC Models at Each Stage of the Patient’s Care Journey – Example: Liver Disease	28
Exhibit 5.	CMS Measure Lifecycle	29
Exhibit 6.	Number of Performance Measures by Program/Model.....	41
Exhibit 7.	Two Examples of Measures Where Programs/Models Apply Different Criteria to the Same Measure.....	42
Exhibit 8.	Top 8 Performance Measures by Number of Programs/Models.....	42
Exhibit 9.	Number of Performance Measures Used for Only One Program/Model.....	44
Exhibit 10.	CMS Definitions of the Seven Performance Measure Types	45
Exhibit 11.	Distribution of Performance Measures by Measure Type for the 24 Programs/Models...	46
Exhibit 12.	Distribution of Performance Measures by Program/Model and Measure Type.....	47
Exhibit 13.	Distribution of Performance Measures by Data Sources for the 24 Programs/Models.....	50
Exhibit 14.	Distribution of Performance Measures by Reporting Level for the 24 Programs/Models.	51
Exhibit 15.	Distribution of Performance Measures by Endorsement Status for 24 Programs/Models	52
Exhibit 16.	Distribution of the 24 Programs/Models by Relationship to Provider Payment	53
Exhibit 17.	Program/Model Type Definitions	53
Exhibit 18.	Distribution of Performance Measures Based on How the 24 Selected Programs/Models are Linked with Payment.....	55
Exhibit 19.	The Number of Measures Associated with a Specific Condition- by Program/Model	56
Exhibit 20.	The Top 15 Clinical Conditions/Subspecialties.....	57
Exhibit 21.	Performance Measure Groupings for Measures Focused on Similar Aspects of Care Across the 24 Programs/Models.....	58
Exhibit 22.	CMMI Models that Include at Least One Quality Measure and One Utilization and/or Spending Measure in Implementation and/or Monitoring including Model, Model Years, and Model Stage.....	59
Exhibit 23.	Patient-Reported Outcome (PRO) or Patient Experience of Care (PEC) Measures among CMMI Models that Include at Least One Quality Measure and One Utilization and/or Spending Measure	60
Exhibit 24.	How Payment is Adjusted for Performance among CMMI Models that Include at Least One Quality Measure and One Utilization and/or Spending Measure in Implementation and/or Monitoring.....	61
Exhibit 25.	PTAC Proposals including Proposal Name, Submitter Name, and Criterion Met	67
Exhibit 26.	How Payment is Adjusted for Performance among PTAC Proposals.....	72
Exhibit 27.	CAHPS Surveys by Domain	85
Exhibit E1.	Model Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Selected CMMI Models.....	121
Exhibit E2.	Lessons Learned Related to Current Performance Measures for Selected CMMI Models	140
Exhibit F1.	Program Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP	145

Exhibit F2.	Lessons Learned Related to Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP	151
Exhibit G1.	Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that Were Found to Meet Criterion 2 and Criterion 4	153
Exhibit G2.	Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that were Found to Meet Criterion 2 and were Found to Not Meet Criterion 4	171
Exhibit G3.	Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that were Found to Not Meet Criterion 2 and were Found to Meet Criterion 4	172
Exhibit G4.	Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that were Found to Not Meet Criterion 2 or Criterion 4, were Withdrawn, or were Determined to be Out of Scope by PTAC.....	176

List of Acronyms

AAFP	American Academy of Family Physicians
AAHPM	American Academy of Hospice and Palliative Medicine
AAN	American Academy of Neurology
ACA	Affordable Care Act
ACAAI	American College of Allergy, Asthma, and Immunology
ACEP	American College of Emergency Physicians
ACO	Accountable Care Organization
ACO REACH	Accountable Care Organization Realizing Equity, Access, and Community Health
ACP	American College of Physicians
ACS	American College of Surgeons
ADI	Area Deprivation Index
AHA	American Hospital Association
AHRQ	Agency for Healthcare Research and Quality
AI	Artificial intelligence
AI/AN	American Indian/Alaska Native
AIP	Advance Investment Payment
AMA	American Medical Association
AMI	Acute myocardial infarction
APM	Alternative Payment Model
AQC	Alternative Quality Contract
AQS	Aggregate quality score
ASC	Ambulatory Surgical Center
ASCO	American Society of Clinical Oncology
AWV	Annual wellness visit
BCBSMA	Blue Cross Blue Shield of Massachusetts
BPCI-A	Bundled Payments for Care Improvement Advanced
CABG	Coronary artery bypass graft
CAD	Coronary artery disease
CAHPS	Consumer Assessment of Healthcare Providers and Systems
CBE	Consensus-based entity
CCC	Clearwater Cardiovascular and Interventional Consultants
CDI	Clostridium difficile infection
CEC	Comprehensive End-Stage Renal Disease Care
CHIP	Children's Health Insurance Program
CHQR	Cancer Hospital Quality Reporting Program
CMF	Care management fee
CMIT	CMS Measures Inventory Tool
CMMI	Center for Medicare and Medicaid Innovation
CMS	Centers for Medicare & Medicaid Services
CNM	Certified nurse midwife
COA	Community Oncology Alliance
CoCM	Collaborative care model
COPD	Chronic obstructive pulmonary disease
CPC+	Comprehensive Primary Care Plus
CQS	Composite quality score

CTAC	Coalition to Transform Advanced Care
DHN	Digestive Health Network
DHTT	Digital health technology tools
DO	Doctor of osteopathic medicine
DOMH	Department of Health and Mental Hygiene
eCQM	electronic Clinical Quality Measure
ED	Emergency department
EHR	Electronic health record
EMR	Electronic medical record
EOM	Enhancing Oncology Model
ESRD	End-Stage Renal Disease
ETC	ESRD Treatment Choices
FFS	Fee-for-service
FHIR	Fast Healthcare Interoperability Resources
FY	Fiscal year
GAO	Government Accountability Office
GPDC	Global and Professional Direct Contracting
HAC	Hospital acquired condition
HCC	Hierarchical condition categories
HCUP	Healthcare Cost and Utilization Project
HDP	Home dialysis payment adjustment
HEDIS	Healthcare Effectiveness Data and Information Set
HESS	Health Equity Summary Score
HF	Heart failure
HHA	Home health agency
HHS	Department of Health and Human Services
HHVBP	Home Health Value-Based Purchasing
HIE	Health information exchange
HIT	Health information technology
HMH	Hackensack Meridian Health
HOS	Health Outcomes Survey
HQRP	Hospice Quality Reporting Program
HRRP	Hospital Readmissions Reduction Program
HRSN	Health-related social need
IAH	Independence at Home
IGG	Illinois Gastroenterology Group
IMPACT	Improving Medicare Post-Acute Care Transformation
IOBS	Innovative Oncology Business Solutions
IOM	Institute of Medicine
IPF	Inpatient Psychiatric Facility
IPPS	Inpatient Prospective Payment System
IQS	Initial quality score
IRF	Inpatient Rehabilitation Facility
KCC	Kidney Care Choices
LIS	Low income subsidy
LTCH	Long-term care hospital

LUGPA	Large Urology Group Practice Association
MA	Medicare Advantage
MACRA	Medicare Access and CHIP Reauthorization Act
MBC	Minnesota Birth Center
MCP	Making Care Primary
MD	Medical doctor
MDTCOC	Maryland Total Cost of Care
MedPAC	Medicare Payment Advisory Commission
MIPS	Merit-based Incentive Payment System
MPFS	Medicare Physician Fee Schedule
MSBP	Medicare spending per beneficiary
MSSP	Medicare Shared Savings Program
MTM	Medication Therapy Management
MVP	MIPS Value Pathways
NF	Nursing facility
NCQA	National Committee for Quality Assurance
NGACO	Next Generation Accountable Care Organization
NP	Nurse practitioner
NQF	National Quality Forum
OCM	Oncology Care Model
OMB	Office of Management and Budget
OP	Outpatient
OPC	Oncology payment category
OQR	Outpatient Quality Reporting
P4P	Pay-for-performance
P4R	Pay-for-reporting
PA	Physician assistant
PAC	Post-acute care
PBIP	Performance-based incentive payment
PBP	Performance-based payment
PBPM	Per beneficiary per month
PBR	Performance-based recoupment
PB-TCOC	Population-based total cost of care
PCF	Primary Care First
PCMH	Patient-centered medical home
PCP	Primary care physician
PEC	Patient experience of care
PFPM	Physician-focused payment model
PHE	Public health emergency
PIP	Performance incentive payment
PMPM	Per member per month
PPA	Performance payment adjustment
PPS	Prospective Payment System
PQM	Partnership for Quality Measurement
PRC	Personalized Recovery Care
PREM	Patient-reported experience measure
PRO	Patient-reported outcome

PROM	Patient-reported outcome measure
PROMIS	Patient-Reported Outcomes Measurement Information System
PRT	Preliminary Review Team
PSI	Patient Safety Indicator
PTAC	Physician-Focused Payment Model Technical Advisory Committee
PY	Performance year
QBP	Quality Bonus Program
QPP	Quality Payment Program
QRP	Quality Reporting Program
RFI	Request for Input
RN	Registered nurse
RPA	Renal Physicians Association
RTS	Report to the Secretary
SDOH	Social determinants of health
SES	Socioeconomic status
SME	Subject matter expert
SNF	Skilled nursing facility
SSI	Surgical site infection
THA	Total hip arthroplasty
TIN	Taxpayer identification number
TKA	Total knee arthroplasty
TPS	Total performance score
TQS	Total quality score
U.S.	United States
UNMHSC	University of New Mexico Health Sciences Center
VBP	Value-based purchasing

I. Introduction and Purpose

Under the bipartisan Medicare Access and Children’s Health Insurance Program (CHIP) Reauthorization Act (MACRA) of 2015, Congress significantly changed Medicare fee-for-service (FFS) physician payment methods. The law also specifically encouraged the development of Alternative Payment Models (APMs) known as physician-focused payment models (PFPMs) and created the Physician-Focused Payment Model Technical Advisory Committee (PTAC) to review stakeholder-submitted PFPM proposals and make comments and recommendations on them to the Secretary of Health and Human Services (HHS; “the Secretary”).

Since its inception, PTAC has received 35 proposals for PFPMs from a diverse set of physician payment stakeholders, including professional associations, health systems, academic groups, public health agencies, and individual providers.ⁱⁱ PTAC evaluates the PFPM proposals based on the extent to which they meet the Secretary’s ten regulatory criteria for PFPMs (specified in federal regulations at 42 CFR § 414.1465).

The goal of the Quality and Cost criterion is to ensure that each proposed model will “improve health care quality at no additional cost, maintain health care quality while decreasing cost, or both improve health care quality and decrease cost”. The goal of the Value over Volume criterion is to “provide incentives to practitioners to deliver high-quality health care.”

Within this context, PTAC has assessed previous submitters’ planned use of performance measures in the implementation and evaluation of their proposed PFPMs. Nearly all of the 35 proposals that were submitted to PTAC between 2016 and 2020 included information about proposed performance measures to some degree. Additionally, the Committee found that at least 16 of the proposed models met both Criterion 2 (Quality and Cost) and Criterion 4 (Value over Volume).

Given the increased emphasis on developing larger, population-based APMs that encourage accountable care relationships, PTAC has conducted several theme-based discussions between 2021 and 2023 that have examined performance measurement in relation to specific topics, such as measuring rural providers’ performance or measurement of care transition management. Relevant topics identified for investigation in this environmental scan include:

- Approaches for measuring and evaluating desired quality, cost, and value-based transformation outcomes in PB-TCOC models and identifying where additional measures may be needed;
- Understanding how the different performance measure types (e.g., spending, utilization, and quality) and data sources (e.g., claims, electronic health records, and surveys) can be used to drive care delivery transformation in PB-TCOC models;
- Best practices for linking performance measures with financial incentives in PB-TCOC models;
- Improving the integration of the patient experience, as well as health equity in performance measures; and

ⁱⁱ The 35 proposals submitted to PTAC represent an unduplicated count (i.e., proposals with multiple submissions are counted only once) of the number of proposals that have been voted and deliberated on by the Committee (28) and the number of proposals that have been withdrawn by stakeholders (7, including one proposal that was withdrawn prior to any review by the Committee).

- Improving the timeliness of sharing resulting data from measurements to inform process improvement.

This environmental scan provides PTAC members with background information and context reflecting expert perspectives on issues related to developing and implementing performance measures. The environmental scan is expected to help PTAC members review strategies in previously submitted proposals. In addition, the environmental scan can inform the Committee’s review of future proposals and future comments and recommendations that Committee members may submit to the Secretary relating to performance measurement in population-based models.

This environmental scan summarizes relevant information from proposals that had been previously submitted to PTAC, findings from relevant literature, selected Center for Medicare and Medicaid Innovation (CMMI) models, and other Centers for Medicare & Medicaid Services (CMS) programs, state models, and demonstrations.

Section II provides key highlights of the findings from the environmental scan. Section III describes the research questions and methods used in the environmental scan. Subsequent sections discuss:

- Defining performance measures, describing the measure lifecycle, and goals of performance measurement for PB-TCOC models (Section IV);
- Challenges related to developing and implementing performance measures (Section V);
- Opportunities for APMs and PB-TCOC models to address challenges related to developing and implementing performance measures (Section VI); trends in existing performance measures (Section VII); performance measures used in CMMI models (Section VIII); performance measures used in the Medicare Shared Savings Program (MSSP) (Section IX); performance measures used in the Medicare Advantage (MA) Star Ratings Program (Section X); performance measures used in PTAC proposals (Section XI); how performance measures are being linked with payment in other programs (Section XII); attribution methods used for models and its performance measures (Section XIII); assessment of the different types of data sources used for construction of performance measures (Section XIV); risk-adjustment methods used for PB-TCOC models (Section XV); summary of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Survey (Section XVI); best practices for measuring performance (Section XVII); opportunities to improve performance measures (Section XVIII); and areas where additional information is needed (Section XIX). Additionally, a list of exhibits and list of abbreviations can be found at the beginning of the environmental scan, following the table of contents.

II. Key Highlights

The following section provides important definitions and highlights key findings from this environmental scan on developing and implementing performance measures in PB-TCOC models.

II.A. Definitions

PTAC developed the following working definitions of “performance measures” and “quality measures” as they relate to PB-TCOC models:

Performance Measures.

Performance measures assess all aspects of participants' performance in models including quality, outcomes, cost, and utilization. Performance measures are intended to help participants monitor their performance and help those accountable for model oversight assess model impact (e.g., improving quality without increasing spending; or decreasing spending without reducing quality).^{1,2,3,4,5,6}

Quality Measures:

Quality measures are a type of performance measure that assesses the safety, timeliness, effectiveness, efficiency, equity, and patient-centeredness of models. Quality measures may capture structures, processes, and patient experiences with health care.^{7,8,9,10}

II.B. Key Findings

Challenges Related to Developing and Implementing Performance Measures

There are several challenges associated with the development and implementation of performance measures. Developing new performance measures can be a lengthy process that requires substantial resources.¹¹ First, implementing performance measures (i.e., measuring and reporting quality) can place significant administrative burden on providers. The relative magnitude of the administrative burden associated with quality reporting can vary by provider type and practice size. For example, research suggests that primary care physicians (PCPs) tend to spend more time per week meeting quality measure reporting requirements compared with other specialties.¹² Furthermore, small and medium practices tend to face greater costs per physician associated with reporting quality measures than large practices.¹³ This may be due to them having less developed reporting infrastructures, having fewer preexisting programs that can help meet reporting requirements, and being unable to benefit from economies of scale.

Second, using information from performance measurement to provide feedback and drive improvement can pose challenges including timeliness of data sharing, measure validity, and data quality. For example, Merit-based Incentive Payment System (MIPS) scores are associated more closely with practice characteristics than quality of care. Some research shows that small, independent practices and safety net providers treating high-risk, medically complex patient populations may receive low MIPS scores despite delivering high-quality care.^{14,15,16} Providers have requested that CMS make adjustments to address these challenges, such as releasing quality measure benchmarks that account for relevant patient and provider characteristics prior to the start of each performance period rather than retrospectively so that providers can monitor and improve their performance for the current year.¹⁷

Third, the proliferation of performance measures can also introduce reporting challenges. Providers may be asked to report on different measures with similar objectives depending on the models they participate in or payment policies relevant to their practice. More measures may also lead to a greater likelihood that the same measure will not be implemented consistently across entities such as providers, health plans, or Accountable Care Organizations (ACOs), as entities may confuse similar measures or make mistakes due to reporting fatigue.

The large number of existing performance measures and the varied definitions for measures looking at similar concepts can confuse providers. Because many of these measures are process measures, using two different measures that point to a common clinical scenario may lead to potentially harmful inconsistencies to care delivery or workflow.¹⁸ Recognizing these challenges, CMS has established the

Universal Foundation to better streamline quality measures across various programs. The Foundation¹⁹ aims to create a parsimonious set of measures that are common to as many CMS quality-rating and value-based care programs as possible.

Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures

APMs and PB-TCOC model design take different approaches to addressing challenges associated with the development and implementation of performance measures. The CMS Quality Measure Development Plan outlines strategies to address challenges associated with developing measures, such as partnering with patients and caregivers during the development process, aligning measures across programs, and reducing the burden associated with data collection.²⁰ Value-based purchasing (VBP) models approach incentives in ways that may influence the models' success with performance measurement. These may include providing larger incentives for meeting performance objectives that can compensate providers for their effort in maintaining and reporting relevant performance measures, engaging providers in program design and implementation, and improving the use of health information technology (HIT) and data registries.²¹

Stakeholders note that other changes to value-based purchasing models and policies may also help address the challenges associated with measure implementation. For example, the Medicare Payment Advisory Commission (MedPAC) advised Congress to replace MIPS with a voluntary value component within Medicare FFS. In this new approach, clinicians would choose to have their performance measured as part of a voluntary, self-organized group that collectively assumes responsibility for patient outcomes. Providers that opt in to being measured would qualify for value-based payment based on their group's performance on a set of population-based measures. Unlike MIPS, this approach would reduce reporting burden and spread benefits and risks across participating providers, eliminating extremes in payment adjustments.²²

Adding on to existing models may also help address measurement reporting challenges. For example, the MIPS Value Pathways (MVP) program that launched in 2023 is a new MIPS reporting option. MVP has potential to reduce administrative burden by accounting for population health quality measures which can be calculated without any new reporting requirement for providers. MVP could also align measures specific to a given specialty or health condition, enabling more efficient comparison of performance across providers.²³

Trends in Current Performance Measures

Information on 618 active, in-development, pending, or suspended performance measures (hereinafter referred to as "current performance measures" or "performance measures") was obtained from the CMS Measures Inventory Tool (CMIT).²⁴ Of the 618 total performance measures identified across the 24 programs/models, 375 of these measures are used by only one program/model, and the remainder are used by more than one program/model. Some measures are more likely to be used across many programs/models compared with others. The most commonly used measures among these 24 programs/models are the COVID-19 Vaccination Coverage Among Healthcare Personnel (n=8 programs/models), Medicare Spending Per Beneficiary (MSPB) (n=6 programs/models), and the

Discharge to Community-Post Acute Care (PAC) (n=6 programs/models).ⁱⁱⁱ In addition to information on each measure, the CMIT classifies measures by measure type (e.g., outcomes, process, patient-reported outcomes). More than half (52 percent) of the measures are process measures, followed by outcome measures (26 percent) and patient-reported measures (9 percent). Most of the 24 programs/models include measures from three to four performance measure types.

Furthermore, the measures included in the CMIT are derived from eight different data sources: registries, claims data, EHRs (electronic health records), electronic clinical data that are not from EHRs, paper medical records, standardized patient assessments, administrative data (non-claims), and patient-reported data and surveys. Registry data are the most common among the 24 models/programs, accounting for 24 percent of measures (n=229). Almost 40% of performance measures (n=246) use more than one data source.

NORC obtained information from CMS program or CMMI model websites to categorize the 24 programs/models as either pay-for-performance, pay-for-reporting, or not related to payment. Fifteen programs/models (63 percent) are characterized as pay-for-performance. Eight programs/models (33 percent) are characterized as pay-for-reporting. One program (4 percent) does not currently tie performance measures to payment.

Most performance measures are reported at the facility, hospital, or agency level (31 percent) or at the clinician group practice level (28 percent). Additionally, the CMIT includes information on whether a performance measure is endorsed by the CMS Consensus-Based Entity (CBE), which is discussed further in Section IV. About 34 percent (n=209) of performance measures are endorsed, 59 percent (n=366) of measures are not endorsed, and endorsement was removed for 7 percent (n=42).

The CMIT also provides information on measures that are related to specific clinical conditions, such as cardiovascular disease, diabetes, and infections. Of the 618 measures, 183 (30 percent) measures are not associated with a specific condition. The remaining 435 measures associated with a specific condition belong to 17 programs and one model. MIPS has the highest count of condition-specific measures with 277 measures. The most common condition-specific measures relate to infection (n=67), malignant neoplasm (cancer) (n=51), and cardiovascular disease (n=51).

Lastly, NORC identified the overlap in existing performance measures across the 24 programs/models. NORC created groupings that roll up many measures into high-level categories; for example, screening measures include all types of screening, such as breast cancer screening and colorectal screening. Screening measures accounted for the most performance measures (n=31) across the 24 programs/models.

Similar performance measures were categorized by conceptual area, including screening measures, therapy-related measures for certain chronic conditions, medication-related measures, and measures of the number or rate of admissions or visits constituting the most substantial groups within the performance measures.

ⁱⁱⁱ The Discharge to Community-Post Acute Care (PAC) measure may be used across several settings, including skilled nursing facilities (SNFs), long-term care hospitals (LTCH), and inpatient rehabilitation facilities.

Performance Measures Used in CMMI Models

Publicly available documentation was reviewed for 14 CMMI models that fit the following criteria:

- Includes at least one quality measure and one utilization and/or spending measure in implementation and/or monitoring;
- Is ongoing, under development, or completed within the last five years; and
- Is operational in more than one state.

The selected CMMI models use a diverse range of performance measures encompassing utilization, spending, quality, and patient experience. Quality and spending measures were included in all 14 selected models, while patient experience and utilization measures were included in 11 and 7 models, respectively. Patient experience measures were predominantly obtained through CAHPS surveys. In the majority of selected CMMI models, financial benchmarks are adjusted for patient risk (e.g., using hierarchical condition categories). In addition, some models risk-adjust their quality benchmarks. Participants typically bear responsibility for reporting on a specific set of common measures. One exception is the Bundled Payments for Care Improvement Advanced (BPCI-A) Model, which allows participants to exercise their choice in measure selection. Selected CMMI models link financial incentives to performance measures in different ways, such as by rewarding an absolute level of performance or rewarding relative improvements.

Researchers have raised concerns in gaps in measures and in how measures are used in CMMI models. For example, some have noted that conditions like depression require validated measures with more specificity.²⁵ Other program evaluators emphasize the need for measures related to health equity.²⁶ In addition, previous implementation of CMMI models suggests that equipping providers with additional resources helps to reduce administrative and data collection costs.²⁷

Other lessons learned show that enhanced payments can serve as incentives for improvement in both quality and spending measures.²⁸ Some stakeholders point out that measures should be updated over time to capture relevant changes in diagnosis/procedure codes and that measures for which performance is topped out or measures that produce negative unintended consequences should be removed.^{29,30} Finally, the use of annual wellness visits as a quality measure can help identify gaps in care, as their inclusion of preventive health screenings encourages discussion of health-related social needs.³¹

Performance Measures in the Medicare Shared Savings Program

The Medicare Shared Savings Program (MSSP)—the first ACO model to be integrated into the Part A and Part B Medicare program—uses quality, patient experience, and spending metrics to measure performance. There is substantial overlap in quality measures used between MSSP and several other CMS models; however, MSSP does not include measures related to utilization.

ACOs that participate in MSSP are eligible to receive shared savings (i.e., performance payments) if they reduce costs for the Medicare program while delivering high-quality care. ACOs participating in MSSP may do so under a basic or enhanced track based on their preferred level of risk. The basic track offers a one-sided risk model with the potential for a financial award with no financial risk. ACOs in the basic track can also choose to slowly increase the level of risk and reward each year. ACOs under the enhanced track are subject to the highest potential financial penalties and rewards. In both tracks,

quality thresholds are used to calculate an ACO's quality scores and determine the percentage of shared savings it will receive at the end of each year.

At its inception, MSSP allowed only retrospective attribution – meaning that performance was assessed relative to the experience of patients that had been predominantly visiting an MSSP ACO's participating providers over time. However, since 2019, MSSP has allowed ACOs to choose between retrospective or prospective attribution, which allows ACOs to be assessed on performance related to care for patients they consider part of the ACO.³² Furthermore, the MSSP has started to give participant providers advanced investment payments and more time for ACOs to transition to downside risk. MSSP has also sought to improve access to ACOs for non-white beneficiaries and those in rural and underserved communities. Beginning in 2024, MSSP will offer a health equity adjustment to ACOs that serve a higher proportion of underserved or dually eligible beneficiaries and meet performance thresholds on three quality measures related to care coordination/patient safety, preventive health, and control of specific conditions – diabetes, depression, and hypertension.

Early results suggest that participating ACOs that have financial risk under MSSP deliver better coordinated and more efficient care for Medicare patients compared with participating ACOs with no financial risk under MSSP.³³ In 2022, ACOs participating in MSSP were found to have higher average performance on quality measures, compared with other similarly sized clinician groups.³⁴

Performance Measures in Medicare Advantage Star Ratings Program

The MA Star Ratings Program, developed by CMS, rates the quality of MA (Part C) and prescription drug (Part D) health plans to improve the quality of care delivered to Medicare beneficiaries. The MA Star Ratings Program, similar to the MSSP, uses quality and patient experience metrics tied to financial incentives. Each health plan is designated a Star Rating based on data from the Healthcare Effectiveness Data and Information Set (HEDIS), CAHPS, the Health Outcomes Survey (HOS), and CMS administrative data. The Star Rating determines benchmarks and rebate amounts for the coming year and eligibility for performance-based bonus payment. Star ratings are presented online,³⁵ which can inform beneficiary decisions.

One report on the MA Star Ratings Program noted that health plans have had historically low investment in patient experience and access improvements, resulting in relatively low scores in these domains. In addition, experts have noted that process and outcome measure thresholds should be updated more frequently, for greater accountability and a more dedicated focus on specific measures for health plans.³⁶ Limitations of the Star Rating system include the fact that the program offers generous bonuses but does not financially penalize health plans with low star ratings. The review also noted a limitation related to the underlying data sets, which tend to include measures focused on the needs of beneficiaries who are relatively young and healthy rather than the needs of beneficiaries with serious illnesses. Further, the review noted that there may be limited usefulness of star ratings for beneficiaries because performance is not measured at the plan or local levels.³⁷

Furthermore, CMS uses Star Ratings for other types of providers delivering services under FFS. For example, the Five-Star Quality Rating System assigns nursing homes a rating between 1 (below average quality) to 5 (above average quality). Each nursing home receives one 5-star rating and separate ratings for health inspections, staffing, and quality measures.³⁸ In addition, the Quality of Patient Care Star Ratings reflect the extent to which Home Health Agencies (HHA) differ from one another on quality.

Whereas Star Ratings of 3 suggest average quality of care, ratings above 3 suggest the HHA performed better than average compared to other HHAs and ratings below 3 suggest the HHA performed worse than average compared to other HHAs.³⁹

Performance Measures used in PTAC Proposals

A total of 34 proposals were submitted to PTAC between 2016 and 2020. Proposals included a variety of performance measures related to utilization, spending, quality, and patient experience. Across the 34 proposals, 28 included utilization measures (82 percent), 32 included quality measures (94 percent), 31 included spending measures (91 percent), and 31 included patient experience measures (91 percent). To support an analysis of these measures, proposals were grouped into four categories based on Criterion 2 (Quality and Cost: ability to maintain/improve health care quality while decreasing costs or improve health care quality at no additional cost) and Criterion 4 (Value over Volume: provision of incentives for delivery of high-quality health care). For the purposes of this analysis, the four categories of PTAC proposals included the following:

1. Proposals that met both Criterion 2 and Criterion 4 (N=19)
2. Proposals that met Criterion 2 but did not meet Criterion 4 (N=5)
3. Proposals that met Criterion 4 but did not meet Criterion 2 (N=1)
4. Proposals that did not meet Criterion 2 or Criterion 4, were withdrawn, or determined to be out of scope by PTAC (N=9)

Among the 19 proposals in the first category (i.e., proposals that met both Criteria 2 and 4), all 19 proposals included a quality measure, 15 included a utilization measure (79 percent), 17 included a spending measure (89 percent), and 17 included a patient experience measure (89 percent). Of the five proposals in the second category (i.e., proposals that met Criterion 2 but not Criterion 4), four (80 percent) included a quality measure, four (80 percent) included a utilization measure, and both spending and patient experience measures were included in all five proposals. The sole proposal in the third category (i.e., proposals that met Criterion 4 but not Criterion 2) included a quality, utilization, spending, and patient experience measure. Of the nine proposals in the fourth category (i.e., proposals that did not meet Criterion 2 or Criterion 4), eight (89 percent) included a quality measure, utilization measure, spending measure, and patient experience measure in their proposals in different ways.

The PTAC proposals included a variety of approaches to link payment to performance with differences across the four categories used in this analysis. Most proposals included approaches that linked payment to performance. However, among the nine proposals in the fourth category (i.e., proposals that did not meet Criterion 2 or Criterion 4, were withdrawn, or determined to be out of scope), only five included approaches that tie payment to performance.

Similar to the selected CMMI model analysis, this analysis revealed significant overlap in performance measures across PTAC proposals, especially among utilization, spending, and patient experience measures.

How Performance Measures are Linked with Payment in Other Programs

In addition to the measures associated with the models and programs described in previous sections, there are many other examples of existing performance measures and measure sets used to determine performance-based financial incentives.

Evidence of the effectiveness of performance-based payment programs is mixed. Effectiveness depends on the types of measures being used to assess performance, clinical setting, level of accountability, risk arrangement, magnitude of incentives, and benchmarking process. In terms of effectiveness as it relates to the focus of specific measures, literature demonstrates that performance-based payment (PBP) programs that tie quality measures to financial incentives have positive effects on care coordination,⁴⁰ preventive health screenings,^{41,42} hospital-acquired infections,⁴³ and antibiotic use.^{44,45} PBP programs that focus on outcomes measures have demonstrated little or no impact on blood pressure or cholesterol levels,⁴⁶ HbA1c control,⁴⁷ and sepsis mortality.⁴⁸ Programs that use utilization measures have shown promise for reducing hospital readmissions,⁴⁹ surgical site infections,⁵⁰ and length of stay.⁵¹ Programs that link reimbursement to cost measures have shown mixed results.^{52,53}

There is evidence in acute care that both the Hospital Readmissions Reduction Program (HRRP) and Hospital Acquired Condition Reduction Program (HACRP) are associated with a decreasing 30-day hospital readmissions rate since 2012.⁵⁴ HACRP has also resulted in cost savings to Medicare and reductions in surgical site infections and length of stay.^{55,56} For ambulatory care, PBP programs outperformed traditional FFS payments⁵⁷ and had positive effects on process measures,⁵⁸ but inconsistent or null findings with respect to utilization and health outcomes.⁵⁹

Other findings on PBP programs show more potential for individual over group-level incentives,⁶⁰ pay-for-performance over pay-for-reporting,⁶¹ and two-sided over one-sided risk.^{62,63} Findings on the ideal magnitude of incentives or the best benchmarking process are mixed.⁶⁴ No evidence was found on whether individual or sets of measures are more effective, or on whether standardized compared with flexible measures are more effective.

PBP programs may have some unintended consequences. Evidence suggests that providers who serve low-income patients are disproportionately penalized under such programs and may avoid treating low-income patients.^{65,66} Other potential concerns include lower job satisfaction among clinicians, opportunities for clinicians to game the system, and challenges with administrative burden and patient attribution.⁶⁷

Attribution Methods Used for Models and Performance Measures

Previous environmental scans have captured information on how patients are attributed to models (e.g., claims-based vs. voluntary attribution, prospective vs. retrospective attribution).⁶⁸

Although patients are attributed to a given model, the group of attributed patients for whom any given performance measure is relevant varies. For example, inclusion criteria for the denominator of a performance measure may define patient population by age, setting, diagnosis, or procedures.

Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

The National Research Council Panel on Performance Measures and Data for Public Health Partnership Grants identified five main types of data that are used for constructing performance measures: administrative data, clinical registries, patient medical records, standardized clinical data, and surveys.⁶⁹ Each data type has benefits and drawbacks. Administrative data are easy to access but lack clinical information to guide care and may have low sample sizes. Clinical registries have larger sample sizes and timely reporting but cover a limited range of measures. Per the Improving Medicare Post-Acute Care Transformation Act of 2014 (the IMPACT Act), CMS implemented specified clinical assessment domains

for standardized patient assessment data required for submission by Medicare post-acute care (PAC) providers to CMS. This requirement applies to Medicare PAC providers including skilled nursing facilities, home health agencies, inpatient rehabilitation facilities, and long-term care hospitals.⁷⁰ Patient medical records are comprehensive but may not be interoperable across sites. Survey data are best for capturing patient experience but can be limited if not collected and analyzed rigorously.

Accurate performance measurement relies on the use of high-quality data sources, and efforts should be made to alleviate issues in existing data sources. Data quality can be improved through removing duplicate medical records and claims, adding social determinants of health (SDOH) indicators, linking disparate data sets, using digital health technology tools (DHTE), and applying electronic clinical quality measures (eCQMs).

Risk Adjustment Models Used for PB-TCOC Models

Risk adjustment in PB-TCOC models is important to determine the appropriate payment amounts for attributed enrollees and to account for differences in risk profiles of patients when comparing performance relative to benchmarks. Most CMS programs use a form of the hierarchical condition categories (HCC) model, which calculates a risk score based on diagnoses in health care administrative data or encounter data and select patient demographics. Most CMS programs also use claims in a prior year to project future expected spending (“prospective”), though the ACO Realizing Equity, Access, and Community Health (REACH) Model uses current-year claims information (“concurrent”) for risk adjustment for high needs ACOs. Prospective risk adjustment can serve as a reasonably accurate predictor of a population’s expected spending when average expected spending does not vary considerably over time. Concurrent risk adjustment may be more appropriate when the population size is smaller or when the health status or spending for the given patient population varies over time.^{71,72}

Challenges to risk adjustment include the cost to identify diagnoses that impact a patient’s risk score, which sometimes leads to higher payments and benchmarks, as well as higher spending. In addition, current risk scores are based largely on FFS claims, which may not be an accurate predictor of expected spending in PB-TCOC models. Current risk adjustment methodologies largely do not incorporate social risk factors or functional status that can both be strong predictors of health spending.

Consumer Assessment of Healthcare Providers and Systems Surveys

At CMS’ request, the Agency for Healthcare Research and Quality (AHRQ) developed Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys to collect data on patient experience. Specifically, the CAHPS surveys ask questions regarding issues that matter the most to patients, and topics on which patients are the best or only source of information.⁷³ CAHPS surveys are designed to reliably assess large samples of patients in four main settings: experience with outpatient providers, experience with facility-based care (e.g., care delivered in the hospital or long-term care setting), experience with condition-specific care, and experience with health plans.⁷⁴ To gain information on patient experiences, the CAHPS surveys ask questions across domains such as doctor-patient communication, care coordination, interactions with staff, access to care, and specific care issues. Patient experience receiving care through telehealth is not captured in the CAHPS but is currently being considered.⁷⁵

Best Practices for Measuring Performance

CMS has developed two sets of recommendations that may help guide development and implementation of performance measures in PB-TCOC models. First, the CMS Alliance to Modernize Healthcare, a federally funded research and development center, developed a list of recommendations for performance measurement in population-based payment models. Recommendations include basing measures on results that matter to patients, improving data collection infrastructure, and aligning performance targets with long-term goals.⁷⁶ Second, CMS used the National Quality Forum (NQF) measure evaluation criteria to develop a checklist for creating population-based performance measures.⁷⁷

There are also several emerging considerations for performance measurement in PB-TCOC models, including accounting for SDOH,⁷⁸ measuring equity,⁷⁹ and incorporating tools like artificial intelligence (AI) into performance measurement.⁸⁰ While there is no consensus yet on how to best incorporate these considerations into performance measurement, researchers and CMMI are testing potential approaches.

Opportunities to Improve Performance Measures

Current gaps in health care performance measurement include addressing or accounting for social determinants of health, measuring and promoting equity, and developing broad population-based outcome measures that are feasible and have relevance to patients and clinicians. There are some promising approaches to addressing these gaps. The ACO REACH Model is working to incentivize and improve data collection for SDOH and equity, which may support future measure development.⁸¹ MedPAC⁸² and ASPE⁸³ have also suggested measuring equity by reporting current performance measures separately for specific subgroups. Finally, Patient-Reported Outcome Measures (PROMs) are a promising approach to measure patient symptoms and physical function (e.g., health status). PROMs could help address gaps in feasibility, relevance to patients, and clinical relevance of population-based performance measures.^{84,85}

III. Research Approach

This section provides a brief review of the research questions and methods that were used in developing this environmental scan.

III.A. Research Questions

ASPE staff collaborated with a subset of Committee members known as a Preliminary Comments Development Team (PCDT) to develop^{iv} the following research questions to inform this environmental scan:

- What are the goals of performance measurement for TCOC organizations?
 - What are the main goals of performance measurement for TCOC organizations (for example, to drive change through financial incentives, to provide actionable information for providers, or to inform beneficiary choices)?

^{iv} A Preliminary Comments Development Team (PCDT) comprised four PTAC members: Jennifer Wiler, MD, MBA (Lead); Lawrence Kosinski, MD, MBA; Soujanya (Chinni) Pulluru, MD; and James Walton, DO, MBA.

- How should goals for performance measurement and the measures for PB-TCOC models differ from what is done for FFS payment systems?
- How are these goals incorporated in the design and implementation of performance measures for PB-TCOC models?
- What are best practices to ensure performance measures are implementable and feasible?
- What is it that should be measured?
 - Should the focus be on measurement of outcomes of care and patient satisfaction (and assume that implied objectives such as care coordination and patient-centered care are being achieved), or should changes in how care is being provided be directly measured?
 - What strategies can be used to measure outcomes and/or changes in how care is being provided?
 - What is the relationship between performance measures and goals for value-based transformation?
- How should it be measured?
 - What is the current landscape of what is being measured (for example, outcomes, processes, patient/caregiver experiences)?
 - What gaps exist between what is currently being measured and the measures that are needed for PB-TCOC models?
 - Are there sufficient performance measures for the most prevalent chronic conditions (such as cardiovascular disease)?
 - Are there examples of important episodes of care where quality measures are not available for all parts of the episode?
 - If system change and how organizations provide care should be directly measured, what should be the mixture of outcome, patient experience, and process measures?
 - To what extent can patient/caregiver experience measures accurately reflect the provision of patient-centered, coordinated care relative to direct measures of those processes?
 - Are certain PB-TCOC models' performance measures suited for specific provider types? If so, which measures are best matched with which provider types?
 - Should some of the process quality measures transition to resultant quality outcome measures over time? If so, what are best practices in developing and implementing those resultant quality outcome measures?
 - Should performance-based financial incentives be determined using individual measures or a set of measures?
 - What are examples of existing performance measures and measure sets used to determine performance-based financial incentives, and what are their advantages and disadvantages?
 - How are organization-wide measures versus specialty-specific or setting-specific measures established in the current landscape?
- How should performance measures be linked with payment and financial incentives?

- What research supports that tying performance to payment drives improvement for beneficiaries?
 - Where has this worked, and where has this not worked, and why?
 - What differences in quality improvement are generated with financial incentives that are upside-only as compared with upside and downside incentives?
 - What has been the relative impact of single-sided risk and two-sided risk associated with performance measures on contributing to improvements for beneficiaries?
 - Is there research identifying or supporting the magnitude of payment tied to quality measures that is most effective at incentivizing improvements? If so, what does this research suggest?
 - Does the use of relative or absolute benchmarks result in better performance? If so, what is the optimal approach? In which provider populations is this approach most effective and why?
 - What is the research indicating that a balanced set of measures, that is, inclusion of measures that are intended to guard against decrements in care or unintended consequences, is effective at avoiding unintended consequences?
 - What is the research indicating the value of using a standardized set of measures versus allowing flexibility in selecting measures?
 - How do potential unintended consequences differ based on type of measures (e.g., process versus outcome versus patient-reported) or level of measurement (e.g., clinician, practice, health system)?
 - What could be the potential impact of moving from process measures to outcome measures (e.g., for providers or programs)? Are there any potential unintended consequences?
 - What has been the relative impact of pay-for-reporting versus pay-for-performance programs on contributing to improvements for beneficiaries?
 - What components are most important for tying performance measures to payment and financial incentives?
- What approaches to design of performance-based payment incentives for PFPMs are most likely to facilitate improvement rather than unintended consequences?
- How should financial incentives be tied to quality measures (including outcomes and patient experience measures)? How should benchmarks be determined?
- How should the magnitude of financial incentives be determined?
- How should the approach to performance-based payment differ by the type of entity that is being measured (e.g., larger entities versus small practices, degree of experience with value-based payment)?
- Should the approach to performance-based payment for quality measures vary based on type of measure/measure selection (e.g., evidence-based process measures versus outcome measures or patient-reported outcome measures)? Should payment incentives be phased in for newer measures or certain types of measures? If so, how?
- What are the advantages and challenges of pay-for-reporting incentives compared with pay-for-performance incentives? To what extent do current models and programs utilize pay-for-reporting and pay-for-performance?

- What are best practices for PB-TCOC models to progress toward incentives that increase participants' financial accountability (e.g., transitioning from pay-for-reporting to pay-for-performance)?
- What role can balancing measures (e.g., measures intended to reduce harm/unintended consequences) have in PB-TCOC models?
- What are best practices for designing performance measurement methodologies and for selecting and implementing performance measurement methodologies for population-based TCOC models?
- What are some of the innovative approaches for measuring and incentivizing value-based care transformations?
- What are the different types of performance targets and benchmarks used across APMs? What are their pros and cons?
- What spending, utilization, and quality measures are used in CMMI model implementation and monitoring, and which are used in the MSSP program?
- What are the different attribution methods (e.g., plurality of primary care services, plurality of specialty care services, anchor stay/visit/procedure; prospective or retrospective; claims-based or voluntary) used for performance measures for APMs?
- What are the different types of data sources that can be used for construction of performance measures (e.g., claims, encounters, assessments, clinical registries, surveys)? What are the pros and cons of the different types of data sources that can be used for the construction of performance measures?
- What are the different types of risk adjustment methods used for PB-TCOC models? What are best practices for incorporating risk adjustment/stratification into model implementation?

The full list of research questions, organized by the environmental scan section, is provided in **Appendix A**.

III.B. Research Methods

The environmental scan included information gathered from a targeted review of the literature, an analysis of previous PTAC proposals, and an analysis of selected value-based CMS programs and CMMI models.

The aim of the targeted internet search was to identify and synthesize information from existing peer-reviewed publications and gray literature from organizations focused on developing and implementing performance measures and related challenges. To conduct this internet search, we used search terms specific to each section. For example, for Section XII "How Performance Measures are Linked with Payment in Other Programs," some of the terms we used included "pay-for-reporting," "pay-for-performance," and "performance-based payment." The inclusion criteria focused the search on publications from health care agencies and research organizations between 2013 and the present, in the English language, and based in the United States. The detailed search terms by section are provided in **Appendix B**.

The analysis of PTAC proposals included a thorough review of past proposals, PTAC reports to the Secretary, and content available in other PTAC process documents (e.g., public meeting minutes, Preliminary Review Team [PRT] reports).

The analysis of selected CMS value-based programs and CMMI models was based on a review of publicly available resources, including the description of and technical documents related to each selected program on CMS websites or model on the CMMI website and recent CMMI model evaluation reports for the model, when an evaluation report was available. Where CMMI model evaluation reports were not available on the CMMI website, an internet search was conducted to identify other relevant evaluations, including those that may have been initiated by the participants themselves.

Finally, discussions were conducted with several subject matter experts. These discussions provide additional insights regarding promising ideas and approaches related to performance measurement. A senior member of the contractor team conducted each discussion, while a junior team member took notes. The notes were thematically analyzed, and findings were incorporated into this environmental scan. **Appendix H** provides a summary of the responses from the discussions with subject matter experts.

IV. Background

There are several definitions of performance measurement and different approaches to measuring quality and performance. This section provides an overview of these definitions, as well as a description of the performance measure lifecycle and the goals of performance measurement.

IV.A. Defining Performance Measures

PTAC has developed the following working definition of performance measures as they relate to PB-TCOC models:

*Performance measures **assess and monitor all aspects of participants' performance in models** including quality (e.g., process and structure), outcomes, cost, and utilization.*

PTAC Working Definitions for Certain Performance Measures

PTAC has developed the following working definitions of the following types of performance measures as they are used in PB-TCOC models:

Quality Measures assess the safety, timeliness, effectiveness, efficiency, equity, and patient-centeredness of models. Quality measures may capture structures, processes, and patient experiences with health care. ^{86,87,88,89}

Outcome Measures focus on the health status of a patient resulting from health care.

Cost Measures quantify the cost of healthcare services provided. Cost measures can measure total cost of care or specific costs.

Utilization Measures address the frequency of health care services provided.

PTAC views quality measures as including both structure and process. Additionally, PTAC views outcome measures as including clinical, utilization and cost outcomes. **Exhibit 1** and **Exhibit 2** include examples of different types of quality and outcome measures.

Exhibit 1. Types of Quality Measures

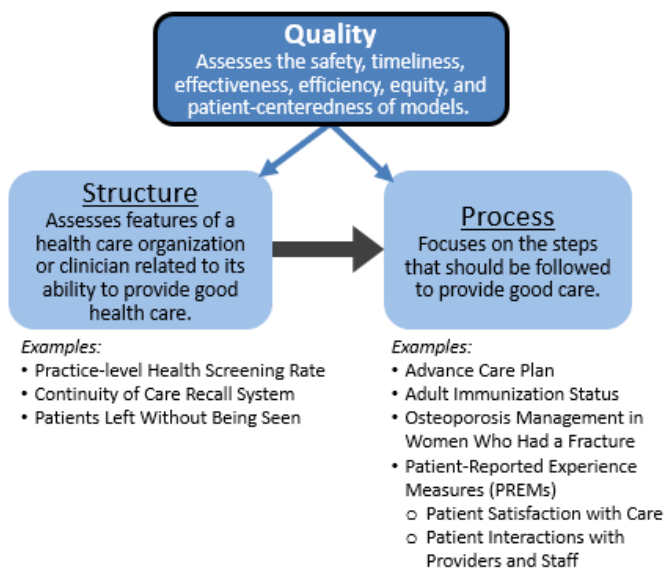
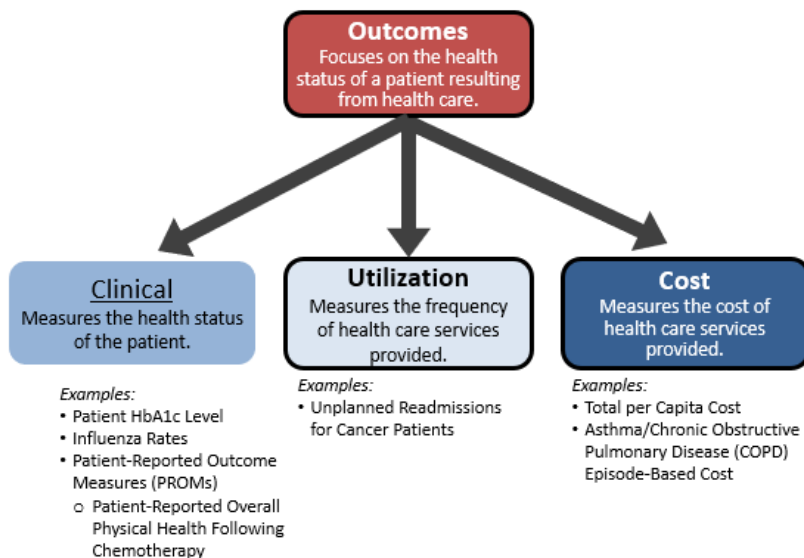


Exhibit 2. Types of Outcome Measures



Please see **Appendix C** for additional information about definitions of various types of performance measures.

Practice-Level Versus Specialty-Specific or Setting-Specific Measures

Several datasets employ physician-level measures, including HEDIS, the CAHPS Clinician & Group Survey, and the Medicaid Child and Adult Health Care Quality Measures. Certain physician specialty societies have created quality measures that are specific to a particular specialty or health condition.⁹⁰ MIPS provides 47 specialty measure sets for eligible clinicians or clinician groups to submit data in 2023.⁹¹

However, most information on quality of performance is available at the medical group or practice level, due to the difficulty in having adequate sample size for individual physician-level measurement.⁹² Measures exist at the setting level for nursing homes, home health agencies, hospice care, acute care hospitals, and long-term care hospitals, among others.⁹³ The 2014 Improving Medicare Post-Acute Care Transformation (IMPACT) Act introduced measures to evaluate post-acute care quality, such as potentially preventable readmissions, beneficiary functional status, and medication reconciliation.⁹⁴ These measures are intended to be used across settings, including skilled nursing facilities (SNFs), home health agencies, inpatient rehabilitation facilities, and long-term care hospitals. Per the IMPACT Act, CMS implemented specified clinical assessment domains for standardized patient assessment data required for submission by Medicare post-acute care (PAC) providers to CMS. In addition, the IMPACT Act required the development of PAC measures under specified quality domains. Measure examples include the Discharge to Community measure referenced above in the Trends in Existing Performance Measures section and the Discharge Function Score.^{95,96,97,98,99}

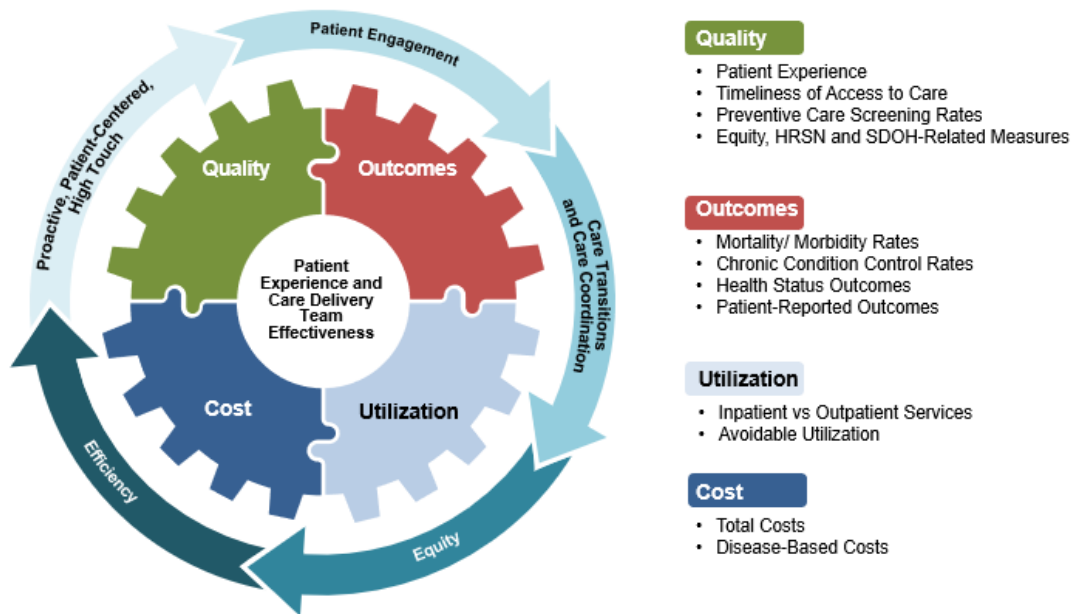
Approaches for Identifying Performance Measures for PB-TCOC Models

PTAC believes that it is important to consider several guiding principles relating to PB-TCOC models when identifying appropriate performance measures for these models, including:

- Providing proactive, patient-centered, high touch care;
- Encouraging patient engagement;
- Managing care transitions and care coordination;
- Achieving equity; and
- Improving efficiency.¹⁰⁰

The guiding principles can be used to identify appropriate types performance measures that can be used to evaluate participating PB-TCOC organizations, and ultimately improve patient experience and care delivery team effectiveness. **Exhibit 3** shows the relationship between the guiding principles and the types of quality, outcome, utilization and cost measure that can be used to evaluate performance related to the guiding principles.

Exhibit 3. Relationship Between Guiding Principles and the Types of Performance Measures for PB-TCOC Models



Additionally, in order to be effective in improving outcomes, PB-TCOC models and providers participating in those models need meaningful performance measures that are related to during the various stages of a patient’s health care journey. **Exhibit 4** provides an example of how the guiding principles of PB-TCOC models can be used to identify meaningful performance measures at various stages in the health care journey for a patient with liver disease.

Exhibit 4. Identifying Meaningful Performance Measures For PB-TCOC Models at Each Stage of the Patient’s Care Journey – Example: Liver Disease

Patient Care Journey	Condition	Guiding Principles for PB-TCOC Models					
		Proactive, Patient-Centered, High Touch	Patient Engagement	Care Transitions and Coordination	Equity	Efficiency	
Health Maintenance	Elevated Liver Enzymes	Access to Preventative Screenings	Patient Understanding of Diagnosis	Timeliness of Consult	Screening Rates Across Populations	Lowered Alanine aminotransferase	
Acute Exacerbation	Liver Failure	Proactive Monitoring	Advance Care Planning with Patient and Caretakers	Timely Access to Hepatologist	Treatment Rates Across Populations	Exacerbation Rate	
Chronic Disease Maintenance	Liver Disease	Proactive Monitoring	Patient Satisfaction with Treatment	Sharing of Patient Data	Outcomes Across Populations	Ambulatory Sensitive Conditions	
Palliative Care	End Stage Liver Disease	Touches with Care Team Members	Shared Decision-making	Successful Transition	Patient-Reported Comfort Across Populations	Reductions in Avoidable Hospitalizations	

Key: Quality Outcomes Utilization Cost

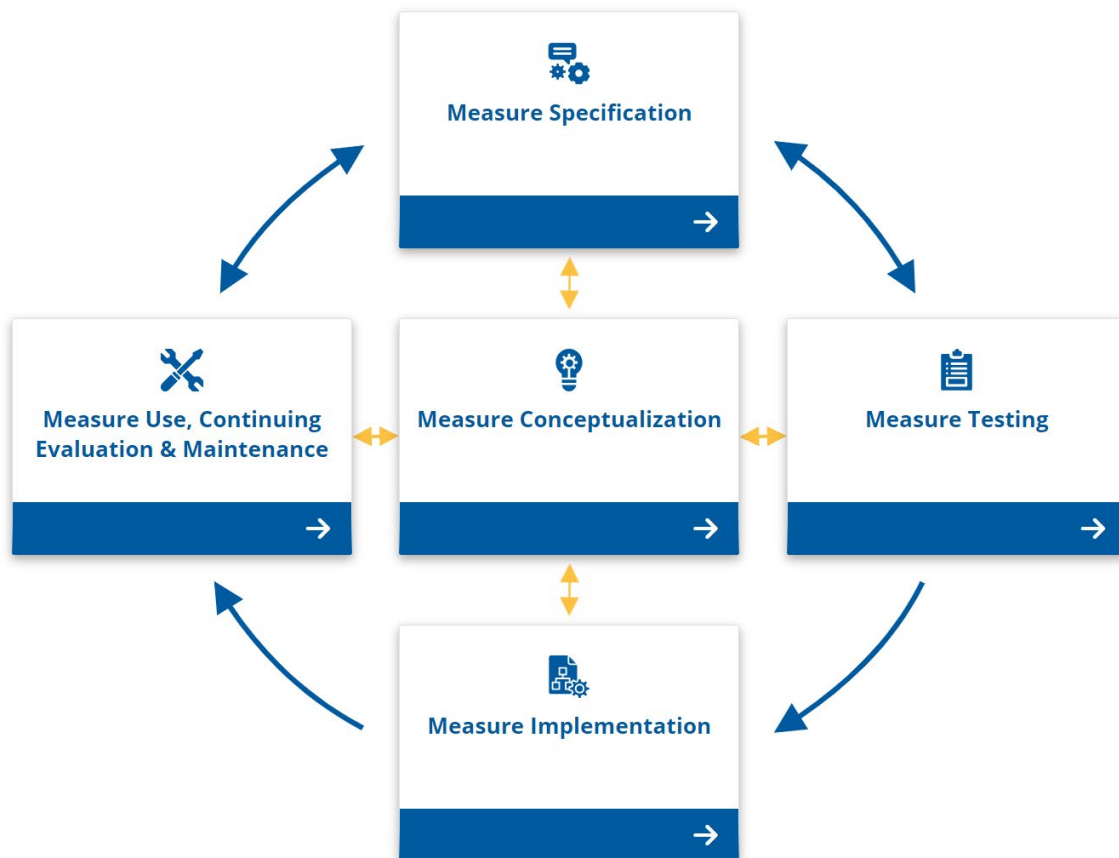
IV.B. Measure Stewards, Measure Developers, and the Measure Lifecycle

Measure stewards are individuals or organizations that “own” measures, meaning they are responsible for maintaining, coordinating, managing, and approving measure content. Measure stewards, who may also be measure developers, are the ongoing point of contact for others interested in a given measure. Stewards have permission to reject or approve and potentially publish measures that their measure developer creates and submits.¹⁰¹

Public agencies, nonprofit organizations, professional medical associations, and other groups may be considered “**measure developers**” who are responsible for the development, implementation, and maintenance of measures, as directed by the measure steward.¹⁰² Measure developers create, edit, and submit measures to a designated steward and are responsible for distributing their measure content for feedback and incorporating potential measure changes suggested by other entities.¹⁰³ For example, CMS serves as the measure steward for most CMS measures and contracts with organizations to develop measures.

Measurement development involves several tasks and stages, which include measure conceptualization; specification; testing; implementation; and use, continuing evaluation, and maintenance.¹⁰⁴ **Exhibit 5** provides an example of the measure development process used by CMS.¹⁰⁵

Exhibit 5. CMS Measure Lifecycle



Measure Conceptualization. Measure conceptualization initiates the measure lifecycle. The measure developer identifies measure concepts by conducting research, analyzing measure gaps, and developing a business case. Identifying gaps in the literature may illustrate areas where measure conceptualization should be prioritized.¹⁰⁶

Measure Specification. The measure developer creates specifications for building and calculating measures. Specifications may include the target population, quality criteria, data sources needed, data protocol, denominator, numerator, proportion of target population needed to meet quality criteria, stratification scheme, timing of measure implementation, definitions of episodes and events, coding system for data collected, measure's documentation, and if applicable, appropriate payment lever for meeting quality criteria.¹⁰⁷

Measure Testing. The measure developer will select the best method to test a measure's feasibility, validity, and reliability. Measure testing occurs simultaneously with measure specification, and developers may refine specifications during the testing stage based on testing results.¹⁰⁸

Measure Implementation. This stage includes all activities involved with moving a measure into the measure use stage, including the endorsement process, measure selection process, and measure rollout. The measure implementation process and the responsible entity vary depending on the program implementing a given measure. The measure developer may provide information to programs on what the measure is collecting and how to interpret results. Data are then collected, analyzed, and used in reports by programs that have implemented a given measure.¹⁰⁹

Measure Use, Continuing Evaluation, and Maintenance. The measure developer is responsible for conducting measure maintenance reviews, including annual updates, comprehensive reevaluation, and early maintenance review. These reviews assure the accuracy and value of a given measure. The measure specifications are reviewed at least annually to ensure that codes are up to date. If the CMS consensus-based entity (CBE) has endorsed the measure, the measure developer reports the results of the maintenance review to re-evaluate its endorsement.¹¹⁰ See Section V for more information pertaining to measure endorsement.¹¹¹

IV.C. Goals of Performance Measurement

Performance measures are used to ensure high-quality patient care and reduce unnecessary hospital expenditures.^{112,113,114} Different health care settings and programs may set specific goals for implementing performance measures, such as reducing patient mortality rates, improving patient satisfaction with care, and increasing the use of EHRs. Programs/models should aim to select measures that address program goals while being mindful of the number of performance measures selected.¹¹⁵ Requiring too many measures could result in staff burden and the collection of unnecessary data, while too few measures may omit reporting of important data related to the goals.

The Committee on Quality Measures for the Healthy People Leading Indicators developed the following criteria for measurement selection for health care delivery:¹¹⁶

- Impact or importance of the condition or outcome to be measured, assessed by the estimated impact on disability, mortality, and economic costs (potentially using predictive and system-based simulation models).

- Improvability, or the extent of the gap between current practice and evidence-based best practice and the likelihood that the gap can be closed. Improvability can be demonstrated by high performance in some sectors or populations, and identification of barriers to improvement.
- Scientific soundness of the measure, including validity and reliability, as evidenced by peer-reviewed literature.
- Geographic, temporal, and population coverage to ensure that the measure has sufficient granularity to be useful in monitoring actions to improve health at different geographic levels in important population subgroups.
- Data availability to ensure that data are readily available in a form useful for quality and performance measurement, meaning that an appropriate national data source exists that can capture relevant performance factors.

V. Challenges Related to Developing and Implementing Performance Measures

The following section describes challenges related to measuring and reporting quality. Additionally, this section describes challenges associated with the proliferation of performance measures, including similarities and differences between current measures and measure definitions.

V.A. Technical Issues Related to Measuring Quality

Providers face several technical challenges when measuring quality, including administrative burden, challenges associated with the timeliness of data sharing, and challenges due to variance across different settings.

Collection, construction, and reporting of quality measures can place significant administrative burden on physician practices. Administrative burden associated with use of quality measures include but are not limited to understanding the measures, reporting data on performance, and comprehending reports of performance from payers.^{117,118} One subject matter expert indicated that practices tend to rely on using claims data for collecting and reporting performance data because claims data are ubiquitous and exchangeable. Extracting clinician information contained in medical records outside of the clinical workflow requires time and resources from providers and care teams.¹¹⁹ The variability that exists in data sources and data quality across settings, as well as variability in staff capacity and analytic resources, creates additional challenges for providers when measuring performance.¹²⁰

A national survey of physician practices indicated physicians from four common specialties—primary care, orthopedics, cardiology, and multispecialty practices that include primary care—and staff spend approximately 785.2 hours per physician annually managing quality measures, including tracking specifications, creating and executing data protocols, entering data into the electronic medical record (EMR), and collecting and transmitting data.¹²¹ Regarding performance measurement activities, physicians and staff reported that the greatest amount of time was spent on “entering information into the medical record only for the purpose of reporting for quality measures from external entities.” The total time spent on coordinating and managing quality measures translated to an average annual cost of \$40,069 per physician. Only about one-fourth of respondents indicated that current performance measures are clearly linked to quality of care.¹²²

In a structured discussion, one subject matter expert suggested that the high cost of quality reporting is, in part, due to the lack of standardization in collecting and reporting quality measures. A lack of

standardization necessitates manual labor, audits, and other workarounds, which can be burdensome and time-consuming for providers. The subject matter expert indicated that overcoming the lack of standardization will increase costs exponentially in the short-term but reduce long-term costs dramatically once the data collection and reporting systems are implemented in a standardized way. Additionally, the subject matter expert stated that the cost and time requirements of data collection and reporting for quality and cost measurement should be removed as much as possible from the clinical workflow so that providers can focus on generating the data.

Participation in specific programs such as MIPS, a track of CMS' Quality Payment Program (QPP), can also place burden on providers. MIPS has complex quality reporting requirements requiring substantial time and financial resources from physician practices. MIPS assesses four domains—quality, improvement activities, promoting interoperability, and cost—to calculate the overall MIPS final score, which is used to determine the payment adjustment. Whereas CMS calculates the measure of cost of patient care for providers, providers are required to submit data for the three other domains. Providers are allowed to select which quality measures they report and which improvement activities they perform.¹²³

A qualitative study examining the time and financial costs for physician practices to participate in MIPS showed that physicians, clinical staff, and administrative staff spent an average of 201.7 hours annually per physician on MIPS-related activities.¹²⁴ In addition, the study showed that practices participating in MIPS spent an average of \$12,811 per physician. Physician practices that were part of a MIPS APM experienced greater per-physician costs compared with practices that were not part of an APM.¹²⁵ Providers participating in MIPS tended to report a low return on investment given the high financial and administrative costs incurred.¹²⁶ As a result of the burden associated with participating in MIPS, some physicians may opt out of reporting their performance data as a requirement of MIPS participation and accept the negative payment adjustment.¹²⁷ Further, because physicians can select which quality measures to report, some physicians may choose to report measures that are easy to achieve or that they performed well on, rather than measures that are clinically relevant to the patient or measures they need to improve upon.¹²⁸

The administrative burden associated with measuring performance may be greater for certain types of providers. On average, PCPs tend to spend more time per week managing quality measures compared with cardiologists, orthopedists, and physicians in multispecialty groups.¹²⁹ PCPs reported concerns that the administrative burden associated with MIPS requirements not only negatively influenced their professional satisfaction but also could lead to patient harm by diverting resources away from patient care.¹³⁰ In addition, because total MIPS scores are based on providers' self-reporting of selected measures rather than improvements in population-level health outcomes, there is concern that MIPS scores better reflect reporting compliance than quality improvement. Stakeholders suggest providers may be encouraged to conduct irrelevant screenings due to incentivizing reporting compliance rather than improving health outcomes.¹³¹ Some measures available for self-reporting as part of MIPS do not measure common clinical activities for specific specialties.¹³²

The administrative burden associated with measuring performance may also vary by practice size. Small and medium physician practices tend to experience greater costs per physician compared with large practices.¹³³ Relatively small practices may lack the administrative capacities and resources that larger practices and health systems have to manage quality measures. Moreover, large practices with many

physicians may be able to benefit from economies of scale and may have more well-developed reporting infrastructures and preexisting programs that can help meet reporting requirements compared with smaller practices. Some physicians in small practices report considering joining larger practices to ease the burdens associated with quality reporting.¹³⁴

Gathering actionable performance feedback to identify gaps in care, reduce variations in care delivery, and improve health outcomes relies on timely data sharing. However, stakeholders have noted that CMS has been slow to meet its statutory obligation of providing timely data and performance feedback as part of the requirements of MACRA.¹³⁵ For example, one challenge of participating in MIPS has been the timeliness of performance feedback and data sharing. Performance feedback on the previous year can be shared with providers six months into the current performance year.¹³⁶ There is also a time lag before claims data are reconciled and mature.¹³⁷ Due to this lag in data sharing, physicians participating in MIPS are unable to monitor their performance throughout each performance year. Participating physicians do not know how they are performing on measures that account for 30 percent of their total MIPS score, on which measures of cost they will be evaluated, or which patients will be attributed to them.¹³⁸ The lag in data sharing also leaves physicians with insufficient time to improve their performance for the current year.¹³⁹

In addition to challenges associated with the timeliness of performance data sharing, there may be issues with the usability and the quality of the data provided. For example, MIPS scores are not comparable across providers because each provider's composite performance score reflects different, self-selected measures.¹⁴⁰ In addition, recent research evidence shows that MIPS scores can be inconsistently associated with performance on process and outcome measures.¹⁴¹ For example, in 2019, nearly one in five PCPs with a low MIPS score (≤ 30 out of 100) had a composite, risk-adjusted outcome score that was in the top quintile of all PCPs participating in MIPS. There is also concern that performance in MIPS is focused on quality of care as it relates to specific patient characteristics rather than the quality of health care provided by physicians.¹⁴² Small, independent practices and safety net providers treating high-risk, medically complex patient populations may receive low MIPS scores despite delivering high-quality care.^{143,144,145} As a result, one subject matter expert indicated that some physicians choose not to treat patients who could lower their MIPS score.

Further, analyses conducted by the American Medical Association (AMA) show that MIPS data files can be incomplete and inconsistent, making it difficult to understand MIPS performance for some types of practices, such as small and/or rural practices.¹⁴⁶ In addition, the data provided in the annual Medicare QPP Experience Report may have limited usability including but not limited to the following:

- Physicians can be counted more than once if they bill services through different organizations;
- Physicians can receive multiple different scores, which are publicly reported (e.g., a low MIPS score for one practice and a high MIPS score for a different practice);
- Performance is not broken down by specialty, site of service, or reporting mechanism; and
- Longitudinal trends on quality and cost over time are not included in the report.¹⁴⁷

To allow physicians to identify opportunities to reduce spending and improve the quality of care, physicians have requested CMS to release prior MIPS cost measures benchmarks, make cost measure benchmarks and attribution information available on a rolling basis, and publish quality measure benchmarks prior to the start of each performance period.¹⁴⁸ Physicians have also suggested that CMS provide more meaningful performance feedback, such as sharing data on how physicians' performance

compares to other physicians of similar specialties or practice sizes.¹⁴⁹ One proposed legislative solution recommends allowing physicians who do not receive at least three quarterly MIPS feedback reports during the performance period to be exempt from MIPS penalties.¹⁵⁰

V.B. Overlap Between Current Measures/Measure Definitions

Before the National Quality Measures Clearinghouse was taken offline in 2018, there were over 2,500 existing measures in its database of evidence-based quality measures and measure sets.¹⁵¹ The CMS Measure Inventory Tool (CMIT), a repository of measures CMS uses to promote health care quality, contains over 500 active measures.¹⁵² Measures still in use may be topped out (i.e., the performance is so consistently high that differences and improvement in performance cannot be observed) and removed from programs and models. The proliferation of performance measures can introduce reporting challenges for clinicians and health care institutions not only because there is a large number of measures but also because there is a lack of standardization in the use of the measures across entities. Further, providers may be asked to report on different measures with similar objectives depending on the models or payment policies relevant to their work. More measures also may lead to a greater likelihood that the same measure will not be implemented consistently across entities, as providers may confuse reporting processes for similar measures or make mistakes due to reporting fatigue.

The definition of performance measurement can differ across entities. Whereas some definitions of performance measurement might focus specifically on outcomes and efficiency, others might include program inputs or outputs (e.g., process measures) and outcomes at different stages.¹⁵³ In addition, the definition of specific performance measures can differ across entities. The use of different performance measure definitions for the same performance measure implemented across health care entities can complicate data collection, reporting, and review processes. For example, the same health condition or area of performance can be defined differently depending on the measure developer.¹⁵⁴ In HEDIS, the National Committee for Quality Assurance (NCQA) defines poor diabetes control as hemoglobin A1C 9 percent or above¹⁵⁵ whereas the American Diabetes Association defines poor diabetes control as hemoglobin A1C 8 percent or above.¹⁵⁶

The large number of measures and variation in measure definitions can lead to confusion among providers.¹⁵⁷ In a national survey of physician practices, 46 percent of practice leaders reported that working with measures that were similar but not identical was a significant burden and recommended addressing this challenge by using measures that are uniform across entities.¹⁵⁸ The multi-payer environment present in the United States has challenged the development of a single, harmonious set of performance measures.¹⁵⁹

Within CMS, each of the agency's quality programs has its own set of quality measures that entities must report. Many models and programs include claims-based measures as claims data can capture valuable information without increasing provider burden. Claims data, however, were not designed for research or performance measurement, include a minimal amount of clinical information, and may not include a complete record of the services provided to a patient.¹⁶⁰ In a discussion, one subject matter expert suggested that having validated and reliable outcome measures will allow population-based models to have fewer measures and be less prescriptive in terms of how care is delivered.

To create a parsimonious set of measures, CMS is streamlining quality measures across its programs through the Universal Foundation. Priorities of CMS' Universal Foundation of quality measures include the following:

- Focusing providers' attention on the measures that are most meaningful for broad segments of the population;
- Reducing provider quality reporting burden;
- Identifying disparities in health care;
- Prioritizing the development of interoperable digital quality measures so that measures are reported automatically;
- Allowing comparisons across quality and value-based programs; and
- Identifying measurement gaps.¹⁶¹

The overarching goals of CMS' Universal Foundation are to select quality measures that are likely to achieve the priorities listed above and will be applicable to as many CMS quality-rating and value-based care programs as possible.¹⁶² In a structured discussion, one subject matter expert suggested that CMS aims to ensure that all measures in the Universal Foundation can be operationalized with United States Core Data for Interoperability (USCDI) as the basis for data collection and with Fast Healthcare Interoperability Resources (FHIR) as the basis for data exchange. The primary domains of the Universal Foundation measures for adults include wellness and prevention, chronic conditions, behavioral health, care coordination, person-centered care, and equity.¹⁶³ CMS indicated that additional population-specific and setting-specific measure sets are needed, such as dialysis care and long-term services.¹⁶⁴

In addition to the Universal Foundation, CMS launched Meaningful Measures 2.0. Meaningful Measures 2.0 is an initiative to reduce the number of measures in CMS programs and improve the impact of quality measures on value-based care. Goals of the initiative include addressing measurement gaps, aligning measures across value-based programs and entities, transforming measures to be fully digital, and developing Patient-Reported Outcome-Based Performance Measures (PRO-PMs) to be incorporated into the workflow.¹⁶⁵ A recent CMS study discussed progress in reducing the number of measures; the number of measures decreased by 15 percent over the past seven years (from 578 measures in 2016 to 492 measures in 2023).¹⁶⁶

V.C. Challenges in Developing New Performance Measures and Obtaining Endorsement of Performance Measures

As mentioned in the previous section, ensuring alignment of measures across programs, payers, and payment systems can be difficult as programs and entities have their own unique needs and requirements.¹⁶⁷ As part of the measure endorsement process, the consensus-based entity (CBE) conducts harmonization reviews on measures to reduce measure duplication and overlap. The CBE will initiate the harmonization process if they find that there are multiple measures that can be combined or streamlined.¹⁶⁸ In addition to harmonizing measures, existing measures may need to be refined or replaced while new measures may need to be developed. There are several challenges associated with developing new performance measures and obtaining endorsement of the measures.

The development of new measures involves multiple steps that can take up to five to six years to complete, such as conducting research, defining measurement concepts and specifications, collecting data to pilot test the measures, conducting data validation, and completing the endorsement

process.^{169,170,171} One challenge with developing new measures is the length of time and the resources required. For example, a 2021 report from the Government Accountability Office (GAO) indicated that a stakeholder group worked with CMS for three years to convert seven pathology-specific registry measures for use in MIPS.¹⁷² The lengthy process required to develop new measures can influence the availability of new measures aimed to address measurement gaps, such as measures for specific specialties.¹⁷³

The length of the measure development process can vary depending on many factors including but not limited to the type of measure developed, the availability of data, and the measure's alignment with the consensus-based review timelines for obtaining measure endorsement.¹⁷⁴ For example, developing measures of shared accountability can pose challenges as developers have to determine the proximal process and/or outcome for which each provider should take responsibility. Recruiting and engaging patients and caregivers to contribute to measure development can help to ensure that measures are person-centered and meaningful. However, including patients and caregivers in the development process can come with challenges as they may be intimidated by the process when participating in technical discussions with subject matter experts.¹⁷⁵

In addition to the challenges associated with the measure development process, challenges can arise while obtaining endorsement of performance measures. The purpose of endorsing performance measures is to ensure that measures used in programs are supported by scientific evidence, are safe and effective, and will increase the likelihood of improving patient health outcomes. Measure developers can support the endorsement process in several ways, including conducting literature reviews to demonstrate the need for the measure, defending the measure's scientific basis or measure testing results, or addressing questions about the measure.¹⁷⁶ The endorsement process involves a pre-submission period, the submission with an automated check for completeness, an internal and public comment period, an endorsement committee review, the dissemination of an endorsement decision, and an appeals process (as needed).¹⁷⁷ Different stakeholders are involved in the endorsement review and decision-making process, including patients, caregivers, clinicians, institutions, purchasers and plans, rural health experts, health equity experts, researchers, and other interested parties (e.g., representatives of EHR vendors). Criteria used for endorsing measures include the importance of the measure (e.g., the measure's anticipated impact on outcomes), the measure's feasibility, the validity or reliability of the measure, the measure's contribution to addressing inequities in health care, and the measure's use and usability. Endorsed measures are re-evaluated every five years to ensure maintenance over time.¹⁷⁸

Obtaining measure endorsement can be a lengthy process. It can take up to six months between the time the measure is submitted for endorsement to the time when an endorsement decision is made.¹⁷⁹ Although measures do not have to be submitted for endorsement, CMS is more likely to use measures that have been reviewed and endorsed by a consensus-based entity. MACRA allows the use of non-endorsed measures if the measures are evidence-based. However, an understanding of what evidence-based entails and how to evaluate evidence has not always been clearly defined.¹⁸⁰

V.D. Challenges in Implementing New Performance Measures

Performance measure implementation includes any activity focused on progressing a performance measure from the development state to an active state. Measure implementation includes measure selection processes and measure rollout.

Measure developers must complete several steps as part of measure selection and rollout processes. Measure selection refers to determining whether a measure should be included in a measure set for a specific health care program. CMS measure selection criteria have been established to ensure that measure developers completed person- or encounter-level testing, reliability testing, and validity testing, and adjust the measure based on the results of these efforts, if needed.¹⁸¹

Measures are rolled out after they are selected and approved for use in CMS programs. The process of rolling out a new measure in a CMS program can vary across measure sets depending on the program, the scope of the measure implementation, the measured entity, the process for data collection, and the planned use for the measure. Preparing for implementation of a measure can include planning for the initial rollout of the measure, data management and production, audits, validation efforts, provision of education to end users on what is being measured and how to interpret results, dry runs, and appeals processes.¹⁸² The purpose of conducting a dry run with a measure is to finalize the methodologies associated with the measure, including case identification/selection, data collection, measurement calculation, and how unintended consequences will be quantified. Following dry runs, developers assess the measure's success in meeting the needs of the program into which the measure was adopted.¹⁸³ In an interview, a subject matter expert indicated that it is often unclear how to implement measures. To ensure measure feasibility, measure developers are encouraged to explicitly describe what is required to implement the new measures, including identifying the technology, changes to workflow, changes to processes, and special skill sets required to implement the measures.

Some program-specific requirements can add steps to the implementation of new performance measures. For example, MIPS requires that new measures are submitted for publication in a specialty-specific, peer-reviewed journal.¹⁸⁴ Writing and publishing a manuscript can be an arduous and lengthy process.

VI. Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures

APMs and PB-TCOC model design may consider ways to address the challenges associated with the development and implementation of performance measures. This section describes potential strategies and opportunities APMs and PB-TCOC models can use to address the challenges, as well as examples of models that may address the challenges.

VI.A. Discussion of Potential Opportunities

MedPAC established several key principles for measuring quality in health care, including encouraging coordination across providers and across time; including population-based measures such as outcomes, patient experience, and value; rewarding performance based on clear, prospectively set targets; and not placing excessive burden on providers.¹⁸⁵

The CMS Quality Measure Development Plan outlines strategies to address challenges associated with developing and implementing measures, including the administrative burden of quality reporting, the timeliness of performance feedback and data sharing, potential issues with data usability and quality, and the length of the measure development and endorsement process. The following strategies can be applied during measure development to ensure that measures are meaningful and appropriate:

- Partnering with patients, families, and caregivers during measure development, including diverse patient and caregiver groups (e.g., children, older adults, individuals who are dually eligible);
- Partnering with frontline clinicians and professional societies to develop measures that will be meaningful to providers;
- Aligning measures (i.e., using the same measure concepts) across programs, payers, and payment systems;
- Reducing the burden associated with data collection for performance measure reporting by collecting data that are part of the clinical workflow (e.g., collecting data directly from patients);
- Shortening the timeframe for performance measure development by incorporating Lean principles^v into the measure development workflow and facilitating transparency and knowledge sharing across measure developers through forums and measure developer libraries;
- Streamlining data acquisition for measure testing by utilizing broader data sources for measure development and increasing data sharing and reuse by measure developers;
- Developing meaningful, person-centered outcome measures by utilizing evidence-based research and incorporating patients', caregivers', and clinicians' perspectives, concerns, preferences, and goals into measure design;
- Developing patient-reported outcome measures by first using building block process measures as part of the Medicare EHR Incentive Program that help to encourage development of future outcome measures;
- Developing appropriate use (e.g., overuse) measures by applying recommendations from the Choosing Wisely campaign^{vi} and encouraging the development of balancing measures to mitigate unintended consequences; and
- Developing measures that encourage shared accountability across providers and settings by improving the adoption and use of health information exchange and encouraging interoperability across health information infrastructures.¹⁸⁶

In addition to the strategies outlined above, a review by RAND summarized several features of VBP models that may influence their success during implementation. These model features included the following:

- Providing larger incentives that adequately compensate physicians for their efforts and investments in improving care;
- Ensuring measure alignment across programs, particularly prioritizing alignment of measures with the patient population's health conditions;
- Considering provider engagement in the design and implementation of programs, including objective targets defined prior to the start of the measurement year and avoiding having providers compete against other providers through relative thresholds; and
- Improving the use of HIT and data registries.¹⁸⁷

By adopting and incorporating the strategies listed above, APMs and PB-TCOC models can help to address the challenges associated with developing and implementing performance measures.

^v Lawal AK, Rotter T, Kinsman L, et al. Lean management in health care: Definition, concepts, methodology and effects reported (systematic review protocol). *Syst Rev.* 2014;3:103. doi:10.1186/2046-4053-3-103

^{vi} Choosing Wisely. An Initiative of the ABIM Foundation. <https://www.choosingwisely.org/>

VI.B. Examples of Successful Models

The strategies and design features of models that link financial incentives to payment in order to make improvements in health care quality have evolved over time based on lessons learned. Understanding how to best design models to achieve the desired goals of improving health care quality and reducing unnecessary hospital expenditures and facilitate implementation is considered a work in progress.¹⁸⁸

New models and programs can be developed to address the challenges associated with measure implementation. For example, MedPAC advised Congress to eliminate MIPS because they believed the program would not fulfill its goals to reliably measure and reward quality and would impede the movement toward high-value care.¹⁸⁹ Instead, MedPAC advised Congress to implement a voluntary value component within Medicare FFS where clinicians can elect to be measured as part of a voluntary, self-organized group that collectively assumes responsibility for patients' outcomes. If physicians elect to be measured, they would be eligible to receive a value-based payment based on performance on a group of population-based measures. In this model, all providers across all specialties would receive incentives to improve population-based outcomes. Quality data would not need to be reported to Medicare as measures would be calculated by CMS using claims data and surveys.¹⁹⁰

Modifying existing models and programs can also help to address the challenges associated with measure implementation. For example, launched in 2023, the MIPS Value Pathways (MVP) program is intended to address some of the challenges faced by providers participating in MIPS and may have a bigger impact on reducing health care costs and improving health outcomes. As the newest MIPS reporting option, MVP has potential to address administrative burden associated with MIPS participation by including population health quality measures that CMS can calculate without requiring health care reporting. In addition, MVP will allow comparisons across providers of the same specialty by aligning measures across different reporting categories for specific specialties and health conditions. MVP will also include patient-centered measures (e.g., patient-reported outcomes, patient experience, patient satisfaction) and encourage reporting of more clinically relevant measures.¹⁹¹ Moving forward, CMMI plans to expand its definition of success for its models and add metrics focused on health equity, person-centered care, and health system transformation.¹⁹²

VII. Trends in Existing Performance Measures

The following section summarizes the number and characteristics of performance measures currently used by 24 CMS Medicare payment programs (including nine CMS value-based care programs and eight CMS pay-for-reporting programs) and seven CMMI models.^{vii} Information for the 24 programs/models

^{vii} The 24 programs/models include 17 programs: Ambulatory Surgical Center Quality Reporting (ASCQR); End-Stage Renal Disease (ESRD) Quality Incentive Program (QIP); Home Health Quality Reporting (QR); Hospice Quality Reporting program (HQR); Hospital Acquired Condition (HAC) Reduction Program; Hospital Outpatient Quality Reporting (OQR); Hospital Readmission Reduction Program (HRRP); Hospital VBP; Inpatient Psychiatric Facility Quality Reporting (IPFQR); Inpatient Rehabilitation Facility Quality Reporting (IRFQR); Long-Term Care Hospital (LTCH) QR; Medicare Advantage (MA) Star Ratings Program; Medicare Shared Savings Program (MSSP); Merit-based Incentive Payment System (MIPS) Program; Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (CHQR); Skilled Nursing Facility (SNF) QR; SNF VBP; and include seven models: Accountable Care Organization (ACO) Realizing Equity, Access, and Community Health (REACH) Model; Bundled Payment for Care Improvement Advanced (BPCI-A) Model; Home Health Value-Based Purchasing (VBP); Independence at Home (IAH) Demonstration; Kidney Care Choices (KCC) Model; Oncology Care Model (OCM); and Primary Care First (PCF) Model.

was obtained from the CMS Measures Inventory Tool (CMIT), which is a repository of performance measure information that includes 46 CMS value-based care programs, CMS pay-for-reporting programs, or CMMI models (as of October 2023 when CMIT data were pulled for this analysis). For each measure, the CMIT includes program/model name, measure name, measure definition, measure type, and measure source.^{viii} The CMS programs (**17**) were selected to ensure a variety of federal reporting programs (e.g., pay-for-performance, pay-for-reporting, quality reporting and other approaches). The CMMI models (**seven**) were selected based on the following criteria: 1) the model must have been active in the last five years; 2) the model must include at least one quality measure and at least one utilization or spending measure in implementation and/or monitoring; 3) the model must be or have been operational in more than one state; and 4) the model is included in the CMIT. High-level findings are summarized below; for additional information, see the full report “Overview of Current Performance Measures Included in Selected Medicare Payment Programs” on the PTAC website under [Resources](#).

VII.A. High-Level Findings and Summary Statistics for Performance Measures in 24 Select Models or Programs

Total Current Performance Measures

There are 618 active, in-development, pending, or suspended^{ix} performance measures^x (hereinafter referred to as “current performance measures” or “performance measures”) in the CMIT for the 24 selected programs and models included in this analysis. Of these 618 measures, 523 (84.6 percent) are actively being used in CMS programs, and 89 (14.4 percent) are in development, pending, or suspended. It is important to note that the 618 measures are not all unique as some measures are used by multiple programs/models (see further discussion below).

Exhibit 6 provides a breakdown of the number of performance measures by program/model. MIPS includes the most performance measures among the 24 programs/models with 309 performance measures (50 percent) of the 618 measures.^{xi} The number of measures included in the other 23 programs/models ranges from 3-33. The MA Star Ratings Program included the second most performance measures with 33 performance measures (5 percent of the 618 measures).

^{viii} Centers for Medicare and Medicaid Services Measures Inventory Tool, <https://cmit.cms.gov/cmit/#/MeasureInventory>

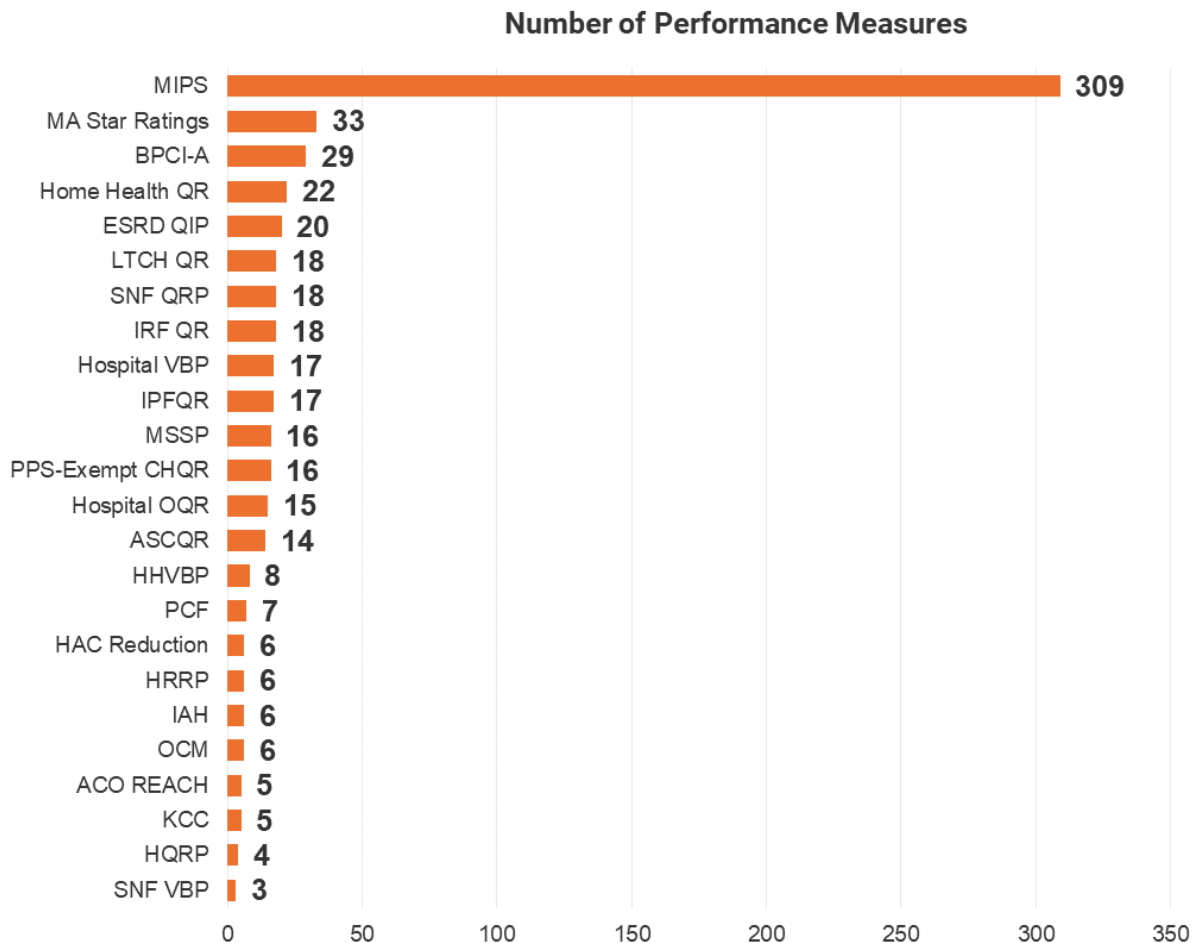
^{ix} Suspended measures may include measures that were temporarily suspended due to the Public Health Emergency (PHE). One example is the Skilled Nursing Facility (SNF) Value-Based Purchasing (VBP) measure, which uses the 30-day all-cause readmission measure; this measure is currently listed in the CMIT as suspended.

^x Inactive measures are not included with the exception of the Oncology Care Model (OCM); because this program is not active, all measures were also inactive. To include this model in this analysis, NORC included the six inactive measures tied to the OCM, which are a part of the 618 measures used in this analysis.

^{xi} Participants in MIPS choose at least six quality measures (one must be an outcome measure) from the full list of measures, and CMS calculates and scores each participant on four administrative claims measures. Participants are not scored on all measures.

Of the seven CMMI models^{xii} included in the CMIT, the Bundled Payment for Care Improvement Advanced (BPCI-A) Model includes the most performance measures with 29 measures (5 percent of the 618 measures); the remaining six CMMI models use five to seven performance measures each.

Exhibit 6. Number of Performance Measures by Program/Model



Note: Current performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Distinct Performance Measures

Most (75%) of the 618 current performance measures in the 24 programs/models included in this analysis are distinct or “unduplicated” measures (n=455).^{xiii} This number includes active measures, as well as measures that are in development, pending, or suspended.

^{xii} The seven models included are the Bundled Payment for Care Improvement Advanced (BPCI-A) Model, Accountable Care Organization (ACO) Realizing Equity, Access, and Community Health (REACH) Model, Home Health Value-Based Purchasing (VBP) Model, Independence at Home (IAH) Demonstration, Kidney Care Choices (KCC) Model, Oncology Care Model (OCM), and the Primary Care First (PCF) Model.

^{xiii} The number of distinct measures represents the number of current performance measures with distinct names (i.e., if each measure name is counted one time).

Further, 375 of the performance measures (61 percent of the 618 total current performance measures) are used for only one program or model.

Additionally, there are 163 measures (26 percent of the 618 total currently performance measures) that are used by more than one program. These 163 measures may use different numerators, denominators, or denominator exclusions. Exhibit 7 provides two examples of measures that are used by more than one program but are defined in slightly different ways.

Exhibit 7. Two Examples of Measures Where Programs/Models Apply Different Criteria to the Same Measure

Measure	Program/Model	Differences in Measure
Colorectal Screening (Measure ID: 139)	Medicare Advantage (MA) Star Ratings Program; Medicare Shared Savings Program (MSSP); Merit-Based Incentive Payment System (MIPS); Primary Care First (PCF) Model	MIPS uses a denominator that includes patients 50-75 years of age while the three other programs/models use a denominator that includes patients 45-75 years of age.
Controlling Blood Pressure (Measure ID: 167)	MSSP; MIPS; PCF	PCF differs in its denominator exclusion criteria from the other two programs: its denominator excludes pregnant women and does not exclude patients 81 years of age or older with an indication of frailty beyond those with advanced illness.

Exhibit 8 lists the eight^{xiv} performance measures that are most often used across the 24 programs/models. The most common performance measure is COVID-19 Vaccination Coverage Among Healthcare Personnel used by eight different programs (33 percent). The top eight measures listed in Exhibit 8 include four outcome measures, three process measures, and one cost/resource use measure.

Exhibit 8. Top 8 Performance Measures by Number of Programs/Models

Measure ID	Measure Name	Measure Type	Number of Programs/Models	Included Programs ^{xv}
180	COVID-19 Vaccination Coverage Among Healthcare Personnel	Process	8	LTCH QRP; PPS-Exempt CHQR; ASCQR; ESRD QIP ¹ ; Hospital OQR; IPF QR; IRF QR; SNF QRP

^{xiv} Top eight performance measures were chosen because it was a clean break from five to four programs; there were 13 performance measures with four programs each.

^{xv} See Appendix D for the full names of each program/model.

Measure ID	Measure Name	Measure Type	Number of Programs/Models	Included Programs ^{xv}
434	Medicare Spending Per Beneficiary (MSPB)	Cost/Resource Use	6	Hospital VBP ² ; IRF QR; MIPS; LTCH QRP; SNF QRP; Home Health QR
210	Discharge to Community-Post Acute Care (PAC)	Outcome	6	Home Health VBP; Home Health QR; IRF QR; LTCH QR; SNF QRP; SNF VBP ³
462	National Health Safety Network (NHSN) Facility-Wide Inpatient Hospital-Onset Clostridium Difficile Infection (CDI) Outcome Measure	Outcome	5	HAC Reduction; Hospital VBP ⁴ ; IRF QR; LTCH QRP; PPS-Exempt CHQR
459	NHSN Catheter-Associated Urinary Tract Infection (UTI) Outcome Measure	Outcome	5	HAC Reduction; Hospital VBP ⁵ ; IRF QR; LTC QRP; PPS-Exempt CHQR
356	Hospital-Wide All-Cause Unplanned Readmission Measure (HWR)	Outcome	5	IAH Demonstration; ACO REACH; BPCI-A; MSSP; MIPS
727	Transfer of Health Information to the Patient Post-Acute Care (PAC) / Transition Record with Specified Elements Received by Discharged Patients (Discharges from an Inpatient Facility to Home/Self-care or Any Other Site of Care) ⁶	Process	5	Home Health QR; LTCH QR; SNF QRP; IRF QR; IPF QR
728	Transfer of Health Information to the Provider PAC / Timely Transmission of Transition Record (Discharges from an Inpatient Facility to Home/Self-care or Any Other Site of Care) ⁷	Process	5	Home Health QR; LTCH QR; SNF QRP; IRF QR; IPF QR

Note: All data are as of October 2023, when CMIT data was pulled. Unless otherwise indicated, measures included in this table are active within each program/model.

¹ The COVID-19 Vaccination Coverage Among Healthcare Personnel Measure (measure ID 180) is a pending measure for the ESRD QIP.

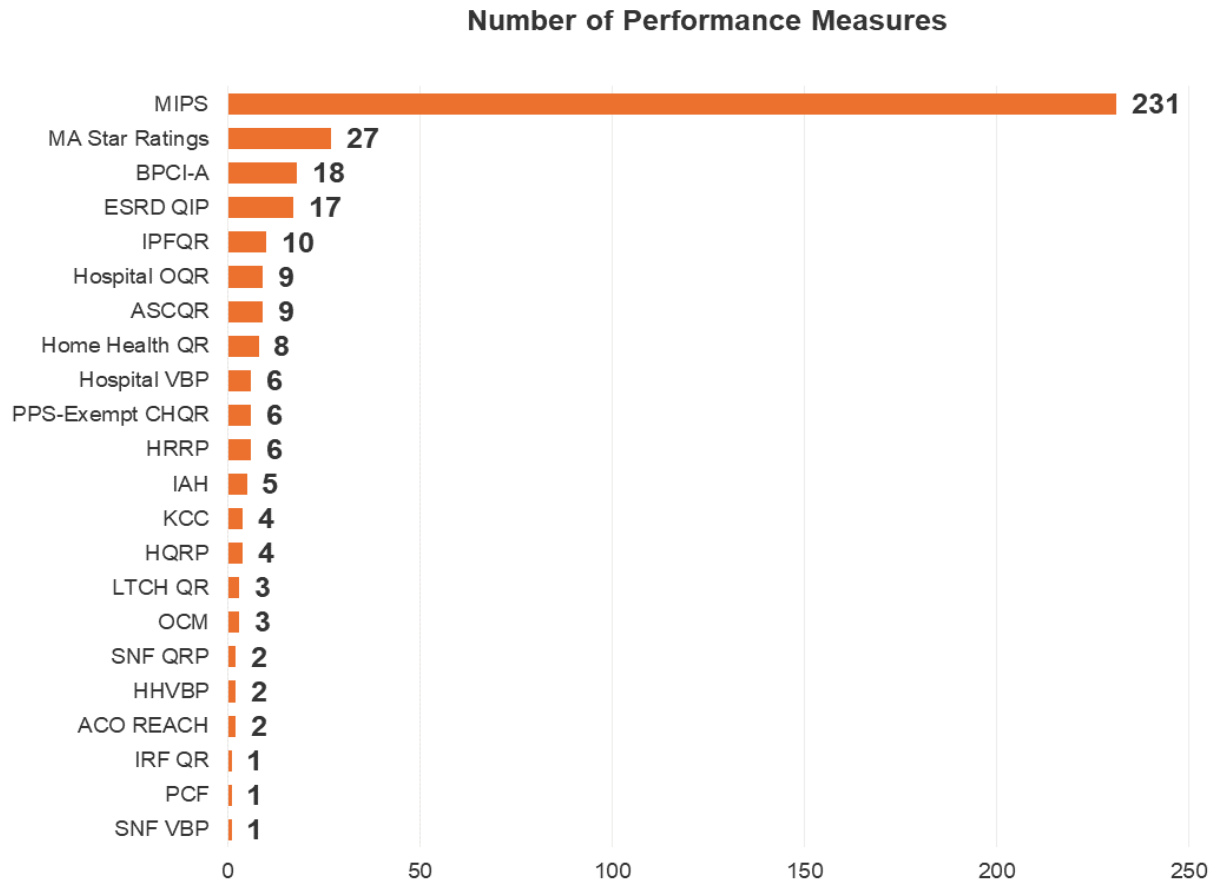
² The MSPB Measure (measure ID 434) is a suspended measure for the Hospital VBP Program.

³ The Discharge to Community-PAC Measure (measure ID 210) is a pending measure for the SNF VBP Program. ⁴ The NHSN Facility-Wide Inpatient Hospital-Onset CDI Outcome Measure (measure ID 462) is a suspended measure for the Hospital VBP Program. ⁵ The NHSN Catheter-Associated UTI Outcome Measure (measure ID 459) is a suspended measure for the Hospital VBP Program. ⁶ For measure ID 727, the IPF QR Program uses the measure name of Transition Record with Specified Elements Received by Discharged Patients (discharges from an inpatient facility to home/self-care or any other site of care); all other listed programs use the measure name of Transfer of Health Information to the Patient PAC. ⁷ For measure ID 728, the IPF QR Program uses the measure name of Timely Transmission of Transition Record (Discharges from an Inpatient Facility to Home/Self-care or Any Other Site of Care); all other listed programs use the measure name of Transfer of Health Information to the Provider PAC.

Of the 618 total measures examined across the 24 programs/models, 375 measures (61 percent), involving 22 programs, are used by only one program/model. The Measure ID column from the CMIT was used to identify these measures. **Exhibit 9** provides the counts of performance measures used for

only one program/model listed by program/model. Of the 375 performance measures used by only one program, 231 measures (62 percent) are used in only the MIPS Program. Other programs/models range from one to 27 measures that are specific only to that program or model. All except two programs/models (Medicare Shared Savings Program [MSSP] and Hospital Acquired Condition [HACRP]) include measures exclusive to their programs/model.

Exhibit 9. Number of Performance Measures Used for Only One Program/Model



Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Note: Two programs do not have measures exclusive to their programs (MSSP and HACRP).

Types of Performance Measures

This analysis focuses on the 618 total performance measures across the 24 programs/models (not the 455 distinct performance measures) in order to assess performance measures at the program level.

The CMIT includes seven types of performance measures: process, outcome, intermediate outcome, patient-reported, cost/resource use, structure, and composite measures. **Exhibit 10** provides CMS definitions for these seven measure types, as well as examples of each measure type as listed in the CMIT.

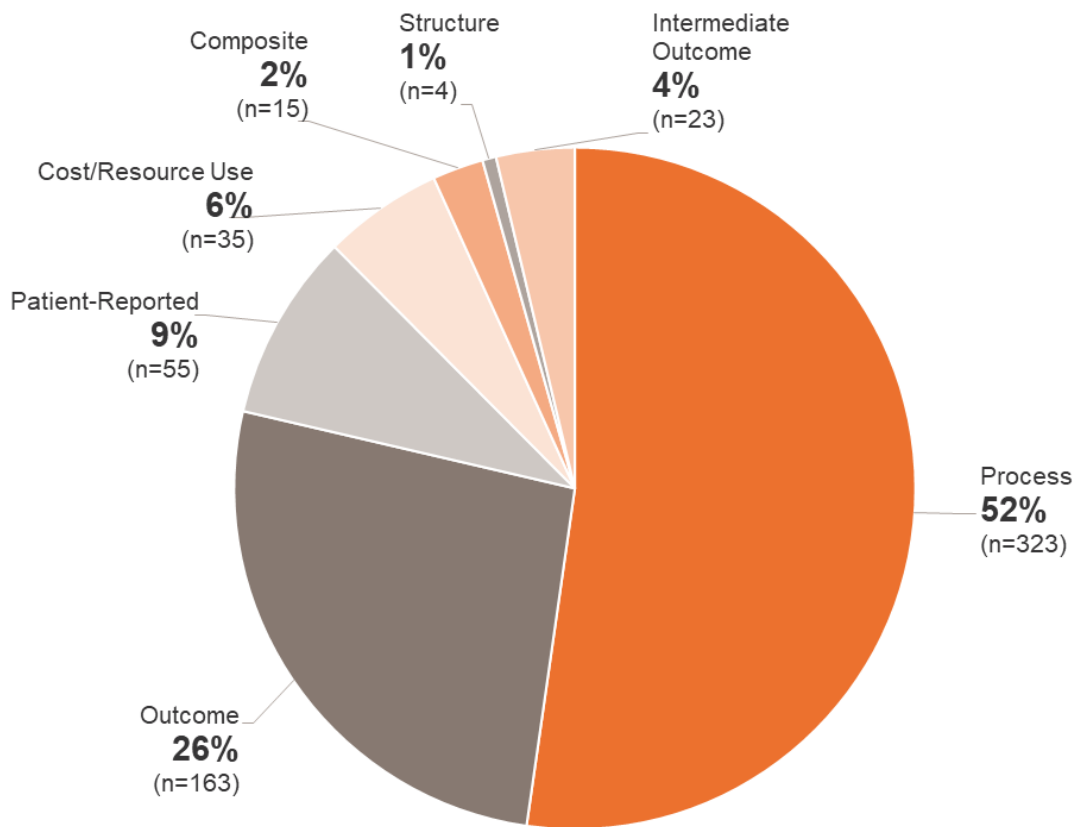
Exhibit 10. CMS Definitions of the Seven Performance Measure Types

Measure Type	Definition ^{xvi}	Examples
Composite Measure	Two or more measures that form a combined measure	Severe Sepsis/Septic Shock: Management Bundle; Patient Safety and Adverse Events Composite; Substance Use Screening and Intervention Composite
Cost/Resource Use Measure	Measures the cost or frequency of health care services provided	Asthma/Chronic Obstructive Pulmonary Disease (COPD) Episode-Based Cost Measure; Medicare Spending per Beneficiary (MSPB); Total Per Capita Cost
Intermediate Outcome Measure	Assesses the change that occurs from treatment resulting in a long-term outcome	Controlling High Blood Pressure; Diabetes: Hemoglobin A1c (HbA1c) Poor Control (> 9%); Kidney Transplant Referral Rate
Outcome Measure	Focuses on the health status of a patient that changes due to care or treatment received	Acute Care Hospitalization; Unplanned Readmissions for Cancer Patients; Patient Fall
Patient-Reported Outcome-Based Performance Measure (PRO-PM)	Based on patient-reported outcome measure (PROM) data aggregated for the responsible health care entity	CAHPS; Functional Status Change for Patients with Hip Impairments; Patient-Reported Overall Physical Health Following Chemotherapy
Process Measure	Focuses on the actions to be followed to provide adequate care	Advance Care Plan; Adult Immunization Status; Osteoporosis Management in Women Who Had a Fracture
Structure Measure	Evaluates health care organizations related to its ability to provide adequate health care	Health Screening Rate; Continuity of Care Recall System; Patients Left Without Being Seen

Exhibit 11 displays the distribution of performance measures by measure type. Of the 618 performance measures included across the 24 programs/models, more than half of the performance measures (323 measures or 52 percent) were process measures. Outcome measures were the second most common measure type (26 percent, n=163); patient-reported measures constitute 9 percent (n=55); cost/resource use measures are 6 percent (n=35). Intermediate outcome, composite, and structure measures constitute only 4 percent, 2 percent, and 1 percent of performance measures, respectively.

^{xvi} <https://mmshub.cms.gov/about-quality/new-to-measures/types>

Exhibit 11. Distribution of Performance Measures by Measure Type for the 24 Programs/Models



Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Exhibit 12 provides a breakdown of the 618 total performance measures by program/model and measure type. Most of the 24 programs/models use about three to four different performance measure types. MIPS is the only program that uses six measure types (all except composite measures). No program/model uses all seven measure types. Twenty-two programs/models (92 percent)—all except Kidney Care Choices (KCC) and Hospice QR—use outcome or intermediate outcome measures. Nineteen programs/models (79 percent) use process measures; 15 (63 percent) use at least one patient-reported outcome measure; five (21 percent) use composite measures.^{xvii} Only three programs/models (12 percent) use structure measures.^{xviii}

^{xvii} BPCI-A (n=7), IPF QR (n=3), HHVBP (n=2), HQR (n=2), and HAC Reduction Program (n=1) use composite measures.

^{xviii} MIPS (n=2), ESRD QIP (n=1), and Hospital OQR (n=1) use structure measures.

Exhibit 12. Distribution of Performance Measures by Program/Model and Measure Type

Program/Model	Composite	Cost/Resource Use	Intermediate Outcome	Outcome	Patient-Reported	Process	Structure	Total
Merit-based Incentive Payment System (MIPS) Program	0 (0%)	24 (8%)	8 (3%)	43 (14%)	31 (10%)	201 (65%)	2 (1%)	309
Medicare Advantage (MA) Star Ratings	0 (0%)	0 (0%)	1 (3%)	1 (3%)	7 (21%)	24 (73%)	0 (0%)	33
Bundled Payment for Care Improvement Advanced (BPCI-A)	7 (24%)	0 (0%)	0 (0%)	8 (28%)	1 (3%)	13 (45%)	0 (0%)	29
Home Health Quality Reporting (QR)	0 (0%)	1 (5%)	0 (0%)	12 (54%)	1 (5%)	8 (36%)	0 (0%)	22
End-Stage Renal Disease (ESRD) Quality Incentive Program (QIP)	0 (0%)	0 (0%)	6 (30%)	5 (25%)	1 (5%)	7 (35%)	1 (5%)	20
Long-Term Care Hospital (LTCH) QR	0 (0%)	1 (6%)	0 (0%)	9 (50%)	0 (0%)	8 (44%)	0 (0%)	18
Skilled Nursing Facility (SNF) Quality Reporting Program (QRP)	0 (0%)	1 (6%)	0 (0%)	9 (50%)	0 (0%)	8 (44%)	0 (0%)	18
Inpatient Rehabilitation Facility (IRF) QR	0 (0%)	1 (6%)	0 (0%)	11 (61%)	0 (0%)	6 (33%)	0 (0%)	18
Hospital Value-Based Purchasing (VBP)	0 (0%)	4 (24%)	0 (0%)	12 (70%)	1 (6%)	0 (0%)	0 (0%)	17

Program/ Model	Composite	Cost/ Resource Use	Intermediate Outcome	Outcome	Patient- Reported	Process	Structure	Total
Inpatient Psychiatric Facility (IPF) QR	3 (18%)	0 (0%)	0 (0%)	1 (6%)	0 (0%)	13 (76%)	0 (0%)	17
Medicare Shared Savings Program (MSSP)	0 (0%)	0 (0%)	4 (25%)	3 (19%)	1 (6%)	8 (50%)	0 (0%)	16
Prospective Payment System (PPS)- Exempt Cancer Hospital Quality Reporting (CHQR)	0 (0%)	0 (0%)	2 (13%)	8 (50%)	1 (6%)	5 (31%)	0 (0%)	16
Hospital Outpatient Quality Reporting (OQR)	0 (0%)	0 (0%)	0 (0%)	3 (20%)	2 (13%)	9 (60%)	1 (7%)	15
Ambulatory Surgical Center (ASC) QR	0 (0%)	0 (0%)	0 (0%)	10 (71%)	2 (14%)	2 (14%)	0 (0%)	14
Home Health VBP (HHVBP) (original)	2 (25%)	0 (0%)	0 (0%)	5 (63%)	1 (12%)	0 (0%)	0 (0%)	8
Primary Care First (PCF) Model	0 (0%)	1 (14%)	2 (29%)	1 (14%)	1 (14%)	2 (29%)	0 (0%)	7
Hospital Acquired Condition (HAC) Reduction Program	1 (17%)	0 (0%)	0 (0%)	5 (83%)	0 (0%)	0 (0%)	0 (0%)	6
Hospital Readmission Reduction Program (HRRP)	0 (0%)	0 (0%)	0 (0%)	6 (100%)	0 (0%)	0 (0%)	0 (0%)	6

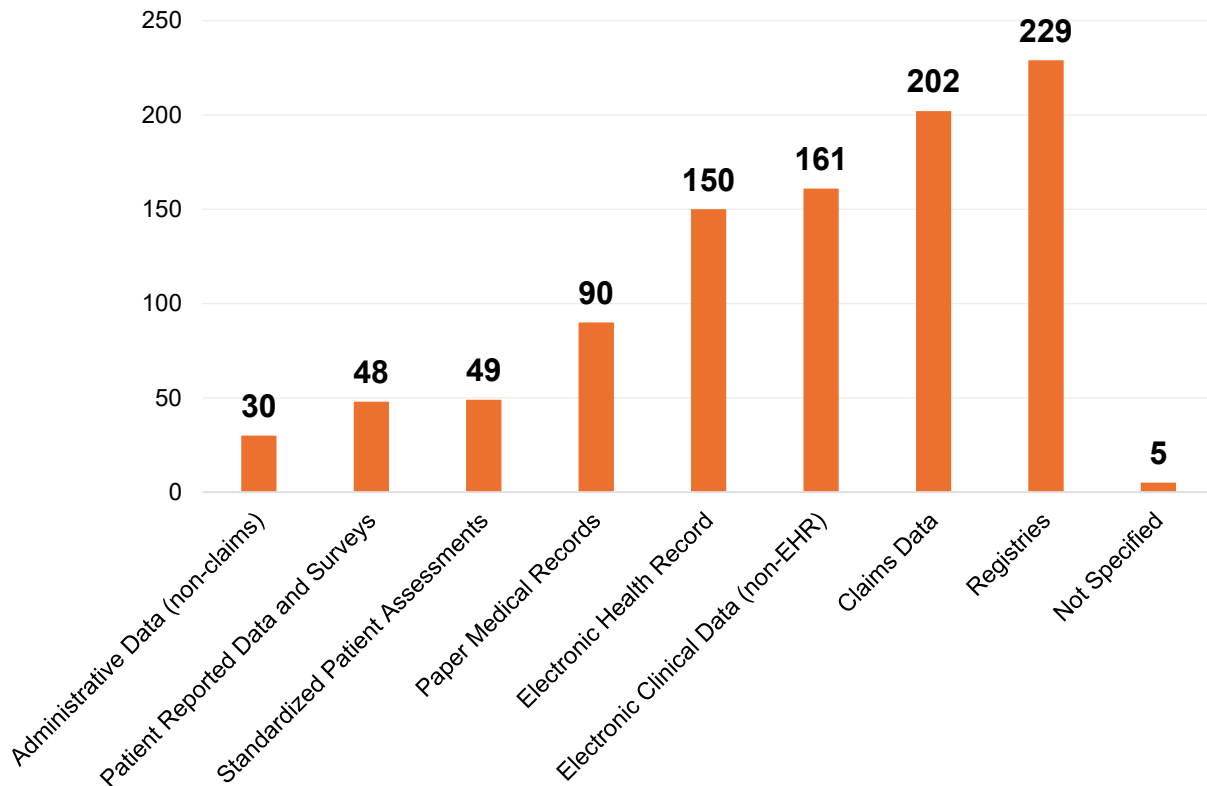
Program/ Model	Composite	Cost/ Resource Use	Intermediate Outcome	Outcome	Patient- Reported	Process	Structure	Total
Independence at Home (IAH) Demonstration	0 (0%)	0 (0%)	0 (0%)	3 (50%)	0 (0%)	3 (50%)	0 (0%)	6
Oncology Care Model (OCM)	0 (0%)	0 (0%)	0 (0%)	2 (33%)	1 (17%)	3 (50%)	0 (0%)	6
Accountable Care Organization (ACO) Realizing Equity, Access, and Community Health (REACH) Model	0 (0%)	0 (0%)	0 (0%)	3 (60%)	1 (20%)	1 (20%)	0 (0%)	5
Kidney Care Choices (KCC) Model	0 (0%)	2 (40%)	0 (0%)	0 (0%)	2 (40%)	1 (20%)	0 (0%)	5
Hospice QRP (HQRP)	2 (50%)	0 (0%)	0 (0%)	0 (0%)	1 (25%)	1 (25%)	0 (0%)	4
Skilled Nursing Facility (SNF) VBP	0 (0%)	0 (0%)	0 (0%)	3 (100%)	0 (0%)	0 (0%)	0 (0%)	3
Total	15 (2%)	35 (6%)	23 (4%)	163 (26%)	55 (9%)	323 (52%)	4 (1%)	618

Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Sources of Measures

The CMIT includes eight performance measure data sources: registries, claims data, electronic clinical data (non-EHR), electronic health records (EHR), paper medical records, standardized patient assessments, administrative data (non-claims), and patient-reported data and surveys. Of the 618 performance measures included among the 24 programs/models, data sources were spread across the eight different sources. Registry data are the most common performance measure data source used among the 24 models/programs, accounting for 24 percent of measures (n=229). Data sources are not specified in the CMIT for 1 percent (n=5) of performance measures. **Exhibit 13** provides the distribution of performance measures by data source. There often are multiple data sources used for a given performance measure; accordingly, Exhibit 13 reflects a total n=964 (as opposed to 618).

Exhibit 13. Distribution of Performance Measures by Data Sources for the 24 Programs/Models

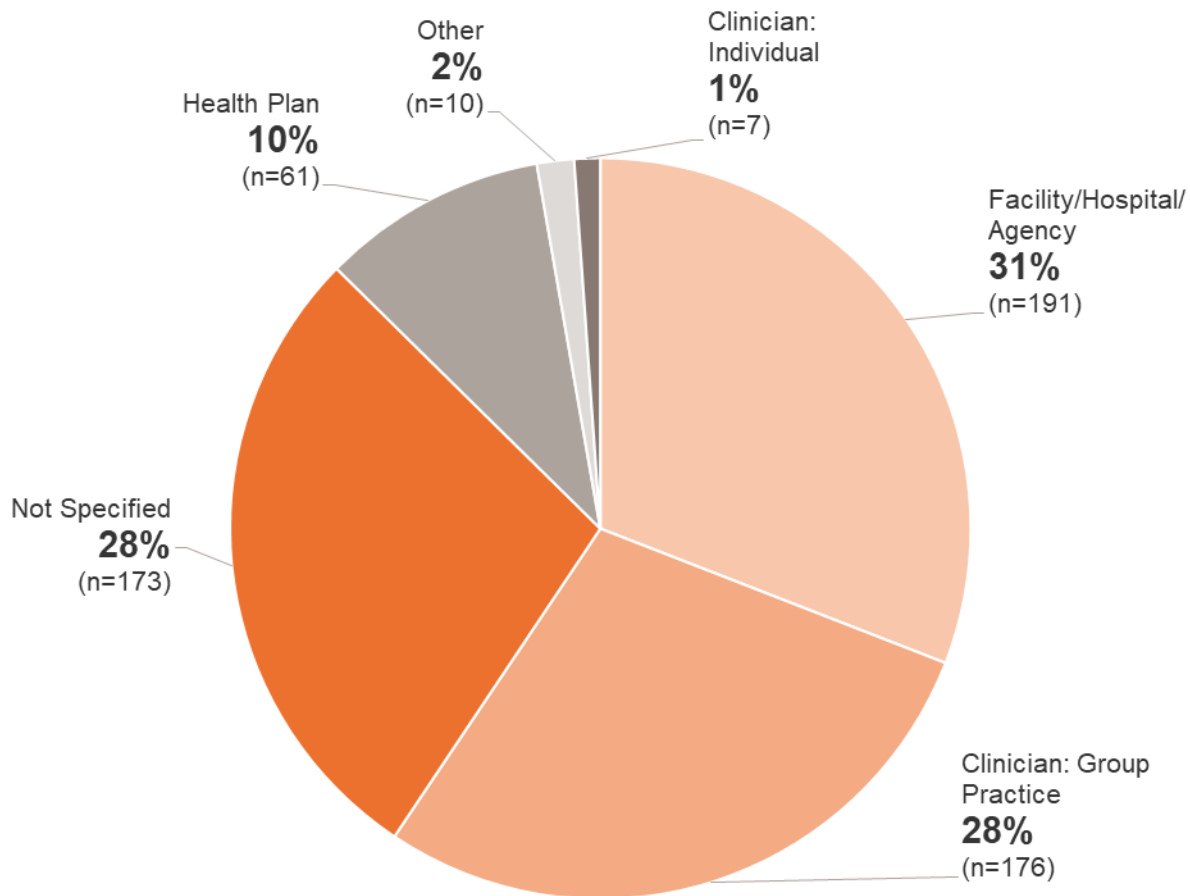


Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Measure Reporting Level

Of the 618 performance measures included among the 24 programs/models, 31 percent (n=191) of the measures are reported at the facility, hospital, or agency level; 28 percent (n=176) are reported at the clinician group practice level. About 28 percent (n=173) of performance measures do not specify level of reporting. **Exhibit 14** provides the distribution of performance measures by reporting level. Reporting level was created using the column “Level of Analysis” from the CMIT.

Exhibit 14. Distribution of Performance Measures by Reporting Level for the 24 Programs/Models



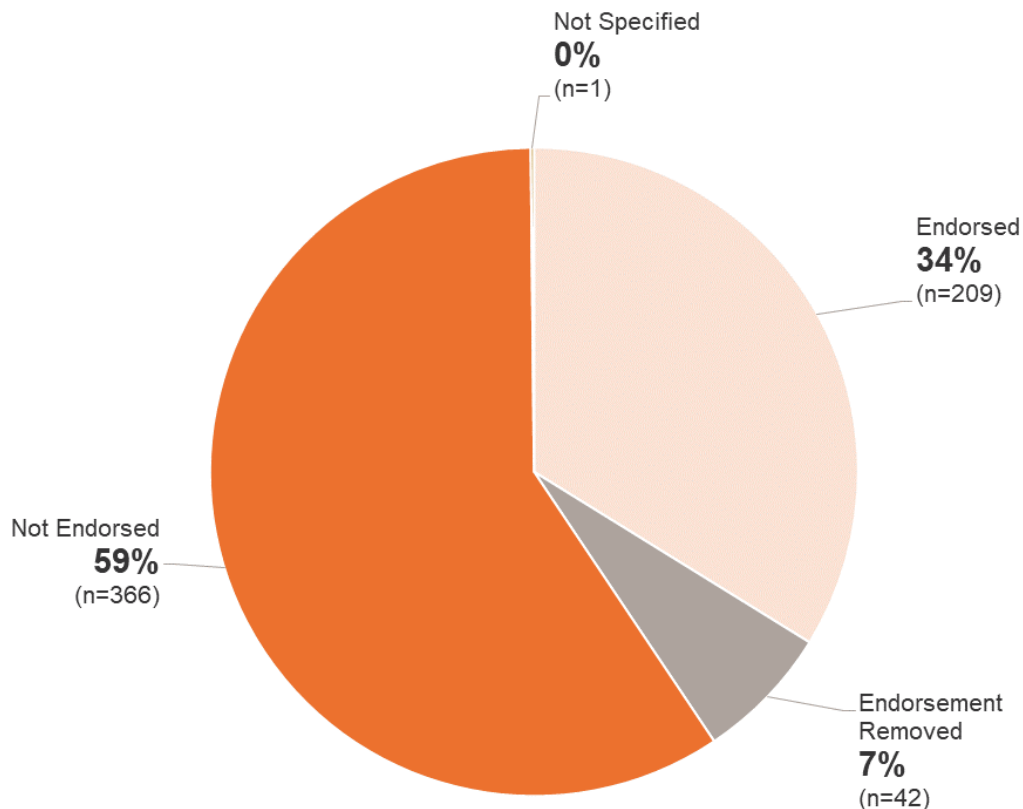
Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Performance Measure Endorsement Status

The CMIT includes information on whether the performance measure is endorsed by the CMS Consensus-Based Entity (CBE).^{xix} **Exhibit 15** provides the endorsement status for the 618 performance measures associated with the 24 programs/models. About 34 percent (n=209) of performance measures are endorsed, 59 percent (n=366) of measures are not endorsed, and endorsement has been removed for 7 percent (n=42). Twenty-three programs/models have at least one endorsed measure; the MA Star Ratings is the only program without any endorsed measures. For two programs – HACRP and HRRP – all active performance measures are endorsed (six measures each).

^{xix} Battelle’s PQM currently serves as the CMS CBE. See <https://mmshub.cms.gov/sites/default/files/Blueprint-CMS-CBE-Endorsement-Maintenance.pdf> and <https://p4qm.org/about> for more information on the CMS CBE process.

Exhibit 15. Distribution of Performance Measures by Endorsement Status for 24 Programs/Models



Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Whether Performance Measures are Tied to Payment

The 24 programs/models were categorized as pay-for-performance, pay-for-reporting, or not related to payment based on information from the CMS program and CMMI Innovation Models websites.^{xx} **Exhibit 16** provides the program/model type as it relates to payment for the 24 programs/models. Fifteen programs/models (63%)^{xxi} were characterized as pay-for-performance, defined as programs/models whose focus is on providing payment to providers based on outcomes of patients; providing better outcomes results in higher payments.^{xxii} Eight programs/models (33%)^{xxiii} were characterized as pay-for-reporting, defined as programs/models that are required to report quality measure data to CMS and result in a decrease to Medicare payments for nonperformance.^{xxiv} One program (4%), the PPS-Exempt CHQR, does not currently tie performance measures to payment.

^{xx} <https://www.cms.gov/priorities/innovation/models#views=models>

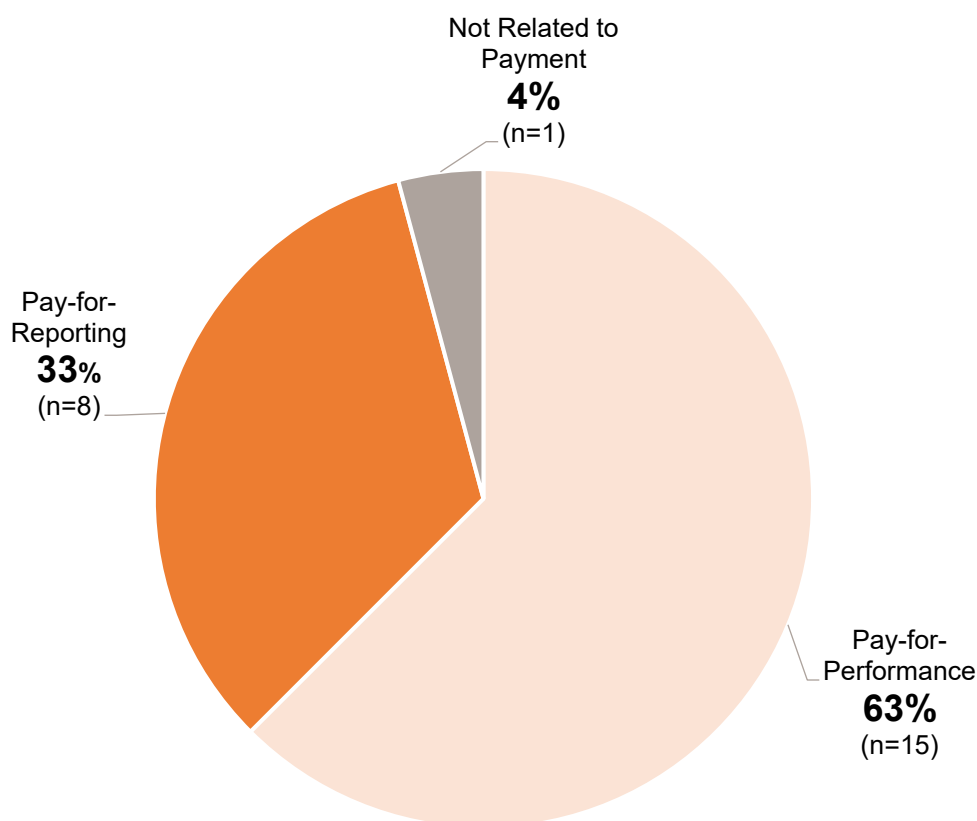
^{xxi} Pay-for-performance programs/models include ACO REACH, BPCI-A, ESRD QIP, HACRP, HHVBP, Hospital VBP, HRRP, IAH Demonstration, KCC Model, MA Star Ratings Program, MIPS, MSSP, OCM, PCF Model, and SNF VBP.

^{xxii} https://www.cms.gov/regulations-and-guidance/guidance/faca/downloads/tab_h.pdf

^{xxiii} Pay-for-reporting programs include ASCQR, Home Health QR, Hospital OQR, HQR, IPFQR, IRFQR, LTCH QR, and SNF QRP.

^{xxiv} <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Downloads/Pay-for-Reporting-Quality-Assessments-Only-Methodology.pdf>

Exhibit 16. Distribution of the 24 Programs/Models by Relationship to Provider Payment



The 15 pay-for-performance programs/models may also be characterized as another type of program, as defined on CMS program and CMMI Innovation Models websites. For example, four pay-for-performance programs/models were also characterized as disease-specific and episode-based programs/models: BPCI-A, ESRD QIP, HHVBP, and OCM. Also, four pay-for-performance programs/models use the term value-based purchasing to define their program/model: HACRP, Hospital VBP, HRRP, and SNF VBP. Three pay-for-performance models – ACO REACH, KCC, and PCF – are also considered Accountable Care Organizations (ACOs). The IAH Demonstration, a pay-for-performance model, is also defined as a statutory model. The PPS-Exempt CHQR Program, which does not currently tie performance measures to payment, is considered a quality reporting program (as opposed to pay-for-reporting). Exhibit 17 provides definitions for all program/model types.

Exhibit 17. Program/Model Type Definitions

Program/Model Type	Definition
Accountable Care	Models in which a doctor, group of health care providers or hospital takes financial responsibility for improving quality of care, including advanced primary care services, care coordination and health outcomes for a defined group of patients, thereby reducing uncoordinated care and unnecessary costs for patients and the health system. ^{xxv}

^{xxv} <https://www.cms.gov/priorities/innovation/models#views=models>

Program/Model Type	Definition
Disease-Specific and Episode-Based	Models which aim to address deficits in care for a defined population with a specific shared disease or medical condition, procedure, or care episode. ^{xxvi}
Pay-for-Performance	Programs/models whose focus is on providing payment to providers based on outcomes of patients; providing better outcomes results in higher payments. ^{xxvii}
Pay-for-Reporting	Programs/models that are required to report quality measure data to CMS and result in a decrease to Medicare payments for nonperformance. ^{xxviii}
Quality Reporting	Programs/models are required to report on certain quality measures; however, measures are not necessarily tied to payment.
Statutory Models	Models and demonstrations requiring testing as determined by Congress and/or the Secretary of Health and Human Services. ^{xxix}
Value-Based Purchasing	Programs/models that reward providers with incentive payments for the quality of care they provide to Medicare beneficiaries. ^{xxx}

NORC also examined the distribution of the 618 performance measures based on how the 24 programs/models are linked with payment (**Exhibit 18**).^{xxxi} Seventy-seven percent (n=476) of the measures correspond with the 15 pay-for-performance programs/models (50% [n=309] are MIPS performance measures, while the remaining 27% [n=167] correspond to the other 14 pay-for-performance programs/models). Twenty percent (n=126) of the measures correspond with the eight pay-for-reporting programs, and three percent of the measures correspond with the one program (PPS-Exempt CHQR) that is not linked with payment.

^{xxvi} <https://www.cms.gov/priorities/innovation/models#views=models>

^{xxvii} https://www.cms.gov/regulations-and-guidance/guidance/faca/downloads/tab_h.pdf

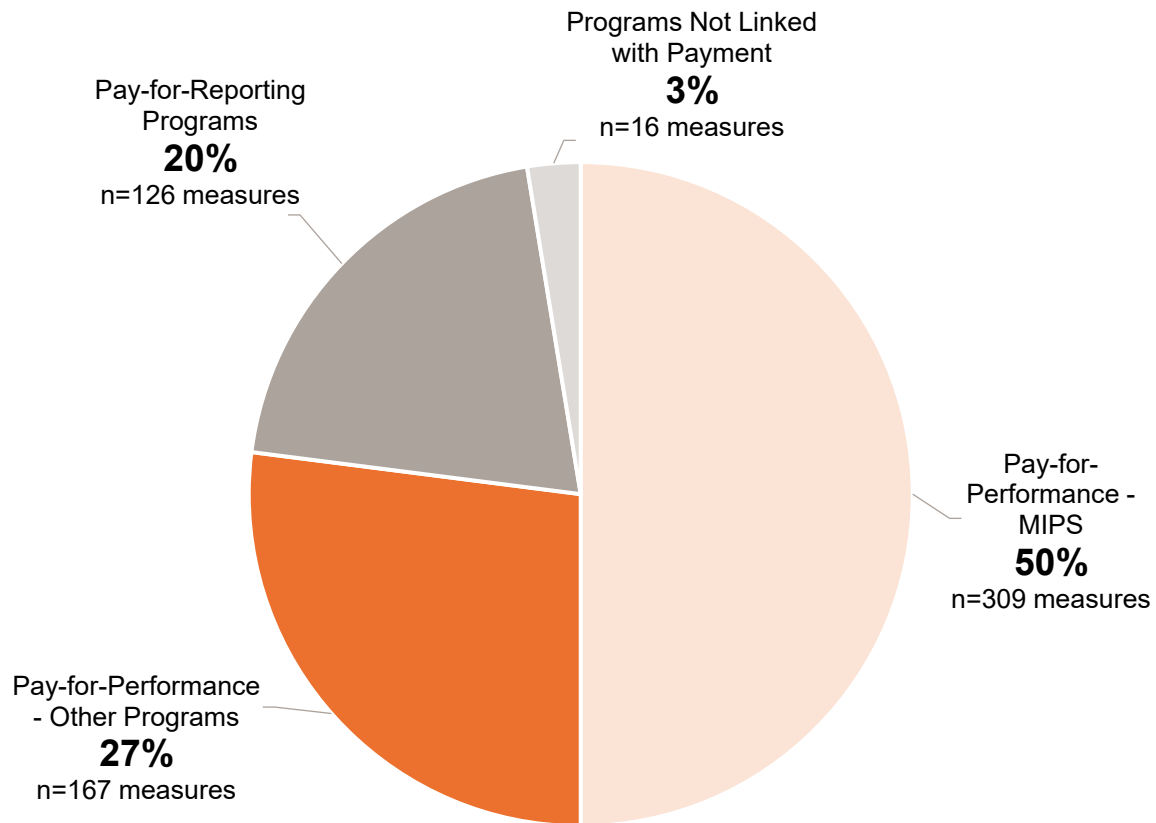
^{xxviii} <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Downloads/Pay-for-Reporting-Quality-Assessments-Only-Methodology.pdf>

^{xxix} <https://www.cms.gov/priorities/innovation/models#views=models>

^{xxx} <https://www.cms.gov/medicare/quality/value-based-programs>

^{xxxi} There are some limitations of this analysis, including 1) not all measures for a given program/model are necessarily tied to payment or required to be reported (e.g., some programs/models have many measures from which providers choose a set of measures); 2) measure-specific requirements can change frequently; and 3) measures may be used differently in different programs/models.

Exhibit 18. Distribution of Performance Measures Based on How the 24 Selected Programs/Models are Linked with Payment



Note: Performance measures include active, in-development, pending, and suspended measures listed in the CMIT as of October 2023. Further, the OCM is an inactive model; the six measures associated with the OCM are inactive measures.

Note: There are some limitations of this analysis, including 1) not all measures for a given program/model are necessarily tied to payment or required to be reported (e.g., some programs/models have many measures from which providers choose a set of measures); 2) measure-specific requirements can change frequently; and 3) measures may be used differently in different programs/models.

VII.B. Assessment of Performance Measures for Most Prevalent Chronic Conditions

Current performance measures reflect various chronic health conditions such as cardiovascular disease, cancer, diabetes, and hypertension. For example, measures related to cardiovascular disease include "Controlling High Blood Pressure," "Coronary Artery Disease (CAD) Management," and "Statin Therapy for Patients with Cardiovascular Disease." Performance measurement for chronic conditions presents unique difficulties due to shifting clinician roles in disease treatment over time and accounting for disease interactions, especially for individuals with multiple chronic conditions.^{193, 194}

Of the 618 measures, 183 (30 percent) measures are not associated with a specific condition. The remaining 435 measures associated with a specific condition belong to 17 programs and one model (**Exhibit 19**).^{xxxii}

^{xxxii} All 17 included programs and one model (Home Health VBP) include condition-specific measures. The six remaining models do not specify condition-specific measures.

Exhibit 19. The Number of Measures Associated with a Specific Condition- by Program/Model

Program/Model Name	Number of Condition-Specific Measures
Merit-based Incentive Payment System Program	277
End-Stage Renal Disease Quality Incentive Program	18
Medicare Shared Savings Program	15
Prospective Payment System-Exempt Cancer Hospital Quality Reporting	15
Inpatient Psychiatric Facility Quality Reporting	15
Medicare Part C Star Rating	14
Hospital Value-Based Purchasing	14
Skilled Nursing Facility Quality Reporting	12
Long-Term Care Hospital Quality Reporting	12
Hospital Outpatient Quality Reporting	10
Ambulatory Surgical Center Quality Reporting	8
Inpatient Rehabilitation Facility Quality Reporting	6
Hospital Readmission Reduction Program	6
Home Health Quality Reporting	5
Hospital Acquired Condition Reduction Program	5
Hospice Quality Reporting	1
Skilled Nursing Facility Value-Based Purchasing	1
Home Health Value-Based Purchasing	1
Total	435

Note: All data are as of October 2023, when CMIT data was pulled.

There are 59 different clinical conditions or subspecialties represented in the 18 programs/model that report condition-specific measures. **Exhibit 20** provides the top 15 most common conditions. Infection is the most common condition among these programs, tying to 15 percent of measures (n=67). Measures from MIPS are included for each of the top 15 clinical conditions or subspecialties.

Exhibit 20. The Top 15 Clinical Conditions/Subspecialties

Clinical Conditions	Number of Measures
Infection	67
Malignant Neoplasm (Cancer)	51
Cardiovascular Disease	51
Behavioral/Mental Health	50
Orthopedics	24
Renal	22
Eye Disease/Disorder	18
Pulmonary Disorder	16
Diabetes Mellitus	12
Pain	12
Unintentional Injury	11
Autoimmune Disorder	7
Neurological Disorder	7
Obstetrics/Gynecology	5
Gastroenterology	5

Note: All data are as of October 2023, when CMIT data was pulled.

VII.C. Measures Focused on Similar Aspects of Care

In addition to measures that are repeated across programs, some programs include performance measures that are distinct, but similar to other measures. **Exhibit 21** provides a summary of distinct measures focused on similar aspects of care. These groupings do not capture all performance measures but offer a look at common measures used among these 24 programs/models.

Exhibit 21. Performance Measure Groupings for Measures Focused on Similar Aspects of Care Across the 24 Programs/Models

Performance Measure Grouping	Number of Performance Measures	Percentage of Performance Measures (Total n=455)
Screening Measures	31	6.8%
Therapy-Related Measures for Certain Chronic Conditions	29	6.4%
Medication-Related Measures	21	4.6%
Measures Related to Number/Rate of Admissions/Visits	20	4.4%
Follow-up-Related Measures after Hospitalizations or ED Visits	15	3.3%
Measures Related to Readmissions	14	3.1%
Surgery-Related Measures	13	2.9%
Immunization-Related Measures	12	2.6%
Pain-Related Measures	11	2.4%
Measures Related to Infections	10	2.2%
Cost of Care Measures	7	1.5%
Measures Related to Mortality Rates	6	1.3%
Measures Related to Care Coordination	4	0.9%

Note: All data are as of October 2023, when CMIT data was pulled.

VIII. Performance Measures Used in CMMI Models

This section provides in-depth information from publicly available model documentation on 14 CMMI models that fit the following criteria:

- Includes at least one quality measure and one utilization and/or spending measure in implementation and/or monitoring;
- Is ongoing, under development, or completed within the last five years; and
- Is operational in more than one state.

For additional details on model years and current stage, see **Exhibit 22**.

Exhibit 22. CMMI Models that Include at Least One Quality Measure and One Utilization and/or Spending Measure in Implementation and/or Monitoring including Model, Model Years, and Model Stage

Model	Model Years	Model Stage
Accountable Care Organization Realizing Equity, Access, and Community Health (ACO REACH)*	2021- present	Participants announced
Bundled Payments for Care Improvement Advanced (BPCI-A)*	2018- present	Ongoing
Comprehensive ESRD Care (CEC) Model	2015-2021	No longer active
Comprehensive Primary Care Plus (CPC+)	2017-2021	No longer active
Enhancing Oncology Model (EOM)	2022- present	Ongoing
ESRD Treatment Choices (ETC) Model	2021- present	Ongoing
Expanded Home Health Value-Based Purchasing Model (Expanded HHVBP)	2022- present	Ongoing
Home Health Value-Based Purchasing (HHVBP) Model*	2016-2021	No longer active
Independence at Home (IAH) Demonstration*	2011- present	Ongoing
Kidney Care Choices (KCC) Model *	2020- present	Ongoing
Making Care Primary (MCP) Model	Launching in July 2024	Accepting applications
Next Generation Accountable Care Organization (NGACO)	2016-2021	No longer active
Oncology Care Model (OCM)*	2016-2022	No longer active
Primary Care First (PCF) Model Options*	2021- present	Ongoing

*These 7 models were also included in the high-level trends analysis in Section VII.

See **Appendix E** for more detailed information on the models’ clinical focus, providers, setting, and patient population; components relevant to performance measurement; technical issues related to performance measurement, including how performance is tied to payment; how measurement is used to determine success; gaps related to current performance measures; and lessons learned related to performance measurement.

VIII.A. Types of Performance Measures

Selected CMMI models included a variety of performance measures related to utilization, spending, quality, and patient experience. Utilization measures were identified in seven out of 14 models (CPC+, EOM, Expanded HHVBP, HHVBP, IAH, MCP, and PCF). Utilization measures were similar across models. Six models included measures of emergency department (ED) utilization (CPC+, EOM, Expanded HHVBP, HHVBP, IAH, and MCP), and six models included measures of acute hospital utilization (CPC+, EOM, Expanded HHVBP, HHVBP, IAH, and PCF).

Quality measures were included in all 14 models. Most measures were unique to a given model. However, some quality measures were seen across multiple models. For example, a hospital readmission measure was included in five models (BPCI-A, CEC, ACO REACH, IAH, and NGACO), and screening for depression and follow-up plan was included in four models (CEC, EOM, MCP, and OCM). In addition, advance care plan (BPCI-A, CEC, PCF), controlling high blood pressure (CPC+, MCP, PCF), and hemoglobin A1c (HbA1c) poor control (CPC+, MCP, PCF) were each included in three models.

Spending measures were included in all 14 models. Total Medicare Parts A and B spending was the most common spending measure, included in five models (CEC, CPC+, IAH, NGACO, OCM). Other spending

measures included total cost of care (BPCI-A, EOM, ACO REACH), total per capita cost (MCP, PCF), condition-specific spending (ETC, KCC), and setting-specific spending (Expanded HHVBP, HHVBP).

Patient experience measures were included in 11 out of 14 models. CAHPS measures were utilized most frequently, appearing in eight of 11 models (73 percent) that included a patient experience measure. MCP included a unique patient experience measure called the Person-Centered Primary Care Measure (PCPCM). Patient-reported outcomes were also incorporated in some models. For example, CEC included the kidney disease quality of life (KDQOL) survey. The IAH Demonstration, KCC Model, and NGACO Model did not include patient experience measures. For more details on each model and their patient experience measures, see **Exhibit 23** and **Appendix E**.

Exhibit 23. Patient-Reported Outcome (PRO) or Patient Experience of Care (PEC) Measures among CMMI Models that Include at Least One Quality Measure and One Utilization and/or Spending Measure

Model	CAHPS Measure	Other PRO or PEC Measure	No PRO or PEC Measure
Accountable Care Organization Realizing Equity, Access, and Community Health (ACO REACH)	X		
Bundled Payments for Care Improvement Advanced (BPCI-A)		X	
Comprehensive ESRD Care (CEC) Model	X		
Comprehensive Primary Care Plus (CPC+)	X		
Enhancing Oncology Model (EOM)	X		
ESRD Treatment Choices (ETC) Model			X
Expanded Home Health Value-Based Purchasing Model (Expanded HHVBP)	X		
Home Health Value-Based Purchasing (HHVBP) Model	X		
Independence at Home (IAH) Demonstration			X
Kidney Care Choices (KCC) Model			X
Making Care Primary (MCP) Model		X	
Next Generation Accountable Care Organization (NGACO)	X		
Oncology Care Model (OCM)		X	
Primary Care First (PCF) Model Options	X		

The performance measures generally capture performance relative to each model’s goals, ensuring that measures encapsulate key aspects of quality, efficiency, and equity. This allows for the performance measures to offer a comprehensive view of a model’s impact. Most programs seek to improve on quality, cost, and satisfaction goals, and their measures reflect these goals. Other models have more specific goals, typically related to a particular condition. For example, the OCM aims to provide higher quality, more coordinated oncology care at similar or lower costs than Medicare FFS. Therefore, it includes measures unique to the program such as “Oncology: Medical and Radiation – Plan of Care for Pain” and “Combination chemotherapy is recommended or administered within 4 months (120 days) of diagnosis.” Other CMMI models have additional goals related to patient engagement (e.g., EOM) or patient freedom of choice in their treatment plan (e.g., ESRD) and do not have measures that directly capture participants’ performance for these specific goals.

VIII.B. How Payment is Linked to Performance

Selected CMMI models use a variety of payment mechanisms to incentivize performance, including performance-based payments, performance-based recoupment, performance-based bonuses, performance incentive payments, performance payment adjustments, shared savings and/or losses, and reconciliation amounts based on quality scores. For more details on each model and their related payment mechanisms, see **Exhibit 24** and **Appendix E**.

Exhibit 24. How Payment is Adjusted for Performance among CMMI Models that Include at Least One Quality Measure and One Utilization and/or Spending Measure in Implementation and/or Monitoring

Name	How Payment is Adjusted for Performance
Accountable Care Organization Realizing Equity, Access, and Community Health (ACO REACH)	<ul style="list-style-type: none"> • Shared savings and losses. • 2% of ACO’s financial benchmark is held at risk, which can be earned back based on performance on four quality measures.
Bundled Payments for Care Improvement Advanced (BPCI-A)	<ul style="list-style-type: none"> • Reconciliation amounts based on quality scores (positive or negative, up to 10%).
Comprehensive ESRD Care (CEC) Model	<ul style="list-style-type: none"> • Shared savings and losses, multiplied by Total Quality Score (TQS).
Comprehensive Primary Care Plus (CPC+)	<ul style="list-style-type: none"> • Prospective performance-based incentive payments (PBIP). • Practices that do not meet the annual performance thresholds are “at risk” for repaying all or a portion of the PBIP.
Enhancing Oncology Model (EOM)	<ul style="list-style-type: none"> • Performance-based payment and performance-based recoupment based on quality and savings.
ESRD Treatment Choices (ETC) Model	<ul style="list-style-type: none"> • Home dialysis payment adjustment (HDP) and performance payment adjustment (PPA) to Medicare claim payments.
Expanded Home Health Value-Based Purchasing Model (Expanded HHVBP)	<ul style="list-style-type: none"> • Performance payment adjustment to FFS payments based on performance on quality measures.
Home Health Value-Based Purchasing (HHVBP) Model	<ul style="list-style-type: none"> • Performance payment adjustment (upward or downward, 3-7%) to Medicare payments based on quality performance and improvement.
Independence at Home (IAH) Demonstration	<ul style="list-style-type: none"> • Shared savings for meeting performance thresholds.
Kidney Care Choices (KCC) Model	<ul style="list-style-type: none"> • Performance-adjusted capitated payments based on quality measures. • Bonus payments for successful kidney transplants. • Shared savings and/or losses based on spending and quality measures, with different options including one-sided risk, 50% shared losses and savings, and 100% shared losses and savings.
Making Care Primary (MCP) Model	<ul style="list-style-type: none"> • Upside-only performance incentive payments.

Name	How Payment is Adjusted for Performance
Next Generation Accountable Care Organization (NGACO)	<ul style="list-style-type: none"> • Shared savings and losses. • Bonus payment for meeting quality requirements. • Quality “withhold” – a portion of ACO’s financial benchmark is held “at-risk” depending on the ACO’s quality score.
Oncology Care Model (OCM)	<ul style="list-style-type: none"> • Performance-based payment, multiplied by an Aggregate Quality Score (AQS).
Primary Care First (PCF) Model Options	<ul style="list-style-type: none"> • Performance-based payment (up to 50% increase or 10% decrease) based on quality performance and improvement.

In nearly all selected CMMI models, target prices are adjusted for risk. For example, five models (BPCI-A, CPC+, IAH, KCC, and PCF) adjust target prices based on hierarchical condition categories (HCC). Some models such as the ETC Model risk-adjust their performance benchmarks, which adjusts the transplant waitlist rate based on beneficiary age. Also, some measures within the MCP performance measure set are risk-adjusted, and the OCM uses a risk-adjusted ED visit and observation stay measure. However, the HHVBP Model does not include risk adjustment. This was updated in the release of the Expanded HHVBP Model, which allows for direct comparison of participants across varied patient populations.¹⁹⁵

Some models hold different participant/provider types or risk tracks to different performance standards, including the ACO REACH Model, PCF, CEC, and Expanded HHVBP. Under the ACO REACH Model, performance benchmarks differ for standard, new entrant, and high needs ACOs. For PCF, performance standards differ across four assigned risk groups. In addition, the CEC Model and Expanded HHVBP use different performance benchmarks for smaller- and larger-volume providers.

Participants are generally held responsible for a set of measures that focus on capturing quality, spending, patient experience, and utilization outcomes. Models rarely allow participants to select the measures for which they are responsible. The only exception was for the BPCI-A model, in which participants can choose two measures from the Administrative Quality Measures Set or three measures from the Alternate Quality Measures Set on which to be scored.

Several CMMI models incentivize performance improvement in addition to achievement of certain performance thresholds. Models incorporate a variety of measurement and incentive strategies, including calculating percent improvement relative to previous year(s), incorporating continuous improvement measures, calculating an improvement score, and utilizing improvement bonuses. Incentive structures can vary across models, depending on how each model aligns its incentives with its goals. One example is the ETC Model, which includes a Health Equity Incentive that provides additional improvement points to participants who show improvement in the home dialysis rate or transplant rate for their attributed dual-eligible or Low-Income Subsidy (LIS) beneficiaries.

VIII.C. Overlap in Current Performance Measures/Measure Definitions

There is substantial overlap in current performance measures across selected CMMI models, especially among utilization and patient experience measures. For example, around three-fourths of selected CMMI models used CAHPS measures of patient experience, and nearly half of selected CMMI models included measures of emergency department utilization and acute hospital utilization. There is more variation among quality measures across models, in part because several models focus on specific

conditions (e.g., oncology). See Section VIII.A. and **Appendix E** for more information about types of performance measures.

VIII.D. Gaps in Current Measures

In 2023, CMS established the Universal Foundation to streamline quality measures and identify gaps in current measures across CMS programs. Identified gaps currently include patient-safety measures for ambulatory settings, holistic well-being measures, and digitally collected outcome and patient-reported measures.¹⁹⁶ In the analysis of publicly available documentation for 14 selected CMMI models, a few evaluation reports identified gaps in existing measures. For example, HHVBP evaluators identified the need for health equity measures.¹⁹⁷ In addition, evaluators of OCM raised concerns about specificity of the validated measures for depression,¹⁹⁸ and PCF evaluators highlighted the lack of outcome measures compared with process measures.¹⁹⁹ For more details on each model and their related gaps in current measures, see **Appendix E**.

VIII.E. Lessons Learned Related to Developing and Implementing Performance Measures

Lessons learned from the implementation of CMMI models highlight several key approaches for performance measurement in population-based TCOC models, including:

- Additional resources may be needed to reduce administrative burden and costs associated with data collection.²⁰⁰
- Enhanced payments incentivize improvement in quality and spending measures.²⁰¹
- Performance measures and their associated benchmarks should be updated over time to capture relevant changes in diagnosis/procedure codes and to remove measures for which performance is topped out or measures that produce negative unintended consequences.²⁰²
- Measures that specifically target pre-existing inequities should be incorporated.²⁰³
- Incorporating annual wellness visits (AWVs) as a quality measure can help identify gaps in care through preventive health screenings and review of health-related social needs.²⁰⁴

Evaluation reports from CMMI models also note challenges with performance measurement that warrant further attention. In the IAH demonstration, participants struggled to meet site-reported measures, particularly timely follow-up and medication reconciliation. Additionally, PCF practices reported that their previous success in reducing preventable hospital utilization could make it difficult to achieve further reductions. Most models reported at least some improvement in quality of care; however, several models contributed to net increases in cumulative Medicare spending, including the NGACO model,²⁰⁵ OCM,²⁰⁶ and CPC+.²⁰⁷ In addition, social determinants of health, patient preferences, and other contextual factors played a role in influencing patient engagement and outcomes across various models. Lastly, evaluation reports identify the importance of guardrails to prevent unintended consequences within models. For example, the final HHVBP evaluation report identified persisting quality gaps by Medicaid status, race, and ethnicity. The report emphasized the need for targeted initiatives to prevent unintended consequences and to address pre-existing disparities.²⁰⁸

IX. Performance Measures Used in the Medicare Shared Savings Program

This section provides in-depth information from publicly available documentation on the MSSP, the first ACO model to be integrated into the Part A and Part B Medicare program. MSSP is an APM but is not considered a CMMI model.

See **Appendix F** for more details on the program.

IX.A. Types of Performance Measures

The MSSP uses quality and patient experience metrics to measure performance. There are several quality measures that are reported, including preventive care and screening rates for a variety of conditions, follow-up care, hospital-wide admission and readmission rates, diabetes hemoglobin A1c poor control (>9 percent), controlling high blood pressure, and care coordination.

Patient experience is measured through the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for the MIPS Clinician/Group Survey.

There is substantial crossover in quality measures used between MSSP and several CMMI models, including hospital readmission, screening for depression, and controlling high blood pressure. Like several CMMI models, MSSP incorporates Medicare Parts A and B spending as a spending measure. However, MSSP does not include performance measures related to utilization, whereas many CMMI models include measures such as emergency department utilization and hospital utilization.

IX.B. How Payment is Linked to Performance

ACOs that participate in MSSP are eligible to receive shared savings, or performance payments, if they reduce costs for the Medicare program while delivering high-quality care. At program inception, MSSP allowed only retrospective attribution; however, since 2019, MSSP has allowed ACOs to choose between retrospective or prospective attribution.²⁰⁹ ACOs may participate under the basic or enhanced track of MSSP. The basic track offers a one-sided risk model, and ACOs may increase the level of risk and reward each year. ACOs under the enhanced track are subject to the greatest level of risk and reward. ACOs must meet an annual spending benchmark; if they spend less than the benchmark, they share the savings with CMS. There is a penalty for spending more than the benchmark under the enhanced track. Further, at the end of every performance year, participating ACOs are required to report quality data to CMS. ACOs must meet quality thresholds and may have shared savings withheld if they do not meet these thresholds. ACOs earn quality points based on performance level for each measure. The higher the level of performance, the higher the number of quality points earned. The total points earned are summed and divided by the total points available to determine the ACO's quality score. The percentage of shared savings varies based on the ACO's quality score.

MSSP also incentivizes performance improvement through a Quality Improvement Reward. ACOs can earn up to four additional points in each quality domain if they show statistically significant improvement in their performance on quality measures from one year to the next.

IX.C. Gaps in Current Measures

In response to health equity concerns for non-white beneficiaries and those in rural and underserved communities, MSSP implemented a health equity adjustment to ACOs' quality performance category scores to reward high-quality care delivered to underserved communities.²¹⁰ Other gaps in current MSSP measures were not identified in publicly available MSSP documentation.

IX.D. Lessons Learned Related to Developing and Implementing Performance Measures

From 2013-2021, the MSSP resulted in net savings to MA and CMS overall, although the program was associated with net losses for the traditional Medicare program.²¹¹ Early performance results from CMS'

MSSP Pathways to Success suggest that ACOs with greater financial accountability (e.g., more accurate financial benchmarks, downside risk) are more likely to deliver more coordinated and efficient care for Medicare patients compared with ACOs with less financial accountability under MSSP.²¹² As of 2022, MSSP had generated overall savings and high-quality performance results for six consecutive years. In the same year, ACOs had a higher average performance on quality measures compared with other similarly sized clinician groups not in the program. Specifically, these ACOs had statistically significant higher performance for quality measures related to diabetes and blood pressure control, breast cancer and colorectal cancer screening, tobacco screening and smoking cessation, and depression screening and follow-up.²¹³ In 2022, low revenue ACOs, which primarily consist of physicians and may include a small hospital or serve rural areas, earned more shared savings than high-revenue ACOs. This result illustrates how well the Shared Savings Program supports primary care providers and underscores the importance of their participation in the program.²¹⁴

X. Performance Measures Used in the Medicare Advantage Star Ratings Program

This section provides in-depth information from publicly available documentation on the MA Star Ratings Program. The MA Star Ratings Program, developed by CMS, rates the quality of MA (Part C) and prescription drug (Part D) health plans, aiming to improve the quality of care delivered to Medicare beneficiaries. Like MSSP, the MA Star Ratings Program is not considered a CMMI model.

See **Appendix F** for more details on the program.

X.A. Types of Performance Measures

Similar to MSSP, the MA Star Ratings Program ties performance on quality and patient experience metrics to payment. Quality measures include care coordination, preventive screenings for chronic conditions, all-cause readmissions, transitions of care, chronic condition care management, and follow-ups for people with multiple high-risk chronic conditions.

The MA Star Ratings Program utilizes patient experience measures selected from HEDIS and CAHPS, including patient ability to get appointments and care quickly, patient experience with customer service, complaints about the plan, patient ratings of the health plan and health care quality, and the number of patients choosing to leave the plan.²¹⁵

The MA Star Ratings Program measures Medicare spending but does not include utilization measures. Many of the quality measures used in the MA Star Ratings Program are also included in CMMI models. Several of the patient experience measures can also be found in CMMI models. However, the metric tracking how many members choose to leave the plan is unique to the MA Star Ratings Program due to the program's unique focus on health plans rather than providers.

X.B. How Payment is Linked to Performance

Star Ratings are used to determine whether a health plan contract is eligible for a performance-based bonus payment and whether there will be a percentage increase in payment benchmarks and rebate amounts. Health plan contracts must obtain a 4-, 4.5-, or 5-Star Rating to be eligible for a performance-based bonus payment. To calculate the overall Star Ratings, MA health plan contracts receive a numeric measure score for each measure (up to 40 measures; number of measures differs by health plan

contract type). Measures come from four sources: HEDIS, CAHPS, the Health Outcomes Survey (HOS), and CMS administrative data. Measures are weighted to reflect CMS priority in scoring MA plans. Process measures receive a weight of 1; patient experience measures receive a weight of 1.5; outcomes and intermediate outcomes receive a weight of 3; and quality improvement measures receive a weight of 5. CMS calculates Star Ratings scores for each measure, as well as a summary Star Rating score based on their performance in five domains: use of screenings, tests, and vaccines; management of chronic conditions; member experience with plans (CAHPS); member complaints and changes in plan's performance; and customer service/appeals. The numeric measure scores are converted to a Star Rating (ranges from 1-5 stars with more stars indicating better performance) based on one of the following methods: clustering or relative distribution and significance testing. The distribution of health plan contracts in each bucket does not change substantially over time. For example, if a 3-star health plan improves, but all other health plans improve as well, the 3-star health plan will not increase to a higher rating.

The MA Star Ratings Program also measures improvement by comparing the health plan contract's current and prior year measure scores; however, improvement is not tied to payment.

X.C. Gaps in Current Measures

According to a report from McKinsey, the MA Star Ratings Program has historically had low investment in patient experience improvements.²¹⁶ According to its analysis, McKinsey attributed this lack of improvement to health insurers not engaging as effectively as they could with providers.²¹⁷ Payers could invest more in provider engagement by creating incentive programs to close care gaps, host roundtables for best practices to be shared by physicians, or develop materials to encourage performance improvements on program measures.

X.D. Lessons Learned Related to Developing and Implementing Performance Measures

The McKinsey report found two key areas of improvement for the MA Star Ratings Program.²¹⁸ First, outcome and process measures need to evolve as cut points change over time, which will require long-term investments from health plan contracts, as well as accountability. Second, patient experience and access scores have seen less improvement overall than other domains and would benefit from focused efforts to improve in this area from health plan contracts. The report also found that improvements in digitization and analytics could help health plans engage beneficiaries and providers and decrease administrative burden.

Additionally, the Urban Institute found that there were many limitations to the Star Rating system and quality bonus program.²¹⁹ First, the program offers generous bonuses and no downside risk or penalties for programs with low Star Ratings. Second, there are limitations in underlying data sets which have led to measures that are focused on the needs of younger and healthier beneficiaries (e.g., prevention-focused measures, use of process rather than outcome measures). Finally, performance is measured at the contract level instead of the plan level, therefore, potentially masking plans with low Star Ratings. This report also found that many beneficiaries do not use Star Ratings when selecting their plans, limiting the usefulness of the MA Star Ratings Program.

PB-TCOC models can also draw on innovative approaches related to performance measures, performance measurement, and financial incentives from other programs, such as Medicaid 1115 waiver programs, commercial/employer coverage, and Marketplace plans. For example, several state

Medicaid 1115 waiver programs have implemented innovative social determinants-related measures that are tied to payment.²²⁰ In addition, some commercial/employer initiatives have reported improved outcomes using global budget models that require providers to meet quality thresholds to be eligible for shared savings.²²¹ Covered California, a Marketplace health plan, implemented a Quality Transformation Initiative that ties substantial financial consequences to a select few clinical measures and is guided by core principles including “less is more” and “equity is quality.”²²²

XI. Performance Measures Used in PTAC Proposals

This section summarizes findings from an analysis of the 34 proposals that were submitted to PTAC between 2016 and 2020.^{xxxiii} Proposals were grouped into four categories based on Criterion 2 (Quality and Cost: ability to maintain/improve health care quality while decreasing costs or improve health care quality at no additional cost) and Criterion 4 (Value over Volume: provision of incentives for delivery of high-quality health care). The four categories of PTAC proposals for the purposes of this analysis included the following:

1. Proposals that met both Criterion 2 and Criterion 4 (N=19)
2. Proposals that met Criterion 2 but did not meet Criterion 4 (N=5)
3. Proposals that met Criterion 4 but did not meet Criterion 2 (N=1)
4. Proposals that did not meet Criterion 2 or Criterion 4, were withdrawn, or determined to be out of scope by PTAC (N=9)

For additional details on included proposals and which criterion they met, see **Exhibit 25**.

Exhibit 25. PTAC Proposals including Proposal Name, Submitter Name, and Criterion Met

Proposal Name	Submitter Name	Met Criterion 2	Met Criterion 4
Advanced Primary Care: A Foundational Alternative Payment Model (APC-APM) for Delivering Patient-Centered, Longitudinal, and Coordinated Care	American Academy of Family Physicians (AAFP)	X	X
Patient and Caregiver Support for Serious Illness	American Academy of Hospice and Palliative Medicine (AAHPM)		X
The Patient-Centered Headache Care Payment (PCHCP)	American Academy of Neurology (AAN)		
Patient-Centered Asthma Care Payment (PCACP): An Alternative Payment Model for Patient-Centered Asthma Care	American College of Allergy, Asthma & Immunology (ACAAI)	X	
Acute Unscheduled Care Model (AUCM): Enhancing Appropriate Admissions	American College of Emergency Physicians (ACEP)	X	X

^{xxxiii} Although 35 proposals were submitted to PTAC, one proposal submitted by the American Association of Hip and Knee Surgeons (AAHKS) was excluded from analysis because PTAC received a letter of intent only.

Proposal Name	Submitter Name	Met Criterion 2	Met Criterion 4
The “Medical Neighborhood” Advanced Alternative Payment Model (AAPM) (Revised Version)	American College of Physicians-National Committee for Quality Assurance (ACP-NCQA)	X	X
The ACS-Brandeis Advanced APM	The American College of Surgeons (ACS)	X	X
Patient-Centered Oncology Payment Model (PCOP)	American Society of Clinical Oncology (ASCO)		X
Intensive Care Management in Skilled Nursing Facility Alternative Payment Model (ICM SNF APM)	Avera Health (Avera Health)	X	X
Bundled PCI Services	Clearwater Cardiovascular and Interventional Consultants, MD, PA (CCC)		
Oncology Care Model 2.0	Community Oncology Alliance (COA)		
Advanced Care Model (ACM) Service Delivery and Advanced Alternative Payment Model	Coalition to Transform Advanced Care (C-TAC)	X	X
Comprehensive Colonoscopy Advanced Alternative Payment Model for Colorectal Cancer Screening, Diagnosis and Surveillance	Digestive Health Network, Inc. (DHN)		
Alternative Payment Model for Improved Quality and Cost in Providing Home Hemodialysis to Geriatric Patients Residing in Skilled Nursing Facilities	Dialyze Direct (Dialyze Direct)		X
An Innovative Model for Primary Care Office Payment	Jean Antonucci, MD (Dr. Antonucci)		X
Medicare 3 Year Value Based Payment Plan (Medicare 3VBPP)	Zhou Yang, PhD, MHP (Dr. Yang)		
Oncology Bundled Payment Program Using CNA-Guided Care	Hackensack Meridian Health and Cota, Inc. (HMH/Cota)	X	X
Community Aging in Place – Advancing Better Living for Elders (CAPABLE) Provider-Focused Payment Model	Johns Hopkins School of Nursing and the Stanford Clinical Excellence Research Center (Hopkins/Stanford)	X	X
Project Sonar	Illinois Gastroenterology Group and SonarMD, LLC (IGG/SonarMD)	X	X
Making Accountable Sustainable Oncology Networks (MASON)	Innovative Oncology Business Solutions, Inc. (IOBS)	X	X

Proposal Name	Submitter Name	Met Criterion 2	Met Criterion 4
LUGPA APM for Initial Therapy of Newly Diagnosed Patients with Organ-Confined Prostate Cancer	Large Urology Group Practice Association (LUGPA)	X	X
A Single Bundled Payment for Comprehensive Low-Risk Maternity and Newborn Care Provided by Independent Midwife Led Birth Center Practices that Are Clinically Integrated with Physician and Hospital Services	Minnesota Birth Center (MBC)		
Annual Wellness Visit Billing at Rural Health Clinics	Mercy Accountable Care Organization (Mercy ACO)		
HaH Plus (Hospital at Home Plus) Provider-Focused Payment Model	Icahn School of Medicine at Mount Sinai (Mount Sinai)	X	X
Multi-Payer, Bundled Episode-of-Care Payment Model for Treatment of Chronic Hepatitis C Virus (HCV) Using Care Coordination by Employed Physicians in Hospital Outpatient Clinics	New York City Department of Health and Mental Hygiene (NYC DOHMH)	X	X
The COPD and Asthma Monitoring Project	Pulmonary Medicine, Infectious Disease and Critical Care Consultants Medical Group (PMA)	X	X
Home Hospitalization: An Alternative Payment Model for Delivering Acute Care in the Home	Personalized Recovery Care (PRC)	X	X
Incident ESRD Clinical Episode Payment Model	Renal Physicians Association (RPA)	X	X
Bundled Payment for All Inclusive Outpatient Wound Care Services in Non-Hospital-Based Setting	Seha Medical and Wound Care (Seha)		
Remote Specialists and Experts on Demand Improving Care and Saving Costs (Revised Version)	Eitan Sobel, MD (Sobel)		
Comprehensive Care Physician Payment Model	University of Chicago Medicine (UChicago)	X	X
Eye Care Emergency Department Avoidance (EyEDA) Model	The University of Massachusetts Medical School (UMass)	X	X
ACCESS Telemedicine: An Alternative Healthcare Delivery Model for Rural Cerebral Emergencies	The University of New Mexico Health Sciences Center (UNMHSC)	X	X

Proposal Name	Submitter Name	Met Criterion 2	Met Criterion 4
CMS Support of Wound Care in Private Outpatient Therapy Clinics: Measuring the Effectiveness of Physical or Occupational Therapy Intervention as the Primary Means of Managing Wounds in Medicare Recipients	Upstream Rehabilitation (Upstream)		X

See **Appendix G** for more detailed information on the proposal.

XI.A. Types of Performance Measures

Proposals included a variety of performance measures related to utilization, spending, quality, and patient experience. Of the 34 proposals, 28 included utilization measures (82 percent), 32 included quality measures (94 percent), 31 included spending measures (91 percent), and 31 included patient experience measures (91 percent). Analyses specific to each category are discussed below.

Proposals that Met Criterion 2 and Criterion 4

Of the 19 proposals that met Criterion 2 and Criterion 4, 15 (79 percent) included a utilization measure. Utilization measures were similar across proposals and were often measures of hospital utilization or ED utilization.

Quality measures were included in all 19 proposals. Almost all quality measures were unique to a given proposal and its goals, as many proposals address a specific condition. However, specific quality measures were used in more than one proposal in a few instances. Advance care planning was used in two proposals (C-TAC and RPA). Cervical cancer screenings were included in two proposals (AAFP and ACP-NCQA), and tobacco screenings were included in two proposals (AAFP and ACS).

Spending measures were included in 17 of the 19 proposals (89 percent). Total cost of care was the most common spending measure, included in seven proposals (C-TAC, HMH/Cota, IOBS, PRC, UChicago, UMass, and UNMHSC). The second most common spending measure was Medicare Parts A and B spending (LUGPA, PMA, and RPA).

Patient experience measures were included in 17 of the 19 proposals (89 percent). Seven proposals (37 percent) used measures from patient satisfaction surveys (Avera Health, C-TAC, IOBS, PMA, PRC, UMass, and UNMHSC). CAHPS measures appeared in six proposals (32 percent) (AAFP, ACP-NCQA, ACS, Hopkins/Stanford, Mount Sinai, and UChicago). Two proposals mentioned the Patient-Reported Outcomes Measurement Information System (RPA, PRC); one had its own patient survey to capture patient experience measures (IGG/SonarMD); one cited the prostate cancer shared decision-making process (LUGPA); and one included patient-reported outcomes and national guidelines concerning pain management (HMH/Cota).

Proposals that Met Criterion 2 but Did Not Meet Criterion 4

Only one proposal (ACAAI) aligned with this category. The ACAAI proposal included measures for all four measure types.

Proposals that Did Not Meet Criterion 2 but Met Criterion 4

Of the five proposals that did not meet Criterion 2 but met Criterion 4, four (80 percent) included a utilization measure. Three proposals included measures of hospital utilization or ED utilization (ASCO, Dialyze Direct, and Upstream).

Four of the five proposals (80 percent) included quality measures. Again, quality measures were often unique to a given proposal and its goals. Only one measure, advance care planning, appeared in more than one proposal (AAHPM and ASCO).

Spending measures were included in all five proposals, with three proposals using total cost of care as their spending measure. One proposal included Medicare Part A and Part B spending, and one proposal included per beneficiary per month (PBPM) payments as its spending measure.

Patient experience measures were included in all five proposals. Three proposals used patient satisfaction surveys (AAHPM, ASCO, and Upstream). One proposal included CAHPS measures (Dialyze Direct).

The proposal submitted by Dr. Antonucci did not include utilization or quality measures; it included one patient experience measure, which was to utilize the How's Your Health survey.

Proposals that Did Not Meet Criterion 2 or Criterion 4, Were Withdrawn, or Determined to be Out of Scope by PTAC

Of the nine proposals that did not meet Criterion 2 or Criterion 4, were withdrawn, or determined to be out of scope, eight (89 percent) included a utilization measure. Four proposals included measures of hospital utilization or ED utilization (AAN, COA, DHN, and Dr. Yang). Mercy ACO was the only proposal to not include a utilization measure.

Eight of the nine proposals (89 percent) included quality measures. Although many quality measures were unique to a given proposal and their goals, depression screenings were included in four proposals (AAN, COA, DHN, Mercy ACO). Preventive care screenings for alcohol use were mentioned in two proposals (AAN and DHN). The DHN proposal applied the MIPS quality measures; the MBC proposal mentioned use of "outcome-level maternity quality measures" but did not provide specific measures. The proposal submitted by Dr. Sobel did not include any quality measures.

Spending measures were included in eight of the nine proposals (89 percent). Episode- or condition-specific cost of care accounted for four proposals (AAN, CCC, DHN, Seha), and total cost of care measures accounted for three proposals (COA, Mercy ACO, Sobel). The MBC proposal did not include any spending measures.

Patient experience measures were included in eight of the nine proposals (89 percent). Five proposals used patient satisfaction surveys (CCC, DHN, Dr. Yang, Mercy ACO, and Sobel). Others mentioned specific surveys, such as the printed post-partum survey (MBC) or the oncology medical home survey (COA). The AAN proposal included the percent of patients rating access to providers and experience of care as "excellent." The Seha proposal did not include any patient experience measures.

Most proposed models aimed to achieve a balance between total cost of care, volume of patients seen, and patient experience. The performance measures included in the proposed models covered various domains, such as neurology, ophthalmology, oncology, and endocrinology.

XI.B. How Payment is Linked to Performance

The PTAC proposals included a variety of approaches to link payment to performance. There were differences among the proposal categories on whether payment was linked to performance. Most proposals in the first three categories included approaches that linked payment to performance (all but two [Dialyze Direct and UNMHSC] out of 25 proposals). In contrast, among nine proposals that did not meet Criterion 2 or Criterion 4, were withdrawn, or determined to be out of scope, five proposals included approaches that tie payment to performance (CCC, Dr. Yang, MBC did not link payment to performance).

For more details on each proposal and their related payment mechanisms, see **Exhibit 26** and **Appendix G**.

Exhibit 26. How Payment is Adjusted for Performance among PTAC Proposals

PTAC Proposal	How Payment is Adjusted for Performance
Proposals that Met Criterion 2 and Criterion 4	
American Academy of Family Physicians (AAFP)	<ul style="list-style-type: none"> Provider repays incentive payments if they do not meet performance benchmarks.
American College of Emergency Physicians (ACEP)	<ul style="list-style-type: none"> Performance on a set of quality measures determines eligibility for reconciliation payments and the size of discount built into each episode’s target price.
American College of Physicians-National Committee for Quality Assurance (ACP-NCQA)	<ul style="list-style-type: none"> Retrospective positive or negative payment adjustments made based on performance on financial benchmarks.
The American College of Surgeons (ACS)	<ul style="list-style-type: none"> Payment is adjusted based on quality measures, incorporating two-sided risk.
Avera Health (Avera Health)	<ul style="list-style-type: none"> Option 1: Payment adjustments based on performance on quality metrics (0%, 50%, or 100% of payment). Option 2: Shared savings only.
Coalition to Transform Advanced Care (C-TAC)	<ul style="list-style-type: none"> Quality bonus funded by shared savings. Downside risk beginning in year 3.
Hackensack Meridian Health and Cota, Inc. (HMH/Cota)	<ul style="list-style-type: none"> Upside only: Physicians will receive higher bundle compensation if performance metrics are achieved.
Johns Hopkins School of Nursing and the Stanford Clinical Excellence Research Center (Hopkins/Stanford)	<ul style="list-style-type: none"> Partial bundled payment with partial upside risk.
Illinois Gastroenterology Group and SonarMD, LLC (IGG/SonarMD)	<ul style="list-style-type: none"> Payment adjustments are based on quality and financial performance, including shared savings and losses.

PTAC Proposal	How Payment is Adjusted for Performance
Innovative Oncology Business Solutions, Inc. (IOBS)	<ul style="list-style-type: none"> Provider receives shared savings if quality parameters are met.
Large Urology Group Practice Association (LUGPA)	<ul style="list-style-type: none"> Participants earn performance-based payments or owe performance-based repayments based on the number of quality performance targets achieved/exceeded.
Icahn School of Medicine at Mount Sinai (Mount Sinai)	<ul style="list-style-type: none"> Shared savings and losses based on performance.
New York City Department of Health and Mental Hygiene (NYC DOHMH)	<ul style="list-style-type: none"> Shared savings (and an annual bonus) and shared losses based on performance on the HCV SVR benchmark.
Pulmonary Medicine, Infectious Disease and Critical Care Consultants Medical Group (PMA)	<ul style="list-style-type: none"> Two-sided risk arrangement with shared savings and losses based on performance.
Personalized Recovery Care (PRC)	<ul style="list-style-type: none"> Shared savings; amount based on performance on five performance metrics (20% of savings per metric).
Renal Physicians Association (RPA)	<ul style="list-style-type: none"> Quality scores determine physician’s eligibility and amount of shared savings. Physicians can choose to participate in downside risk. One-time financial incentive/bonus payment for patient receiving a kidney transplantation.
University of Chicago Medicine (UChicago)	<ul style="list-style-type: none"> Care continuity fee given to providers who meet benchmarks for providing their patients with both inpatient and outpatient care. Providers continue to be subject to financial incentives/penalties under their current model (e.g., MIPS, MSSP).
The University of Massachusetts Medical School (UMass)	<ul style="list-style-type: none"> Shared savings based on performance on ED-avoidable visits and other quality performance. If providers do not meet performance thresholds, their financial loss will equal the minimum of 8% of performance year payments.
The University of New Mexico Health Sciences Center (UNMHSC)	<ul style="list-style-type: none"> Performance measures are not linked to payment.
Proposals that Met Criterion 2 and Did Not Meet Criterion 4	
American College of Allergy, Asthma & Immunology (ACAAI)	<ul style="list-style-type: none"> Provider receives default payment if they achieve at least “good” on all performance measures. Payments are increased or decreased (up to 5-9%) if team scores “high” or “low” on some performance measures.
Proposals that Did Not Meet Criterion 2 and Met Criterion 4	
American Academy of Hospice and Palliative Medicine (AAHPM)	<ul style="list-style-type: none"> Track 1: Positive and negative performance incentives of up to 4% of total care management fees based on quality and spending performance. Track 2: Shared savings and losses based on total cost of care (beginning in year 3).
American Society of Clinical Oncology (ASCO)	<ul style="list-style-type: none"> Performance incentive payments will be positively or negatively adjusted based on provider success in adherence to clinical treatment pathways, quality metrics, and cost reduction.

PTAC Proposal	How Payment is Adjusted for Performance
Dialyze Direct (Dialyze Direct)	<ul style="list-style-type: none"> • Performance measures are not linked to payment.
Jean Antonucci, MD (Dr. Antonucci)	<ul style="list-style-type: none"> • 15% of annual income will be withheld; if participants do not meet quality and cost benchmarks, they may lose this income.
Upstream Rehabilitation (Upstream)	<ul style="list-style-type: none"> • Claim refunded to CMS if minimum standards of improvement are not met. • Clinicians can receive a 3% savings bonus for achieving average reimbursement costs below risk-adjusted thresholds.
Proposals that Did Not Meet Criterion 2 or Criterion 4, were Withdrawn, or Determined to be Out of Scope	
American Academy of Neurology (AAN)	<ul style="list-style-type: none"> • Provider receives default payment if they achieve at least “good” on all performance measures. • Payments are increased or decreased (up to 4%) if team scores “high” or “low” on some performance measures.
Clearwater Cardiovascular and Interventional Consultants, MD, PA (CCC)	<ul style="list-style-type: none"> • Performance measures are not linked to payment.
Community Oncology Alliance (COA)	<ul style="list-style-type: none"> • Episode-based payment with shared risk based on quality and spending performance.
Digestive Health Network, Inc. (DHN)	<ul style="list-style-type: none"> • Prospective episode-based model with retrospective reconciliation; payments adjusted based on the quality of care delivered.
Zhou Yang, PhD, MHP (Dr. Yang)	<ul style="list-style-type: none"> • Performance measures are not linked to payment.
Minnesota Birth Center (MBC)	<ul style="list-style-type: none"> • Performance measures are not linked to payment.
Mercy Accountable Care Organization (Mercy ACO)	<ul style="list-style-type: none"> • 10% withhold in provider compensation if less than 40% of attributed Medicare patients complete an annual wellness visit or if patient satisfaction standards are not met.
Seha Medical and Wound Care (Seha)	<ul style="list-style-type: none"> • The proposal mentions that the model will create incentives for best outcomes but does not specify details.
Dr. Sobel (Sobel)	<ul style="list-style-type: none"> • Quality, outcomes, and patient satisfaction affect fee schedule and future participation eligibility.

Of the 34 proposals, 26 (76 percent) mentioned establishing target prices or benchmarks for comparison. Most proposals indicated that historical performance data would be used to set benchmarks. Eight proposals (24 percent) did not specify if or how benchmarks would be established (Hopkins/Stanford, IGG/SonarMD, UNMHSC, AAN, DHN, Dr. Yang, Mercy ACO, and Sobel); these were dispersed across the four proposal categories.

Twenty-six proposals (76 percent) also included risk adjustment methodologies to adjust target prices based on patients’ risk. For example, six proposals (ACEP, ACP-NCQA, Avera Health, LUGPA, RPA, and CCC) planned to adjust target prices based on hierarchical condition categories (HCCs); five of these six were proposals that met both Criterion 2 and Criterion 4. Eight of the 34 proposals (24 percent) did not apply risk adjustment (Hopkins/Stanford, UChicago, UMass, UNMHSC, Dialyze Direct, Mercy ACO, Seha, and Sobel); these were distributed across the four proposal categories.

Proposals generally recommended a set of utilization, quality, spending, and/or patient experience measures aimed at achieving the specific goals of the model proposed.

XI.C. Overlap in Current Performance Measures/Measure Definitions

Similar to the selected CMMI models, the analysis reveals substantial overlap in current performance measures across PTAC proposals, especially among utilization, spending, and patient experience measures. There is more variation among quality measures across proposals, as many proposals are disease- or condition-specific. However, a few quality measures did appear across multiple proposals, such as preventive screenings and advanced care planning. See **Appendix G** for more information about types of performance measures.

XII. How Performance Measures are Linked with Payment in Other Programs

In addition to the models and programs described in previous sections, there are other examples of existing performance measures and measure sets used to determine performance-based financial incentives. This section describes evidence on the effectiveness of performance measures in programs and potential unintended consequences of performance measurement.

XII.A. Evidence on the Effectiveness of Performance-based Payment Programs

Evidence of the effectiveness of PBP programs is mixed across measure type, care setting, and patient condition. There has been a wealth of studies on PBP programs that vary in methodological rigor and in the detail provided about the measures included in PBPs or the measures used to determine the effectiveness of the program. The findings presented in this section provide insight across several important characteristics of performance measures, including measure type, clinical setting, level of accountability, risk arrangement, magnitude of payment, and benchmarking process.

Measure type. Evidence on the effectiveness of PBP varies by the type of measure tied to financial incentives.

The effectiveness of PBP varies for quality measures. Key findings for PBP programs that use quality measures include:

- No impact on composite high-value or low-value care.²²³
- Slight improvement in use of selected tests and treatments in outpatient facilities.²²⁴
- Reduced antibiotic prescriptions in outpatient facilities²²⁵ and among adults with acute bronchitis.²²⁶
- Increased coordination of care.²²⁷
- Increased provider confidence in providing high-quality care.²²⁸
- Improved interaction among office staff.²²⁹
- Increased number of patients asked details about their disease by a pharmacist.²³⁰
- Improved child immunization status.²³¹
- Increased breast cancer screening.²³²
- Increased cervical cancer screening.²³³
- Increased screening for tobacco and use and provision of smoking cessation interventions.²³⁴
- Increased vaccination rates.²³⁵
- Mixed results on prescriptions for guideline-recommended antihypertensive medications.^{236,237}

- Wide variation in testing rates by patient.²³⁸
- Increased broad-spectrum antibiotic use for sepsis, but potential overuse.²³⁹
- Increased lactate measurement and fluid resuscitation for sepsis.²⁴⁰
- Lower rate of cardiac screening without indication at general medical examination (indicator of low-value care).²⁴¹
- Slight reduction in antibiotic prescriptions in outpatient facilities.²⁴²

The effectiveness of PBP is mixed for outcome and process measures. Key findings for PBP programs that use outcome and process measures include:

- Little or no impact on high blood pressure or cholesterol in outpatient facilities.²⁴³
- Greater blood pressure control and response to high blood pressure for individual physician incentives.²⁴⁴
- Increased number of patients with record of total cholesterol or blood pressure.²⁴⁵
- Increased rate of recommendations for medications to prevent clotting to 12 percent compared with 6 percent for clinics without P4P incentives.²⁴⁶
- Inconsistent results for HbA1c control.²⁴⁷
- No change in sepsis mortality.²⁴⁸

The effectiveness of PBP varies for utilization measures. Key findings for PBP programs that use utilization measures include:

- Decreased 30-day hospital readmissions rate.²⁴⁹
- Reductions in surgical site infections and length of stay.²⁵⁰
- Little existing evidence on reducing adverse events in hospitals.²⁵¹

The effectiveness of PBP is mixed for efficiency or cost measures. Key findings for PBP programs that use efficiency or cost measures include:

- Cost savings to Medicare.²⁵²
- No impact on composite high-value or low-value care.²⁵³
- Increased prescription of opioids for low back pain, a measure of low-value care.²⁵⁴
- Poorer provider communication.²⁵⁵
- Poorer communication among office staff.²⁵⁶

Clinical setting. For CMS' main hospital-based PBP programs, evidence is limited but shows some promise. Both HRRP and HACRP^{xxxiv} are associated with a decreasing 30-day hospital readmissions rate since 2012.²⁵⁷ HACRP has also resulted in cost savings to Medicare and reductions in surgical site infections and length of stay.²⁵⁸ However, a recent study on the interaction between the HRRP and Hospital VBP found an association between participation in the Hospital VBP and higher readmission rates in the HRRP.²⁵⁹ This may reflect hospital administrators' challenges in balancing multiple sets of performance measures. A Cochrane Review noted that HACRP has low-quality evidence that it is reducing adverse events in the hospital, such as pressure ulcers.²⁶⁰ The authors found no studies exploring impacts on mortality, quality of care, equity, or resource use for HACRP, Hospital VBP, or HRRP.

^{xxxiv} The [Environmental Scan on Encouraging Rural Participation in Population-Based Total Cost of Care \(TCOC\) Models](#) has additional information on HRRP and HACRP.

In terms of ambulatory care, another recent Cochrane Review evidence synthesis and meta-analysis found that PBP mechanisms in outpatient health settings may outperform traditional FFS-related mechanisms on quality of care for hypertension.²⁶¹ A prior systematic review found that ambulatory care PBP programs in the U.S. had primarily positive effects on process-of-care outcomes, inconsistent effects on patient utilization, and limited evidence on patient intermediate outcomes or health outcomes.²⁶²

Level of accountability. According to a systematic review of pay-for-performance programs, there is a non-significant increase in odds of achieving a positive effect if an incentive is paid to individual providers compared with groups.²⁶³ The Health Care Payment Learning & Action Network (HCP LAN) framework considers pay-for-reporting a time-limited step in the evolution to APMs and population-based payments in which providers can familiarize themselves with quality measures and reporting mechanisms before moving into pay-for-performance arrangements.²⁶⁴ While CMS maintains some pay-for-reporting programs, the focus of CMS' and other value-based payment initiatives is pay-for-performance. Further, one study found that physicians preferred a P4P payment bonus made every six months compared to an annual payment.²⁶⁵

Risk arrangements. Penalties for poor performance may have a stronger effect on provider and hospital behavior than rewards for good performance.^{266,267}

Magnitude of incentives. Findings are inconclusive regarding the impact size and direction of incentives have on outcomes. One systematic review suggested that smaller incentives are associated with greater improvement in provider communication and interaction among staff.²⁶⁸ However, another metanalysis found that PBP programs with over five percent of a salary or usual budget tied to performance measures had three times the effect of programs with smaller incentives.²⁶⁹ Still another study found no association between the size of incentive payments and outcomes.²⁷⁰

Benchmarks for performance measures. In one systematic review, there were no conclusive findings as to whether assessing performance relative to a participant's previous performance, relative to their peers, or compared with an absolute target has a stronger influence on performance.²⁷¹ Another literature review found that paying for improvement was less likely to be effective compared with programs that did not reward performance improvements.²⁷²

XII.C. Unintended Consequences of Pay-for-Performance Measures

PBP programs have been associated with some deleterious unintended consequences, particularly for patients with high clinical and social needs. Evidence suggests that providers who serve low-income patients are disproportionately penalized under PBP programs if compared with the national average, and thus may avoid treating low-income patients.^{273,274} One potential solution is to compare providers who see patients with complex needs to one another instead of to the general population of providers. For example, HRRP transitioned to socioeconomic peer groups in 2019 wherein providers are compared with other providers with similar patient panels.²⁷⁵ By 2022, this resulted in reduced performance penalties for hospitals treating larger proportions of minority patients within HRRP.²⁷⁶

Overall, PBP programs have had mixed effects on health disparities.²⁷⁷ A recent CMS study reported that disparities were most commonly noted in racial and ethnic groups and dual-eligible beneficiaries, specifically regarding readmissions and medication adherence.²⁷⁸ Researchers found commonalities

among PBP programs that reduced health disparities and suggested six ways to leverage PBP to avoid negative consequences:

- 1) Risk/Case-Mix Adjustment;
- 2) Stratified Performance Measures/Stratification;
- 3) Disparity Reduction Metrics;
- 4) Exception Reporting (compensation for extra hours worked and safety hazards);
- 5) Pay-for-Improvement; and
- 6) Population-Specific Metrics.²⁷⁹

Researchers have expressed concerns about other potential unintended consequences of pay-for-performance (P4P) programs, including:

- Reduced job satisfaction and motivation for clinicians because of perceived loss of autonomy;
- Providers “gaming the system” or driving care toward processes covered by P4P programs and stinting on other care;
- The administrative burden to gather and verify performance metrics; and
- Attribution challenges given that patients seek care from multiple providers.²⁸⁰

XII.D Considerations for Future PBP Models

As PBP programs expand, researchers have noted several factors to consider to improve the effectiveness and patient-centeredness of these programs. Researchers at Mayo Clinic have observed that PBP programs too often focus on secondary data that do not consider the needs of complex patients.²⁸¹ They note that performance measures and payment mechanisms may need to differ among primary/community-based care, acute care, and complex care. They propose a framework based on the six quality domains, which include patient-reported quality of life and experience with care measures.

In another assessment of PBP programs, authors described several dimensions across which to consider optimal PBP schemes.²⁸² One element is the program scope, suggesting it is important to understand the tradeoff between a broad focus that spurs system-wide change and narrow focus that reduces administrative burden. With evidence to support both approaches, a strategy that transitions from narrow to broad over time may be ideal. Another consideration is whether to set benchmarks based on absolute performance or performance relative to prior trends. While it might be easier to give providers specific targets to meet, relative benchmarking may be more financially predictable for payers and requires less historical data. Rewarding improvement or achievement in PBP models also requires a balanced approach to encourage providers who have not achieved the level of performance of more advanced VBP participants without reducing motivation to improve quality. Finally, careful consideration must go into balancing penalties and bonuses, and to the size of the incentives.

XIII. Attribution Methods Used for Models and Performance Measures

For information on how patients are attributed to models (e.g., claims-based versus voluntary attribution, prospective versus retrospective attribution), please refer to the March 2022, March 2023, and June 2023 environmental scans.

Although patients are attributed to a given model, it is important to note that not all patients for that model may be included in certain performance measures. For example, inclusion criteria for the

denominator of a performance measure may define patient population by age, setting, diagnosis, or procedures (e.g., include in the denominator patients at least 18 years of age who have had a diagnosis of congestive heart failure). Further, exclusions may occur that remove specific patients from the denominator of a given performance measure (e.g., exclude women who have had a mastectomy from the denominator of a screening for breast cancer measure).²⁸³

XIV. Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

When constructing performance measures, researchers and policymakers can draw from several different data sources. In a 1999 report by the National Research Council Panel on Performance Measures and Data for Public Health Partnership Grants, the panel identified five main types of data for measurement: administrative data, clinical registries, patient medical records, standardized clinical data, and surveys.²⁸⁴ The Agency for Healthcare Research and Quality (AHRQ) adopts a similar classification system for performance measurement data sources in its overview of data sources for quality measurement, as does the CMS MMS Blueprint.^{285, 286}

- **Administrative data**, such as claims, billing records, and encounter data, are collected as part of the operation of a program. Common data elements include type of service, number of units (e.g., days of service), diagnosis and procedure codes for clinical services, location of service, and amount billed and amount reimbursed.
- **Clinical registries** attempt to capture information about all events of interest on such matters as health status (e.g., births, deaths, cases of disease) or risk factors (e.g., immunizations, environmental contaminants). Qualified Clinical Data Registries and Qualified Registries are both CMS-approved vendors that collect clinical data from clinicians and submit to CMS on their behalf.
- **Patient medical records** contain clinical information obtained in the course of providing health care.
- **Standardized clinical data** are collected by certain types of facilities on patient status at set time intervals and stored in datasets. Standardized clinical data could include patient assessment data, such as Standardized Patient Assessment Data Elements (SPADEs) for Medicare beneficiaries in skilled nursing facilities.
- **Surveys** obtain data through the systematic collection of information from a representative sample of a population of interest. They capture self-reported information from patients or caregivers about their health care experiences and from providers about their perceptions and activities. Surveys are specifically fielded to collect data from a sample of patients, often through the phone, mail, or via the internet.

In its seminal 2015 report *Vital Signs*,²⁸⁷ the Institute of Medicine (IOM) adopted a different approach to classifying data sources for performance measurement. The IOM identified data sources available for assessing progress along each of the four study dimensions: population health, quality of care, cost of care, and engagement in health and health care. Within each of these dimensions, the IOM described potential data sources for performance measurement:

- Population health: individual-level social data, population surveys, reportable diseases, and vital statistics.

- Quality of care: patient-level clinical care data (e.g., EHRs, registries), population-level safety data (e.g., adverse event reporting registries, public health surveillance), population-level clinical data (cancer, chronic condition, and screening registries), claims data, patient-reported outcomes, surveys, and operational and financial data for health care organizations.
- Cost of care: single-payer claims data (e.g., Medicare or private payer data), multi-payer claims databases, surveys, organizational operational data, organizational chargemasters, and the Healthcare Cost and Utilization Project (HCUP) statistics and databases.
- Engagement in health and health care: primarily collected through surveys.

Both AHRQ and the IOM describe the benefits and drawbacks of different data sources for quality measurement. Advantages of administrative data (including claims and clinical encounter data) are that the data are available electronically, are cheaper to collect than medical record data, are available for an entire population of patients, report across multiple payers, and are fairly uniform in coding practices and systems. One challenge AHRQ and IOM identified with respect to claims data is limited clinical information that could be useful for determining the appropriateness of care or identifying the target patient population. For example, claims may include only diagnoses that impact payment levels. Other disadvantages include questionable accuracy for public reporting due to the primary purpose of billing, incompleteness, and time lags between service provision and data availability. The usability of claims data will also be limited when dealing with small populations of patients, particularly those with rare conditions; users will need to proceed with caution when reporting results or drawing conclusions based on small sample sizes that may not be representative of the greater population. Social determinants of health are also often missing from claims data, despite the availability of Z codes, which are discussed in **Section XII.B.**

AHRQ and IOM note that patient medical record data, which document a patient's medical history and care, are rich in clinical detail. Providers also view medical record data as a valid source of information. However, medical record data may originate from different clinical sites and contain data in different formats (e.g., paper-based data and digital data from different EHR systems), potentially requiring trained data abstractors. Thus, medical record data can be costly and complex to collect for quality measurement.

AHRQ notes that patients are the best source for collecting some types of data, such as patients' perceptions of the outcomes of the care they received. There are well-established methods for designing and administering patient surveys. Survey results are also easy for respondents to understand. If not conducted rigorously, disadvantages of surveys include the cost of administration, misleading results if questions are worded poorly, non-standardized survey administration procedures, sampling bias (e.g., the population sampled is not representative of the population as a whole), and response bias (the population is not represented in the responses).

Registries may help providers achieve sufficient sample size by reporting performance for their practice group.²⁸⁸ Additionally, registries may provide more timely feedback to providers, compared with other data sources like claims or surveys. For instance, providers receive quarterly reports from the Qualified Clinical Data Registry, compared with semiannual updates from the Quality and Resource Use Reports.²⁸⁹ Performance measures can be developed using standardized clinical data from existing data sets to assess facility performance across multiple domains of care. However, one disadvantage of standardized clinical data is that the data may not address all topics of interest. The remainder of this

section discusses considerations for selecting and addressing quality issues in data sources for performance measurement.

XIV.A. Key Considerations for Selecting Data Sources to Measure Outcomes

In its 1999 report, the National Research Council Panel emphasized that a performance measurement program must identify outcome goals to then guide the selection of the appropriate measures.²⁹⁰²⁹¹ From there, operationalizing measures requires access to appropriate data and analytic resources. While many data sources may be available to measure performance outcomes, few data sources are optimally suited for the task. Measure developers often must rely on data that were not collected for the purpose of performance measurement, and thus consideration of the benefits and limitations of different data sources guide selection.

Measure data require a high degree of validity and reliability to be used in APMs. To ensure that comparisons among providers and health plans are fair and that the results represent actual performance, it is critical to collect data in a careful, consistent way using standardized definitions and procedures.²⁹²

XIV.B. Best Practices/Trends in Addressing Quality Deficiencies in Data Sources

An initial step in addressing quality deficiencies in health care data sources includes improving the accuracy of the data source itself.²⁹³ This may include using software to identify duplicative or incorrect information in medical records or claims. A single data source could also be improved by adding critical data elements to it. For example, CMS introduced Z codes to the *International Classification of Diseases, Tenth Revision, Clinical Modification* (ICD-10-CM) in 2015 to document SDOH data in Medicare claims. These codes are not used for payment and, as a result, their use has been low for both MA²⁹⁴ and FFS enrollees.²⁹⁵ CMS stated that the collection of Z codes is part of the journey toward better health outcomes, enriching claims information with SDOH data relevant to a patient's health condition. The inclusion of Z codes in the data marks a step toward an information source that yields a more complete view of a patient's health and health-related social needs. CMS is currently engaged in efforts, such as provider education, to increase the use of Z codes.²⁹⁶

Data quality can also be improved by linking disparate data sets.²⁹⁷ The IOM and others have advocated for linking data to evaluate performance. While data linkages have some limitations, such as missing cases or mismatched records, methods continue to improve.

Health care leaders can be overwhelmed by suboptimal data inputs and struggle to derive insights from the data. A 2022 Harvard Business Review article outlined strategies to help overcome data overload including²⁹⁸ identifying the end users for each performance measure and developing ways to use the data to create value for each specific user. Then, integrating data across quality performance categories and settings in the organization can generate more insights that can be used to establish priorities for the organization's efforts. The final strategy is to communicate the information in the format that is most helpful to its users.

Investment in and the use of digital health technology tools (DHTT) has expanded in recent years, presenting a potential source of information for performance measurement. DHTT are devices such as apps, smartphones, and wearable remote monitoring technology devices that collect health information from patients and feed data to provider offices. Information from these tools is collected electronically,

which is cost-effective. Additional advantages of DHTT are objectivity and sensitivity of measurement, richness of high-frequency sensor data, and the opportunity for passive collection of health-related data.²⁹⁹ Measures developed from DHHT can be analyzed using either a data-centric approach, summarizing sensor information maximally sensitive to the condition of interest, or a patient-centric approach, summarizing sensor features that are optimally relevant to a patient's functioning in everyday life.

However, patients with limited digital literacy are less likely to use digital tools, which could omit an important segment of the population. Little is known about which patients actually use DHTT.³⁰⁰ As novel clinical measures using digital technology are developed, it is important to test for measure reliability early in the development process to inform potential usefulness of the measure.³⁰¹ A 2022 literature review focused on the potential use of digital health tools for sleep apnea found that the digital tools showed promising discriminatory properties compared with the gold standard test (polysomnography) in the study. The study also found that further quality testing in additional populations and environments was needed before more widespread use in clinical settings.³⁰² Novel digital clinical quality measures for other conditions will need to undergo the same testing and validation processes.

Electronic Clinical Quality Measures. According to CMS, electronic clinical quality measures (eCQMs) are measures specified in a standard electronic format that use data electronically extracted from EHRs and/or HIT systems to measure the quality of health care provided. Currently, eCQMs cover clinical processes and effectiveness, patient safety, care coordination, efficient use of health care resources, patient and family engagement, and population health. The main difference between eCQMs and MIPS CQMs is that eCQMs require structured data to be captured electronically, often at the point of care, while in most cases, MIPS CQMs have some level of additional manual data collection from the medical record often referred to as data abstraction.³⁰³

CMS began accepting eCQMs on a voluntary basis in 2013, and reporting became mandatory in 2016 for certain eligible hospitals in two programs, the Hospital Inpatient Quality Reporting Program, and the Medicare Promoting Interoperability Program. For MSSP participants, CMS is transitioning from reporting via a web interface (which will end in 2025) to reporting via eCQM.

Fast Healthcare Interoperability Resource. The Fast Healthcare Interoperability Resource (FHIR) is a standard that defines how health care information can be exchanged between different computer systems, regardless of how it is stored in those systems.³⁰⁴ It allows health care information, including clinical and administrative data, to be available securely to those who have a need to access it and who have the right to do so for the benefit of a patient receiving care.

XV. Risk Adjustment Methods Used for PB-TCOC Models

Risk adjustment in PB-TCOC models is important to determine the appropriate payment amounts for attributed enrollees and to account for differences in risk profiles of patients when comparing performance relative to benchmarks. As with other types of APMs, there are different methodologies for risk adjustment in PB-TCOC models. Most CMS programs use a form of the HCC model, which calculates a risk score based on diagnoses in health care administrative data or encounter data and select patient demographics. Extensive research and years of experience with the HCC model have led to developments and refinements over time and modified applications for different programs. For example, most CMS programs use claims in a prior year to project future expected spending

(“prospective”), including MA Star Ratings, MSSP, NGACO Model, and the IAH Demonstration. However, the ACO REACH Model uses current-year claims information (“concurrent”) for risk adjustment for high needs ACOs.

Prospective risk adjustment can serve as a reasonably accurate predictor of a population’s expected spending when average expected spending does not vary considerably over time. Further, it can be useful for a population with stable health status over time on average. Prospectively-calculated risk scores provide predictability in payment levels for providers, organizations, and payers. However, prospective risk scores may underpay for patients who historically have faced barriers to health care and have lower utilization and documented diagnoses.^{305, 306}

Concurrent risk adjustment may be more appropriate when the population size is smaller or when the health status or spending for the given patient population varies over time. The ACO REACH program uses concurrent risk adjustment for High Needs Population ACOs.

An analysis from the Duke Margolis Center for Health Policy points out that risk adjustment can have a large impact on payment amounts.³⁰⁷ As more models implement prospective payments and performance benchmarks—such as PB-TCOC models—the limitations of current risk adjustment methods become more pronounced and costly. Health care organizations expend resources to collect and identify diagnoses that impact a patient’s risk score, leading to higher payments and benchmarks, as well as higher spending. In addition, current risk scores are based largely on FFS claims, which may not be an accurate predictor of expected spending in PB-TCOC models and MA. There are fewer incentives to document all of a patient’s diagnoses in FFS claims, and MA enrollment is increasing. Current risk adjustment methodologies largely do not incorporate social risk factors or functional status that also can be a strong predictor of health spending.

XV.A. Alternatives to Risk Adjustment

There are alternatives to the predominant HCC model to consider when developing risk adjustment approaches in PB-TCOC models, including modifications to a risk score approach that uses administrative data, as well as more novel approaches like survey-based risk, risk grouping, and risk stratification.

Including social risk data, either at the individual or local level, through a measure such as the Area Deprivation Index (ADI) is one approach to improve risk adjustment. The ACO REACH Model uses this approach to calculate health equity benchmarks, which provides an additional benchmark adjustment based on the ADI and on dual-eligible status.³⁰⁸

Another approach uses self-reported data on health and social needs via a survey to create measures of individual risk. One PTAC proposal³⁰⁹ adopted a survey-based approach to risk adjustment, using the “How’s Your Health” survey to collect data and adjust for patient risk.

Additional approaches to account for patient risk outside of the HCC model include risk stratification, outlier payments, and risk corridors.³¹⁰ Risk stratification is the process of segmenting patients into groups of similar complexity and care needs. The first step in risk stratification is to identify high-risk patients. After stratifying patients into groups, practices can more easily make targeted care management decisions and identify those patients who may have particular care needs, with payments calibrated to account for patient spending within the risk strata. Outlier payments are additional

payments by insurers to physicians or organizations to account for encounters and patients that are exceptionally costly. Some adjustment for patient risk is therefore delivered to providers retrospectively instead of prospectively. Outlier payments function as a form of stop-loss insurance. Stop-loss insurance protects the provider against significantly higher than intended patient costs. This strategy is particularly useful for providers who care for vulnerable populations. Risk corridors are another mechanism that can protect against adverse selection and insufficient physician payments by limiting losses and gains beyond an allowable range. They set a target spending amount, and insurers pay into the program to compensate those physicians with patient costs exceeding the target. Risk corridors mirror aggregate stop-loss insurance in that physicians are protected against higher than expected total spending.

XVI. Consumer Assessment of Healthcare Providers and Systems Survey

AHRQ and CMS create and implement several patient experience surveys that inquire about a patient's experience with hospitals, doctors, health care providers, health and drug plans, and more. These surveys focus on matters important to the patient, such as coordination of health care needs, communication with doctors, and understanding medication instructions. The surveys highlight the patient care experience as well as patient satisfaction.³¹¹ Many CMS-administered patient experience surveys are part of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) suite of surveys.³¹² This section describes the goals and domains of CAHPS, including gaps in topics, as well as how CAHPS is applied.

XVI.A. CAHPS Goals and Domains

The CAHPS program was developed in 1995 by AHRQ to increase scientific understanding of patient experience with health care. In 1999, the CAHPS Consortium, overseen by AHRQ, developed design principles to guide the creation of the CAHPS surveys.³¹³ The CAHPS Consortium oversees the technical quality of the work conducted by the CAHPS program; the Consortium consists of staff from AHRQ, Yale University, RAND Corporation, and Westat.³¹⁴ In collaboration with AHRQ, other federal agencies within HHS, such as CMS, play a role in the development of CAHPS surveys. AHRQ is responsible for the overall CAHPS program; however, both AHRQ and CMS fund and develop testing for CAHPS surveys.³¹⁵ CMS requires and administers certain CAHPS surveys, while AHRQ does not specifically require the implementation of any CAHPS surveys, nor does it directly administer any CAHPS surveys.³¹⁶

CAHPS surveys allow organizations to learn about their patients' experiences and subsequently improve care processes. The main goal of the CAHPS surveys is to advance knowledge, measurement, and improvement of patients' experiences with health care.³¹⁷

CAHPS surveys measure patients' experiences of care across four areas:

- Providers (e.g., clinicians and medical groups, hospices, home health care, and surgical care).
- Condition-specific care (e.g., cancer care and mental health care).
- Facility-based care (e.g., hospitals, dialysis centers, nursing homes, and outpatient ambulatory surgical centers).
- Health plans (e.g., health plans, dental plans, and home and community-based services).³¹⁸

There are currently 11 different CAHPS surveys that CMS administers: Hospital, Home Health, Home and Community-Based Services, Fee-for-Service, Medicare Advantage and Prescription Drug Plan, In-Center

Hemodialysis, Nationwide Adult Medicaid, Hospice, Outpatient and Ambulatory Surgery, MIPS, and Emergency Department CAHPS.³¹⁹

Other CAHPS Consortium approved surveys that are not administered by CMS include the following CAHPS surveys: Clinician & Group, Surgical Care, American Indian, Cancer Care, Mental Health Care, Health Plan, Dental Plan, Emergency Department, Child Hospital, and Nursing Home.³²⁰

Domains are functional areas covered by CAHPS surveys.³²¹ Domains are generally broad and are selected based on their importance in reaching and maintaining a high level of quality care, importance to consumers in selecting health care, and whether consumers are the best or sole source of information.³²² Within each survey, domains are tailored to fit the specific facility or type of care being delivered. While domains are tailored to fit each survey, many surveys cover similar topics such as access to care, communication, care coordination, and interactions with staff.³²³ For an overview of the domains covered by the four areas of care CAHPS evaluates, see **Exhibit 27**.

Exhibit 27. CAHPS Surveys by Domain

Survey	Domains
Clinician & Group Survey ³²⁴	<ul style="list-style-type: none"> • Access to care • Communication • Care coordination • Customer service • Overall rating
Surgical Care ³²⁵	<ul style="list-style-type: none"> • Information to help prepare for surgery • Communication • Surgeon’s effectiveness on day of surgery • Information to help recover from surgery • Helpful, courteous, and respectful office staff • Overall rating of the surgeon
American Indian ³²⁶	<ul style="list-style-type: none"> • Getting care quickly • Getting needed care • Communication • Helpful, courteous, and respectful office staff • Guidance about your personal health • Perceived discrimination because of tribal affiliation • Shared decision-making • Care coordination • Rating of the provider • Rating of the clinic

Survey	Domains
Cancer Care ³²⁷	<ul style="list-style-type: none"> ● Getting timely appointments, care, and information ● Communication ● Care coordination ● Helpful, courteous, and respectful office staff ● Cancer care team support in managing the effect of treatment ● Involvement of family members and friends ● Availability of interpreters ● Overall rating of cancer care team ● Overall rating of cancer care
Mental Health Care ³²⁸	<ul style="list-style-type: none"> ● Getting treatment quickly ● Communication ● Getting treatment and information from the plan ● Perceived improvement ● Information on treatment options ● Overall rating of counseling and treatment ● Overall rating of the health plan ● Cultural competency
Health Plan ³²⁹	<ul style="list-style-type: none"> ● Getting needed care ● Getting care quickly ● Communication ● Customer service ● Overall rating of health plan
Dental Plan ³³⁰	<ul style="list-style-type: none"> ● Care from dentists and staff ● Access to care ● Dental plan information and services ● Overall rating of dentist ● Overall rating of care ● Overall rating of ease of finding dentist ● Overall rating of the dental plan
Emergency Department ³³¹	<ul style="list-style-type: none"> ● Getting timely care ● Communication ● Overall rating ● Willingness to recommend
Child Hospital ³³²	<ul style="list-style-type: none"> ● Communication with parent ● Communication with child ● Attention to safety and comfort ● Hospital environment ● Global rating
Nursing Home ³³³	<ul style="list-style-type: none"> ● Meeting basic needs: Help with eating, drinking, toileting ● Nurses'/aides' kindness/respect toward residents ● Nursing home provides information/encourages respondent involvement ● Nursing home staffing, care of belongings, cleanliness ● Overall rating

Survey	Domains
Hospital* ³³⁴	<ul style="list-style-type: none"> ● Communication ● Responsiveness ● Discharge information ● Hospital environment ● Overall rating ● Willingness to recommend
Home Health* ³³⁵	<ul style="list-style-type: none"> ● Care of patients ● Communication ● Specific care issues ● Overall rating ● Willingness to recommend
Home and Community-Based Services* ³³⁶	<ul style="list-style-type: none"> ● Access to care ● Communication ● Experience with case managers ● Choice of services ● Medical transportation ● Personal safety ● Community inclusion and empowerment
Fee-for-Service* ³³⁷	<ul style="list-style-type: none"> ● Communication ● Care coordination ● Patient experiences getting needed health care ● Experiences with personal doctors and specialists ● Customer service
Medicare Advantage and Prescription Drug Plan* ³³⁸	<ul style="list-style-type: none"> ● Access to care ● Ease of getting appointments quickly ● Communication ● Care coordination ● Access to information ● Ease of getting prescriptions ● Overall rating of health/drug plan ● Rating of health care quality
In-Center Hemodialysis* ³³⁹	<ul style="list-style-type: none"> ● Nephrologists' communication and caring ● Quality of dialysis center care and operations ● Providing information to patients ● Rating of kidney doctors ● Rating of dialysis center staff ● Rating of dialysis center
Nationwide Adult Medicaid* ³⁴⁰	<ul style="list-style-type: none"> ● Demographic and health characteristic ● Access to care ● Barriers to care ● Global ratings of care

Survey	Domains
Hospice* ³⁴¹	<ul style="list-style-type: none"> • Communication • Getting timely help • Treating patient with respect • Emotional and spiritual support • Help for pain and symptoms • Training family to care for patient • Rating of hospice • Willingness to recommend
Outpatient and Ambulatory Surgery* ³⁴²	<ul style="list-style-type: none"> • About facilities and staff • Communication • Preparations for discharge and recovery • Overall rating of facility • Willingness to recommend
MIPS* ³⁴³	<ul style="list-style-type: none"> • Timely care • Communication • Patient’s rating of provider • Access to specialist • Health promotion and education • Shared decision-making • Health status/functional status • Courteous and helpful office staff • Care coordination • Stewardship of patient resources
Emergency Department* ³⁴⁴	<ul style="list-style-type: none"> • Timeliness of care • Communication • Overall rating of emergency department • Willingness to recommend

*These surveys are administered by CMS.

XVI.B. Domains Not Captured in the CAHPS

Telehealth services are not referenced as a specific domain within the CAHPS surveys.³⁴⁵ Telehealth services include virtual patient portals, virtual office visits, and home health monitoring tools. Many health care providers now incorporate web-based systems into care strategies, and consequently patients are increasingly interacting with telehealth services. While communication is a domain within the CAHPS, telehealth itself is not referenced. Further, domains could be added to the CAHPS to capture the availability, effectiveness, quality, and timeliness of telehealth care. As of 2022, the CAHPS Consortium is working to develop this area and has released question wording that is inclusive of in-person, video, and telephone-based doctors’ visits.³⁴⁶

Preventive measures and health equity measures are also missing from CAHPS domains.³⁴⁷ When creating and analyzing quality measures, it is important to choose measures that are meaningful to the population being managed. While this is the goal of the CAHPS surveys, some measures do not apply to specific populations.³⁴⁸ The access to care domain used in some CAHPS surveys could be expanded to incorporate access to preventive forms of care and equitable access to a healthy lifestyle. Measures of SDOH could also capture disparities in access to care.

XVI.C. What CAHPS Measures are Appropriate to Use in PB-TCOC Models

Patient-reported experiences of care can be used to more comprehensively inform PB-TCOC models. Specific CAHPS measures that are appropriate to use in PB-TCOC models include provider communication; getting timely care, appointments, and information; health promotion and education; access to specialists; health status/functional status; and shared decision-making.³⁴⁹ These measures can provide insight on the components of PB-TCOC models patients value most as well as their experiences with the components.

CAHPS measures have been used to describe patient experience in care delivery programs, such as Medicaid ACOs in Maine, Massachusetts, Minnesota, New Jersey, Oregon, Rhode Island, Utah, and Vermont, as well as in CMMI models such as ACO REACH, CEC, CPC+, EOM, Expanded HHVBP, HHVBP, NGACO, and PCF.³⁵⁰ CAHPS surveys are also used in some federal value-based purchasing initiatives in which CMS provides financial incentives for providing high-quality services, such as Hospital VBP, HHVBP, MSSP, and MIPS.³⁵¹

XVII. Best Practices for Measuring Performance

Because PB-TCOC models are relatively new, there is currently little evidence in the literature supporting best practices for performance measurement for PB-TCOC models. However, several best practices have been developed in performance measurement more generally that may apply to this new type of model.

Best Practices in Developing PB-TCOC Performance Measures

A Population-Based Payment Work Group was convened as part of the CMS Alliance to Modernize Healthcare. This Work Group developed a set of recommendations for performance measurement in population-based payment models including:³⁵²

- **Base measures on results that matter to patients:** This recommendation suggests focusing on outcomes with high relevance to patients, like functional status, or intermediate outcomes and processes that research supports directly result in better long-term outcomes.
- **Rely on core measure sets but leave room for innovation:** The Work Group recommends using existing measure sets in PB-TCOC models for now, with a focus on pursuing innovation in measure development and use as part of model testing and implementation.
- **Develop a governance process:** HHS, in partnership with stakeholders, should lead efforts to establish measurement priorities and oversee the development, testing, and use of new measures.
- **Improve data collection infrastructure:** This recommendation supports improved national infrastructure to collect and analyze data to support development and use of population-based outcomes measures and patient-reported outcomes measures.
- **Ensure that providers have meaningful incentives for improvement:** Performance measures should align with patient and provider goals and be tied to financial incentives for providers.
- **Align performance targets with long-term goals:** The Work Group recommends developing performance targets that support ongoing improvement, allowing for sharing of best practices across providers and health systems, and not imposing a forced set of winners and losers.

- **Use best practices in measurement science and implementation:** The Work Group recommends an accelerated approach to measure development and testing that allows for testing in the field and feedback from providers prior to tying performance to payment.

NQF set forth criteria for evaluating quality measures that are applicable to developing PB-TCOC performance measures.³⁵³ These criteria were set by NQF as part of its endorsement process, but measures do not need to be endorsed to follow these best practices. NQF’s criteria for performance measures included:

- Importance to measurement and reporting: Is the measure evidence-based, does it address a known performance gap, and does it address a specific national health goal or other high priority in health care?
- Scientific acceptability of measure properties: Is the measure reliable and valid? Is risk adjustment or identification of disparities necessary?
- Feasibility: Are the data needed for the measure routinely collected?
- Usability and use: Is the measure already in use, or is there an evidence-based rationale for how it could be used to support performance improvement?
- Related or competing measures: Are related or competing measures available and, if so, is this measure superior or sufficiently different to justify its use?

CMS had further developed the NQF evaluation criteria into a checklist for creating population-based measures.³⁵⁴ While the checklist was initially developed for ESRD-related models, it is applicable to other models as well. CMS’ checklist fills a gap in the literature by helping to translate NQF’s criteria into specific steps relevant to developing population-based measures.

In early 2023, Battelle was awarded the National Consensus Development and Strategic Planning for Health Care Quality Measurement contract. At this time, Battelle’s PQM replaced NQF as the CMS CBE for performance measurement.

The AMA has also developed “A Guide To Physician-Focused Alternative Payment Models.” This guide includes suggestions for performance measurement for each model type, including episode of care models.³⁵⁵ The Maryland Total Cost of Care Model is perhaps the largest PB-TCOC model currently being tested. It is a state-wide APM designed to reduce costs and improve the quality of care for Medicare beneficiaries in Maryland. The model includes a limited set of spending, utilization, quality, patient satisfaction, and population health measures with a focus on measures for which Maryland ranks low among states.³⁵⁶ Maryland’s measurement approach could serve as a template for other TCOC models, as its large size and lengthy demonstration period allow for exploration of the long term effectiveness of measures.

Best Practices in Number of Measures

Across Alternative Payment Models, the number and breadth of measures included in performance measure sets varies widely. As noted in a 2017 paper, there is little evidence to support either broad or narrow performance measurement in terms of how they may affect health outcomes.³⁵⁷ Narrow performance measurement can potentially encourage providers and systems to ignore processes and outcomes that are important to patients but not part of the APM, while most APMs struggle to show substantial improvement on broad measures of outcomes.³⁵⁸ Performance measures should, as a general rule, be aligned with the goals of the model.

In 2015, the National Academies recommended a set of 15 core measures across four domains designed to drive health care improvement.³⁵⁹ The domains were healthy people, quality of care, costs of care, and people's engagement in health and health care. The core measures were all broad, population-based measures rather than narrow measures of specific specialties or settings.

Similarly, CMS has developed a new Universal Foundation, which is designed to align and consolidate performance measures across CMS programs. The Universal Foundation focuses on broad measures for specific subpopulations, such as people with diabetes or people with asthma.³⁶⁰

Incorporating Prescription Drug Spending and Utilization

There is very little literature on how best to incorporate performance measures designed to improve prescription drug spending and utilization into PB-TCOC models. However, several studies have explored each of these areas for other APMs. For prescription drug spending, some commercial insurers have tested performance-based pharmacy payment, though data are not available to assess results on spending.³⁶¹ Measures used in these models could be applicable for PB-TCOC models, including 90-day fill rates, medication adherence, and medication therapy management post-discharge.³⁶²

Additionally, Blue Cross and Blue Shield of Massachusetts (BCBSMA) implemented a global payment system called the Alternative Quality Contract (AQC), but the model did not have any statistically significant effect on either total prescription drug spending or prescriptions filled.³⁶³ Other APMs have used similar measures of prescription drug spending and utilization. One study of a commercial ACO in California found no effect of the ACO model on prescription drug utilization, spending, or use of generics.³⁶⁴ The measures used in the study, including total prescription drug spending, average scripts filled, average days supplied, and average generic share of total spending, scripts filled, or days supplied, could be replicated in PB-TCOC models. Another study of a commercial patient-centered medical home (PCMH) model found small reductions in total prescription drug spending attributable to the model.³⁶⁵ Overall, population-based payment models like global budgets, ACOs, and PCMHs do not appear to lead to significant reductions in prescription drug spending, but the effects of PB-TCOC models could be tested using available measures of spending and utilization.

Best Practices to Measure or Account for Equity and Social Determinants of Health

Approaches to using performance measures to improve equity are still being tested. Researchers have suggested reporting common quality measures, like hospital readmissions, by race-specific subgroups to help identify gaps and promote equity,³⁶⁶ although this approach works for only large subgroups.³⁶⁷ NQF has also developed a road map to promote health equity through performance measurement.³⁶⁸ The ACO REACH Model is designed to promote equity through requirements to submit equity plans and collect equity data, but the model does not currently include direct performance measures for equity. Many equity gaps may be related to SDOH. Performance measurement approaches that do not account for SDOH can inadvertently penalize providers and health systems that treat high-risk populations.³⁶⁹ Measure sets that do not adjust for SDOH could make participation in PB-TCOC models particularly unattractive for these providers, given the significant financial risks.

Many approaches to incorporating SDOH into performance measurement have been proposed and studied, but there is no consensus yet on best practices. However, academic and gray literature generally agree that SDOH affects health; therefore, it is a best practice to account for SDOH in performance measurement in some way.

In general, there are two recommended approaches to incorporating SDOH in existing performance measures: peer grouping and risk adjustment. MedPAC has recommended peer grouping in Medicare pay-for-performance or public reporting programs.^{370,371} This approach groups providers by the share of their patients facing certain social risks, often by dual-eligibility as a proxy for low income and high social risk. Providers are then evaluated within their peer group to determine rewards and penalties. CMS currently uses the peer-grouping approach for the HRRP, whereby a hospital's performance is measured against other hospitals that treat a similar share of dual-eligible patients.³⁷²

Under risk adjustment for SDOH, measurable social risk factors are added to diagnosis-based risk adjustment formulas to adjust measure performance for underlying patient populations.³⁷³ NQF conducted a trial of these approaches, finding that data availability was a significant barrier to implementing risk adjustment for SDOH. NQF also noted that adjusting for SDOH did not appear to have a large effect on model performance and often did not significantly change results.³⁷⁴ Additionally, risk adjustment for SDOH can further entrench inequities by relying on historical patterns of health care use and spending, which can be artificially low for vulnerable populations due to poor access to care.³⁷⁵

In 2020, ASPE proposed a framework for incorporating social risk into value-based payment that blended the risk adjustment and peer-group approaches.³⁷⁶ The basic tenets of this framework included measuring and reporting quality for the subgroups of beneficiaries with social risk factors, setting high- and fair-quality standards for all beneficiaries, and rewarding improvements in outcomes specifically for beneficiaries with social risk factors.³⁷⁷ ASPE also recommended reporting quality and resource use measures separately for dually eligible and non-dually eligible beneficiaries, and developing a standard risk-adjustment framework for social risk that includes functional status.³⁷⁸

Finally, CMMI's ACO REACH Model and the Advance Investment Payments (AIP) option under MSSP aim to incorporate social risk and health equity into the design of APMs.³⁷⁹

Best Practices in Attribution

Because of the broad reach of PB-TCOC models, there is no literature on the application of performance measures to specific provider types in these models. However, NQF's suggestions to improve attribution practices may apply. NQF notes that attributing quality measures to specific providers or provider types requires consistent and transparent attribution models, adequate data, sufficient sample size, and a clear connection between the level of attribution and the goals of the program.³⁸⁰

To that end, many common process measures focus on patients with specific conditions (e.g., hemoglobin A1C testing, statin use for patients with diabetes). Such process measures could be attributed to the physician type that typically manages those patients, but such attribution can be complicated, and measures may not be valid at all levels of attribution.³⁸¹ For example, one study found that common attribution methods can leave out vulnerable populations.³⁸² Additionally, all commonly used measures included in one review of osteoporosis measures had major limitations, and only one was valid at all levels of attribution (individual physician, group practice, or health plan).³⁸³

Best Practices for Outcome Measures

Many performance measure sets use both process and outcome measures. While some process measures are strongly associated with outcomes (e.g., regular testing of hemoglobin A1c),³⁸⁴ this varies by measure and condition.³⁸⁵ CMS approaches, such as the MA Quality Bonus Program (QBP), tend to

place higher weight on outcome measures.³⁸⁶ However, there is little evidence on best practices for transition from process to outcome measures. One difficulty in such transitions is that outcome measures often require adequate risk adjustment, particularly for broad measures like mortality. Risk adjustment helps ensure that outcome measures do not transfer resources among providers or health systems based on the patients they treat rather than the quality of care they provide.³⁸⁷ Outcome measures may also require new data collection, whereas many process measures are often captured in EHRs. A recent CMS study discussed progress in increasing the percentage of outcome measures over the past seven years (from 36 percent in 2016 to 41 percent in 2023).³⁸⁸

Some researchers are working to develop new population-based performance measures that focus on outcomes that are important to patients but that do not require significant new data infrastructure. For example, Healthy Days at Home is a measure of how many days Medicare beneficiaries spend at home without home health or other post-acute care support. The measure was developed in conjunction with MedPAC and could help payers and providers track how well payment models perform in keeping healthy patients out of health care institutions.³⁸⁹

Another promising option for outcome measures may be those directly reported by patients. Patient-reported performance measures fall into two primary groups: PROMs and patient-reported experience measures (PREMs). PROMs focus on patient outcome experiences like symptoms, functional status, and quality of life as directly reported by the patient.³⁹⁰ In 2012 and 2013, the AMA convened a Technical Expert Panel to identify best practices for developing PROMs.³⁹¹ Their resulting nine best practices included that there should be a rationale for use of the measure, the measures should be adequately developed and psychometrically valid, and the measures should be sensitive to change and clinically actionable.³⁹² In addition, PROMs are frequently used internationally,³⁹³ particularly for joint replacement,^{394,395} and increasingly used to measure outcomes in palliative care.³⁹⁶ There is also growing use of PROMs in initiatives to promote patient-centered care.³⁹⁷ One synthesis study found that providers work to improve care based on public reporting of PROMs if they perceived the data to be credible, timely, and provide actionable feedback. In that study, patients also found PROMs useful as a structure for raising issues with their providers.³⁹⁸ Finally, in a structured discussion, one subject matter expert suggested PROMs should be integrated into EHRs to ease the burden of using patient-reported measures for providers in population-based models.

In contrast to PROMs, which focus on outcomes, PREMs focus on patients' experiences while receiving care. Common PREMs measures include satisfaction with care, wait times, and interactions with providers and staff.³⁹⁹ CMS uses CAHPS surveys to measure PREMs, including wait times, difficulties accessing care, and patient satisfaction with care and with health plans.⁴⁰⁰ In a structured discussion, one subject matter expert suggested patient experience measures can help to prevent stinting on care in a TCOC model because it would be difficult to stint on care while having patients who report a positive experience. PREMs can signal problems with care quality, but the research is mixed on the level of correlation between PREMs and health outcomes.^{401,402,403}

Using Artificial Intelligence in Performance Measurement

Artificial intelligence (AI) applications may also hold promise for performance measurement. Some medical specialties, particularly radiology, have begun using AI as a tool to improve quality and make quality assurance processes more efficient.⁴⁰⁴ However, researchers note that substantial additional

work is needed to ensure that such approaches are standardized and provide results that are meaningful to patients and providers.^{405,406}

More broadly, AI has shown promise in classifying patients into appropriate clinical subgroups and predicting disease course or exacerbations.^{407,408,409} These advancements could help improve performance measurement by ensuring appropriate denominators for disease-specific measures or by allowing comparison of predicted disease course to actual disease course, which may reveal promising approaches to treatment. AI and machine learning techniques can also be used to help reduce measurement error due to sampling bias or missing data,⁴¹⁰ which are common in survey-based measures of patient experience and access to care, such as those captured in CAHPS and HOS. Predictive AI models can also incorporate patient-reported data to improve prediction accuracy and ensure patient-centeredness.⁴¹¹ Natural Language Processing could also allow health systems and payers to process open-ended feedback about patient experiences into usable data.⁴¹² In a structured discussion, one subject matter expert suggested natural language processing and other AI methods can be used to extract data from narrative fields. However, predictive technologies like AI or machine learning are only as good as the data that feed them, and using AI trained on large datasets that inadequately represent some populations could potentially further health inequities.⁴¹³

In the 2022 Digital Quality Measurement Strategic Roadmap, CMS also suggested that AI and similar techniques could be used to validate quality measures and ensure that measures rely on large-scale patterns rather than narrow data points.⁴¹⁴ The CMS “Artificial Intelligence Health Outcomes Challenge” also awarded prizes to developers and data scientists to develop AI tools to predict health outcomes for Medicare beneficiaries. Participants developed algorithms to predict 12-month mortality among Medicare beneficiaries for potential use in CMMI payment models,⁴¹⁵ but it is not clear how this work may be incorporated into future CMMI models.

VIII. Opportunities to Improve Performance Measures

Overall, there is little evidence that pay-for-performance and public reporting of quality measures have improved overall quality of care in the United States.^{416,417,418,419} The MA QBP, for example, was not associated with improvements in quality of care in MA plans.⁴²⁰ Researchers have identified areas where quality measures have not kept up with changing evidence, potentially harming patient care. For example, two measures focused on medication for heart failure patients took more than a year to be updated after introduction of a new, more effective medication.⁴²¹ Researchers have also raised concerns about the adequacy of quality measures to counterbalance incentives to stint on care in PB-TCOC models.^{422,423} Because PB-TCOC models hold providers responsible for total spending, without adequate quality measures, these models can incentivize providers to provide less of all types of care, lowering costs but potentially reducing patients’ access to necessary care that would improve their quality of life or overall health.

There is also inconsistency in results across performance measure sets. A study of medical group performance found that designation as “high performing” was inconsistent across measurement systems, and the “high performing” designation was particularly sensitive to the domains included in the measure set, the thresholds for each level of performance, and the classification approach.⁴²⁴

Finally, the way benchmarks are calculated can influence providers’ participation in APMs. For example, benchmark ratcheting can reduce incentives for ACOs to continue participating in future performance

periods. Many ACO models reset benchmarks each performance period to ACO spending in the prior period. Using this method to calculate the benchmark can make successful ACOs “victims of their own success.” The ratchet effect could be mitigated with program and model changes, such as using multiyear baselines, incorporating regional spending with the organization’s benchmark, and adding prior savings back to the benchmark that has already been reset based on spending in the prior contract period. However, making these types of changes to programs and models may not completely eliminate the ratchet.⁴²⁵ In a structured discussion, one subject matter expert indicated that benchmarks should be set for a three to five-year period of the contract so that providers can plan for their improvement journey without having expectations continually shift over time.

XVIII.A. Areas Where New Measures May be Needed

The literature identifies several current gaps in performance measurement in health care. In a New England Journal of Medicine article announcing the Universal Foundation, CMS identified two key gaps in current performance measures. First, there are no widely used patient safety measures focused on ambulatory care. Current patient safety measures instead focus primarily on hospital-based care. Second, CMS noted that there are gaps in measures of holistic patient well-being.⁴²⁶ In a structured discussion, one subject matter expert listed several broad clinical areas that have few endorsed outcome measures where the data could be used in population-based models. These outcome measures include oncology, obstetrics, mental health, musculoskeletal care, and cardiovascular procedures. Another subject matter expert indicated a need for more measures focused on recognition and management, such as measures that identify symptoms early before they manifest into disease. One subject matter expert suggested that risk selection, where providers optimize their enrolled population under a TCOC model to reduce costs without having to change the clinical management of the population, is an important factor to consider when designing quality measures to be used in TCOC models. Other gaps noted in the literature include addressing or accounting for SDOH, measuring and promoting equity, and developing broad population-based outcome measures that are feasible and have relevance to patients and clinicians.

As noted in the previous section, there is not yet a consensus on how to measure or adjust for SDOH. Studies by ASPE and NQF noted that universal, consistent data on SDOH are lacking, making it difficult to develop measures or measurement approaches that address social risks.^{427,428} Further, risk adjustment for SDOH is not always appropriate as it may further entrench historical inequities, and data are scarce.⁴²⁹

There is also a gap in performance measurement related to health equity. In a 2021 review of approaches to health equity measurement, ASPE noted that Medicare VBP programs do not currently include health equity measures.⁴³⁰ The report identified 10 potential approaches to measuring health equity and found that experts had the most favorable view of the CMS Office of Minority Health’s Health Equity Summary Score (HESS) approach. However, direct measures of equity like the HESS are not currently used in any Medicare APMs. For example, the ACO REACH Model does not currently include an explicit measure of health equity, but it instead adjusts overall quality scores for the level of completeness of equity data collection, including demographic information and social determinants of health.⁴³¹ While researchers have suggested reporting common performance measures separately by race and ethnicity or by SDOH needs to address equity,^{432,433} more direct measures of equity are needed to ensure adequate quality of care across subgroups regardless of group size.

There are also relatively few population-based outcome measures that have been validated and tested.⁴³⁴ Common measures like mortality may not fully align with patient or provider goals for some PB-TCOC models. In addition, outcomes like mortality are also affected by many aspects of patients' lives beyond the quality of care provided by a specific clinician, making it difficult to attribute outcome measures to a particular intervention.⁴³⁵ In a structured discussion, one subject matter expert indicated a need for better measures for high-risk populations, such as people with complex care needs, chronic health conditions, and behavioral health issues. The subject matter expert suggested that improving care in these areas can help to reduce avoidable service use, reduce adverse events, reduce poor prescribing practices, and change care in a way to promote better health outcomes.

XVIII.B. Opportunities to Modify Existing Measures

In 2014, MedPAC suggested modifying existing Medicare performance measure sets to focus on a core set of population-based measures. These included potentially preventable admissions and ED visits, 30-day mortality after an inpatient hospital stay, readmission rates, healthy days at home, and patient experience.⁴³⁶ MedPAC also suggested further research into PROMs, which it considered a promising option for distinguishing quality among MA, traditional Medicare, and ACOs.

CMS has also developed episode-based measures of costs of care focused on common, high-cost procedures. These measures could be adapted to PB-TCOC models but require updates and harmonization to guard against potential unintended consequences.⁴³⁷ Additional measures are also needed for treatment of chronic conditions, which reflect complex care processes and present difficult risk adjustment and measurement timing issues.⁴³⁸

On health equity and SDOH, several approaches have been suggested to modify existing measures to better address these important domains. To better address SDOH and equity, performance on existing measures can be reported separately for key subgroups.^{439,440,441} Measures can also be risk-adjusted for SDOH to ensure that they do not penalize providers who see a high volume of at-risk populations, although ASPE notes that this approach should primarily be used for spending and utilization outcomes.⁴⁴²

Finally, PROMs or PREMs measures collected in the CAHPS and HOS surveys could also be modified and used in PB-TCOC models. For example, the Maryland TCOC Model includes PREMs measures focused on patients' perceptions of their doctor and their hospital.⁴⁴³ In addition, prior to 2021, the MA QBP included two population-based PROMs measures that drew from the HOS: improving or maintaining physical health, and improving or maintaining mental health.⁴⁴⁴ The HOS could be used as a template for collecting PROMs data for specific interventions, or the overall health improvement measures could be applied to PB-TCOC models. Additional measures of functional status could also be added to CAHPS or HOS to support measures of patient-reported improvements in functional outcomes.

XIX. Areas Where Additional Information is Needed

This section includes a summary of some areas for consideration to guide future research on developing and implementing performance measures in PB-TCOC models. **Appendix I** further describes additional areas for future exploration and research.

Best Practices for Measuring Performance

There are currently few reports documenting best practices for performance measurement for PB-TCOC models. Reports cite that there is little evidence in support of either broad or narrow performance measurement sets, and approaches using measures to improve equity are still being tested. Future research could be conducted to identify the ideal set of performance measures that can be used in most populations, including performance measures to improve equity. Additional questions could be included that are specific to certain patient populations (e.g., condition-specific measures).

Opportunities to Improve Performance Measures

Further research could examine specific gaps in existing performance measures where new performance measures need to be developed or existing measures could be modified. Few evaluation reports have cited gaps in performance measurement, such as a lack of specificity in validated measures for depression or few measures related to access or patient experience. Further, there has been little, if no, research identifying performance measures that should be included in specific models or programs but are not included.

Appendix A. Research Questions by Environmental Scan Section

Section	Research Questions
Section V. Challenges Related to Developing and Implementing Performance Measures	<ul style="list-style-type: none"> • What is the reporting burden for performance measures among model participants? • What challenges do the providers face in attempting to achieve substantial changes in measure rates to attain the relative benchmark from year to year? • How can they address the challenges?
Section VI. Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures	<ul style="list-style-type: none"> • What gaps exist between what is currently being measured and the measures that are needed for PB-TCOC models? • Are there sufficient performance measures for the most prevalent chronic conditions (such as cardiovascular disease)? • Are there examples of important episodes of care where quality measures are not available for all parts of the episode? • What are the different types of performance targets and benchmarks used across APMs? What are their pros and cons?
Section VII. Trends in Existing Performance Measures	<ul style="list-style-type: none"> • What is the current landscape of what is being measured (for example, outcomes, processes, patient/caregiver experiences)? • How are organization-wide measures vs. specialty-specific or setting-specific measures established in the current landscape? In which contexts does it make sense to have organization-wide vs. specialty-specific or setting-specific measures? How do they differ? Do they differ in effectiveness? • What research exists regarding the relationship between existing measures and quality of care? • What methods do payers use to rebase performance measures when a majority of participants meet relative or absolute measure benchmarks? • What are the different types of data sources that can be used for construction of performance measures (e.g., claims, encounters, assessments, clinical registries, surveys)? • What are the pros and cons of the different types of data sources that can be used for the construction of performance measures? • What are the key considerations for selecting data sources to measure outcomes (e.g., level and geographic units, ability to link across datasets, timeliness, data quality, completeness, and reliability)? What are some of the common trends/best practices in addressing any data sources' quality deficiencies? • What are the advantages of using digital health technology tools for gathering health-related information from individuals (e.g., objectivity and sensitivity of measurement, richness of high-frequency sensor data, opportunity for passive collection of health-related data)? Should information from digital health technology be harnessed to create performance measures? If so, how? • How can patient medical records be extracted, cleaned, and standardized in an automated manner so that they can be used effectively for the

Section	Research Questions
	<p>construction of performance measures? How can EHRs advance these efforts?</p> <ul style="list-style-type: none"> • What are electronic clinical quality measures (eCQMs)? What are the benefits and challenges of using eCQMs? What are some of the components of eCQMs that differ from more traditional performance measures (e.g., quality data model [QDM], clinical quality language [CQL], health quality measure format [HQMF])? What models and programs include eCQMs in their implementation? • How can the Fast Healthcare Interoperability Resources (FHIR) and the Value Set Authority Center (VSAC) be leveraged for design and implementation of eCQMs? • How should performance measures be designed given the potential for systematic bias of incomplete data (e.g., missing not at random [MNAR])? • How can measure results be efficiently disseminated to providers?
<p>Section VIII. Performance Measures Used in CMMI Models</p>	<ul style="list-style-type: none"> • What spending, utilization, and quality measures are used in CMMI model implementation and monitoring? • Which measures are common across models? Which measures are unique to certain models? • How do these measures capture performance with respect to the model’s goals? • To what extent can measures be modified to meet the model’s goals, as opposed to creating a new measure? • What is the relationship between measure performance and payment under selected models? • Are different participant/provider types or risk tracks held to different performance standards? If so, how? • Are participants held responsible for a single measure or a set of measures? Which domains do these measures capture? Can participants select measures for which they are responsible? • How is performance improvement incentivized, and how do incentives vary across models? • Are there “guardrails” in place to prevent unintended consequences (e.g., worsening disparities)? If so, what are they, and how do they support equity? • Have evaluators identified any gaps in existing performance measures? If so, what gaps in existing performance measures have evaluators identified? Are key performance measures missing that should be developed? Could existing performance measures be modified to address these gaps?
<p>Section IX. Performance Measures Used in the Medicare Shared Savings Program</p>	<ul style="list-style-type: none"> • What spending, utilization, and quality measures are used in the MSSP program? • Which measures are common with other CMS and CMMI programs? Which measures are unique to MSSP? • How do these measures capture performance with respect to the program’s goals? • To what extent can measures be modified to meet the model’s goals, as opposed to creating a new measure?

Section	Research Questions
	<ul style="list-style-type: none"> • Does MSSP include patient-reported outcomes or measures of patient experience with care? What measures do they include? • What is the relationship between measure performance and payment under the MSSP program?
Section X. Performance Measures Used in the Medicare Advantage Star Ratings Program	<ul style="list-style-type: none"> • Has tying payment incentives to performance measures contributed to disparities in care delivery and health outcomes for Medicare beneficiaries? If so, what disparities have been identified, and how have model evaluators identified them? • How can PB-TCOC models draw on innovative approaches and lessons learned from other programs such as Medicare Advantage, MSSP, state Medicaid 1115 waiver programs, commercial/employer coverage, and Marketplace plans related to performance measures, performance measurement, and financial incentives? • What are strengths and weaknesses of performance measures in these programs? • What are the implications for multi-payer alignment of performance measures?
Section XI. Performance Measures Used in PTAC Proposals	<ul style="list-style-type: none"> • What spending, utilization, and quality measures are used in PTAC proposals to address Criterion 2 (Quality and Cost) and Criterion 4 (Value over Volume)? • Which measures are common across proposed models? Which measures are unique to certain proposed models? • How do these measures capture performance with respect to the model's goals? • To what extent can measures be modified to meet the model's goals, as opposed to creating a new measure? • Which proposals include patient-reported outcomes or measures of patient experience with care? What measures do they include? • What is the relationship between measure performance and payment in the proposed model?
Section XII. How Performance Measures are Linked with Payment in Other Programs	<ul style="list-style-type: none"> • Should performance-based financial incentives (e.g., performance-based payments) be determined using individual measures or a set of measures? • What are examples of existing performance measures and measure sets used to determine performance-based financial incentives? • What are the advantages and disadvantages of each? Is one more effective? • What research supports that tying performance to payment drives improvement for beneficiaries? • Where has this worked, and where has this not worked, and why? What can we learn from each example? • What differences in quality improvement are generated with financial incentives that are upside-only as compared with upside and downside incentives? What has been the relative impact of single-sided risk and two-sided risk associated with performance measures on contributing to improvements for beneficiaries?

Section	Research Questions
	<ul style="list-style-type: none"> • Is there research identifying or supporting the magnitude of payment tied to quality measures that is most effective at incentivizing improvements? If so, what does this research suggest? • Does the use of relative or absolute benchmarks result in better performance? If so, what is the optimal approach? In which provider populations is this approach most effective and why? • What is the research indicating that a balanced set of measures, that is, inclusion of measures that are intended to guard against decrements in care or unintended consequences, is effective at avoiding unintended consequences? • What is the research indicating the value of using a standardized set of measures vs. allowing flexibility in selecting measures? • How do potential unintended consequences differ based on type of measures (e.g., process vs. outcome vs. patient-reported) or level of measurement (e.g., clinician, practice, health system)? • What could be the potential impact of moving from process measures to outcome measures (for providers, programs, etc.)? Are there any potential unintended consequences? • What has been the relative impact of pay-for-reporting vs. pay-for-performance programs on contributing to improvements for beneficiaries? • What components are most important for tying performance measures to payment and financial incentives? • What approaches to design of performance-based payment incentives for PFPMs are most likely to facilitate improvement rather than unintended consequences? • How should financial incentives be tied to quality measures (including outcomes and patient experience measures)? How should benchmarks be determined? How should the magnitude of financial incentives be determined? Should upside and downside incentives be utilized? Should absolute or relative thresholds be used? Should assessment of achievement and improvement differ by type of measure and, if so, how? • How should the approach to performance-based payment (PBP) differ by the type of entity that is being measured (e.g., larger entities vs. small practices, degree of experience with value-based payment)?
<p>Section XIII. Attribution Methods Used for Models and Performance Measures</p>	<ul style="list-style-type: none"> • What are the different attribution methods (e.g., plurality of primary care services, plurality of specialty care services, anchor stay/visit/procedure; prospective or retrospective; claims-based or voluntary) used for performance measures for APMs? • Are specific attribution methods better for assessing spending, utilization, or quality using certain performance measures? If so, which performance measures or types of performance measures should be used with which attribution methods? • What are the pros and cons of these different attribution methods with respect to performance measurement in APMs? • What are best practices for attributing patients to providers for the purposes of performance measurement?

Section	Research Questions
Section XIV. Assessment of the Different Types of Data Sources Used for Construction of Performance Measures	<ul style="list-style-type: none"> • How are organization-wide measures vs. specialty-specific or setting-specific measures established in the current landscape? In which contexts does it make sense to have organization-wide vs. specialty-specific or setting-specific measures? • If we want to directly measure systems change and how organizations provide care, what should be the mixture of outcome, patient experience, and process measures? • Do we want to measure outcomes of care and patient satisfaction (and assume that implied objectives such as care coordination and patient-centered care are being achieved), or do we want to directly measure changes in how care is being provided? What strategies can be used to measure outcomes and/or changes in how care is being provided?
Section XV. Risk Adjustment Methods Used for PB-TCOC Models	<ul style="list-style-type: none"> • What are the different types of risk adjustment methods used for PB-TCOC models? What are best practices for incorporating risk adjustment/stratification into model implementation? • What are the different risk factors (demographic characteristics, clinical characteristics, organizational characteristics) that should be included in risk adjustment methodologies? • Given that risk adjustment typically accounts for measurable and available factors, what are some approaches that payers use to improve risk adjustment methodologies? • Should performance measures be adjusted to account for social and functional status-related risk factors? If so, how? • What are the alternatives to risk adjustment (e.g., peer grouping/risk stratification)? • Given the potential variation across PB-TCOC participants, what are the pros and cons of using peer groupings (or stratifications), compared with risk adjustment? • Should performance measures and benchmarks be designed and implemented to account for health disparities and SDOH? If so, how? • What are risk score caps? Do they unfairly penalize certain participants? If so, how?
Section XVI. Consumer Assessment of Healthcare Providers and Systems Survey	<ul style="list-style-type: none"> • Which models include patient-reported outcomes or measures of patient experience with care (e.g., CAHPS)? What measures do these models include? • How are patient-reported outcomes currently being used in payment models? • What have been some lessons learned? • What are potential implications and options for incorporating patient-reported outcomes in future PB-TCOC models?
Section XVII. Best Practices for Measuring Performance	<ul style="list-style-type: none"> • What are some best practices for managing the life cycle of a performance measure? When is it appropriate to remove performance measures from model implementation? • What are best practices to ensure that performance measures are implementable and feasible?

Section	Research Questions
	<ul style="list-style-type: none"> • How should our goals for performance measurement and the measures for PB-TCOC models differ from what we do for FFS payment systems? • How are the goals incorporated in the design and implementation of performance measures for PB-TCOC models? • What are best practices for designing performance measurement methodologies for population-based TCOC models? • What are the best practices in developing new measures? Are there opportunities for this process to be improved? • What are best practices for selecting and implementing performance measurement methodologies for population-based TCOC models? • What are some of the innovative approaches for measuring and incentivizing value-based care transformations? • What are some best practices for incorporating patient-centered outcome measures, such as patient preferences and quality of life, into PB-TCOC model implementation? • How can artificial intelligence tools be leveraged to predict health outcomes that can then be incorporated into the development of performance measures? • How can performance measures incorporate social determinants of health (SDOH), including access to housing, nutrition, and transportation, that can influence health care use and outcomes for vulnerable populations? • How can PB-TCOC models leverage performance measures to improve prescription drug spending, utilization, and equity? How can payment incentives be tied to performance measures that encourage appropriate and equitable use of prescription drugs? • To address the lengthy process of endorsing new measures, what are best practices for moving measures from the development process through the endorsement process? • What are best practices for implementing performance measures? • What are the best approaches for assessing the effectiveness of performance measures within PB-TCOC models? • Based on the six aims for care improvement from the Institute of Medicine (IOM) Quality Chasm Report in 2001 (safe, timely, effective, efficient, equitable, and patient-centered), have we achieved the goals, or moved closer to achieving the goals, these six aims set out to achieve? If so, how have we moved closer to achieving these goals? If not, what are the obstacles to achieving these goals? • How should APMs balance monitoring performance across a comprehensive set of performance measures with mitigating administrative burden of performance reporting on participants?
Section XVIII. Opportunities to Improve Performance Measures	<ul style="list-style-type: none"> • What are the major gaps in design and implementation of performance measures for population-based TCOC models? • What has contributed to cost savings in PB-TCOC models? Have efforts to improve quality or access to care been achieved while also maintaining costs? Have efforts to reduce costs been achieved while also maintaining quality of care?

Section	Research Questions
	<ul style="list-style-type: none"><li data-bbox="456 247 1421 384">• How does the implementation of the performance measures under the model create opportunities for participants to “game” the system? How can performance measures be designed and implemented to mitigate the risk of model participants following perverse incentives?<li data-bbox="456 390 1421 457">• To what extent should performance measures be standardized across payers and models?<li data-bbox="456 464 1421 531">• How does benchmark ratcheting reduce incentives for providers to participate in APMs? What are some strategies to address the ratchet effect?<li data-bbox="456 537 1421 600">• How can payers improve the transparency of performance targets and benchmarks?

Appendix B. Search Strategy

Research Questions	Search Terms
Section V. Challenges Related to Developing and Implementing Performance Measures	
<ul style="list-style-type: none"> • What is the reporting burden for performance measures among model participants? • What challenges do the providers face in attempting to achieve substantial changes in measure rates to attain the relative benchmark from year to year? • How can they address the challenges? 	<ul style="list-style-type: none"> • CMS websites, PubMed and Google Scholar searches: • Measuring quality CMS • Endorsement of performance measures health care • Challenges with Medicare Quality Payment Program • Standardization of performance measures health care • Challenge implementing new performance measures health care • Definition of poor diabetes control • Partnership for Quality Measurement (PQM) Endorsement & Maintenance Process
Section VI. Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures	
<ul style="list-style-type: none"> • What gaps exist between what is currently being measures and the measures that are needed for PB-TCOC models? • Are there sufficient performance measures for the most prevalent chronic conditions (such as cardiovascular disease)? • Are there examples of important episodes of care where quality measures are not available for all parts of the episode? • What are the different types of performance targets and benchmarks used across APMs? What are their pros and cons? 	<p>CMS websites, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • Measuring quality CMS • Successful model • Medicare performance measurement program • Developer of quality measures
Section VII. Trends in Existing Performance Measures	
<ul style="list-style-type: none"> • What is the current landscape of what is being measured (for example, outcomes, processes, patient/caregiver experiences)? • How are organization-wide measures vs. specialty-specific or setting-specific measures established in the current landscape? In which contexts does it make sense to have organization-wide vs. specialty-specific or setting-specific measures? How do they differ? Do they differ in effectiveness? • What research exists regarding the relationship between existing measures and quality of care? 	<p>CMS Measures Inventory Tool (CMIT), CMS and CMMI websites, and associated evaluation and model overview documents, performance measurement for chronic conditions</p>

Research Questions	Search Terms
<ul style="list-style-type: none"> • What methods do payers use to rebase performance measures when a majority of participants meet relative or absolute measure benchmarks? • What are the different types of data sources that can be used for construction of performance measures (e.g., claims, encounters, assessments, clinical registries, surveys)? • What are the pros and cons of the different types of data sources that can be used for the construction of performance measures? • What are the key considerations for selecting data sources to measure outcomes (e.g., level and geographic units, ability to link across datasets, timeliness, data quality, completeness, and reliability)? What are some of the common trends/best practices in addressing any data sources' quality deficiencies? • What are the advantages of using digital health technology tools for gathering health-related information from individuals (e.g., objectivity and sensitivity of measurement, richness of high-frequency sensor data, opportunity for passive collection of health-related data)? Should information from digital health technology be harnessed to create performance measures? If so, how? • How can patient medical records be extracted, cleaned, and standardized in an automated manner so that they can be used effectively for the construction of performance measures? How can EHRs advance these efforts? • What are electronic clinical quality measures (eCQMs)? What are the benefits and challenges of using eCQMs? What are some of the components of eCQMs that differ from more traditional performance measures (e.g., quality data model [QDM], clinical quality language [CQL], health quality measure format [HQMF])? What models and programs include eCQMs in their implementation? • How can the Fast Healthcare Interoperability Resources (FHIR) and the Value Set Authority Center (VSAC) be leveraged for design and implementation of eCQMs? • How should performance measures be designed given the potential for systematic bias of incomplete data (e.g., missing not at random [MNAR])? • How can measure results be efficiently disseminated to providers? 	

Research Questions	Search Terms
Section VIII. Performance Measures Used in CMMI Models	
<ul style="list-style-type: none"> • What spending, utilization, and quality measures are used in CMMI model implementation and monitoring? • Which measures are common across models? Which measures are unique to certain models? • How do these measures capture performance with respect to the model’s goals? • To what extent can measures be modified to meet the model’s goals, as opposed to creating a new measure? • What is the relationship between measure performance and payment under selected models? • Are different participant/provider types or risk tracks held to different performance standards? If so, how? • Are participants held responsible for a single measure or a set of measures? Which domains do these measures capture? Can participants select measures for which they are responsible? • How is performance improvement incentivized, and how do incentives vary across models? • Are there “guardrails” in place to prevent unintended consequences (e.g., worsening disparities)? If so, what are they and how do they support equity? • Have evaluators identified any gaps in existing performance measures? If so, what gaps in existing performance measures have evaluators identified? Are key performance measures missing that should be developed? Could existing performance measures be modified to address these gaps? 	<p>CMS and CMMI websites and associated evaluation and model overview documents</p>
Section IX. Performance Measures Used in the Medicare Shared Savings Program	
<ul style="list-style-type: none"> • What spending, utilization, and quality measures are used in the MSSP program? • Which measures are common with other CMS and CMMI programs? Which measures are unique to MSSP? • How do these measures capture performance with respect to the program’s goals? • To what extent can measures be modified to meet the model’s goals, as opposed to creating a new measure? • Does MSSP include patient-reported outcomes or measures of patient experience with care? What measures do they include? • What is the relationship between measure performance and payment under the MSSP program? 	<p>CMS websites and associated evaluation documents</p>

Research Questions	Search Terms
Section X. Performance Measures Used in the Medicare Advantage Star Ratings Program	
<ul style="list-style-type: none"> • Has tying payment incentives to performance measures contributed to disparities in care delivery and health outcomes for Medicare beneficiaries? If so, what disparities have been identified, and how have model evaluators identified them? • How can PB-TCOC models draw on innovative approaches and lessons learned from other programs such as Medicare Advantage, MSSP, state Medicaid 1115 waiver programs, commercial/employer coverage, and Marketplace plans related to performance measures, performance measurement and financial incentives? • What are strengths and weaknesses of performance measures in these programs? • What are the implications for multi-payer alignment of performance measures? 	<p>CMS websites and associated evaluation documents</p>
Section XI. Performance Measures Used in PTAC Proposals	
<p>What spending, utilization, and quality measures are used in PTAC proposals to address Criterion 2 (Quality and Cost) and Criterion 4 (Value over Volume)?</p> <ul style="list-style-type: none"> • Which measures are common across proposed models? Which measures are unique to certain proposed models? • How do these measures capture performance with respect to the model’s goals? • To what extent can measures be modified to meet the model’s goals, as opposed to creating a new measure? • Which proposals include patient-reported outcomes or measures of patient experience with care? What measures do they include? • What is the relationship between measure performance and payment in the proposed model? 	<ul style="list-style-type: none"> • PTAC proposal documents
Section XII. How Performance Measures are Linked with Payment in Other Programs	
<ul style="list-style-type: none"> • Should performance-based financial incentives (e.g., performance-based payments) be determined using individual measures or a set of measures? • What are examples of existing performance measures and measure sets used to determine performance-based financial incentives? • What are the advantages and disadvantages of each? Is one more effective? • What research supports that tying performance to payment drives improvement for beneficiaries? 	<p>CMS websites, HCP LAN website, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • “pay-for-reporting” AND health in Title/Abstract • “pay-for-performance” AND “systematic review” OR “meta-analysis” AND health in Title/Abstract • “performance-based payment” AND “systematic review” OR “meta-

Research Questions	Search Terms
<ul style="list-style-type: none"> • Where has this worked and where has this not worked, and why? What can we learn from each example? • What differences in quality improvement are generated with financial incentives that are upside-only as compared with upside and downside incentives? What has been the relative impact of single-sided risk and two-sided risk associated with performance measures on contributing to improvements for beneficiaries? • Is there research identifying or supporting the magnitude of payment tied to quality measures that is most effective at incentivizing improvements? If so, what does this research suggest? • Does the use of relative or absolute benchmarks result in better performance? If so, what is the optimal approach? In which provider populations is this approach most effective and why? • What is the research indicating that a balanced set of measures, that is inclusion of measures that are intended to guard against decrements in care or unintended consequences, is effective at avoiding unintended consequences? • What is the research indicating the value of using a standardized set of measures vs. allowing flexibility in selecting measures? • How do potential unintended consequences differ based on type of measures (e.g., process vs. outcome vs. patient-reported) or level of measurement (e.g., clinician, practice, health system)? • What could be the potential impact of moving from process measures to outcome measures (for providers, programs, etc.)? Are there any potential unintended consequences? • What has been the relative impact of pay-for-reporting vs. pay-for-performance programs on contributing to improvements for beneficiaries? • What components are most important for tying performance measures to payment and financial incentives? • What approaches to design of performance-based payment incentives for PFPs are most likely to facilitate improvement rather than unintended consequences? • How should financial incentives be tied to quality measures (including outcomes and patient experience measures)? How should benchmarks be determined? How should the magnitude of financial incentives be 	<p>analysis" AND health in Title/Abstract</p> <p>Snowball sample approach of Cited by and Related Articles</p>

Research Questions	Search Terms
<p>determined? Should upside and downside incentives be utilized? Should absolute or relative thresholds be used? Should assessment of achievement and improvement differ by type of measure and, if so, how?</p> <ul style="list-style-type: none"> • How should the approach to performance-based payment (PBP) differ by the type of entity that is being measured (e.g., larger entities vs. small practices, degree of experience with value-based payment)? 	
<p>Section XIII. Attribution Methods Used for Models and Performance Measures</p>	
<ul style="list-style-type: none"> • What are the different attribution methods (e.g., plurality of primary care services, plurality of specialty care services, anchor stay/visit/procedure; prospective or retrospective; claims-based or voluntary) used for performance measures for APMs? • Are specific attribution methods better for assessing spending, utilization, or quality using certain performance measures? If so, which performance measures or types of performance measures should be used with which attribution methods? • What are the pros and cons of these different attribution methods with respect to performance measurement in APMs? • What are best practices for attributing patients to providers for the purposes of performance measurement? 	<p>CMS websites, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • Performance measure attribution • Attribution methods • Performance measure exclusion criteria
<p>Section XIV. Assessment of the Different Types of Data Sources Used for Construction of Performance Measures</p>	
<ul style="list-style-type: none"> • How are organization-wide measures vs. specialty-specific or setting-specific measures established in the current landscape? In which contexts does it make sense to have organization-wide vs. specialty-specific or setting-specific measures? • If we want to directly measure systems change and how organizations provide care, what should be the mixture of outcome, patient experience, and process measures? • Do we want to measure outcomes of care and patient satisfaction (and assume that implied objectives such as care coordination and patient-centered care are being achieved), or do we want to directly measure changes in how care is being provided? What strategies can be used to measure outcomes and/or changes in how care is being provided? 	<p>CMS websites, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • Digital health technology tool • Electronic clinical quality measures • eCQMs vs. MIPS CQM • Fast Healthcare Interoperability Resource • Novel digital clinical quality measure • Quality measure

Research Questions	Search Terms
Section XV. Risk Adjustment Methods Used for PB-TCOC Models	
<ul style="list-style-type: none"> • What are the different types of risk adjustment methods used for PB-TCOC models? What are best practices for incorporating risk adjustment/stratification into model implementation? • What are the different risk factors (demographic characteristics, clinical characteristics, organizational characteristics) that should be included in risk adjustment methodologies? • Given that risk adjustment typically accounts for measurable and available factors, what are some approaches that payers use to improve risk adjustment methodologies? • Should performance measures be adjusted to account for social and functional status-related risk factors? If so, how? • What are the alternatives to risk adjustment (e.g., peer grouping/risk stratification)? Given the potential variation across PB-TCOC participants, what are the pros and cons of using peer groupings (or stratifications), compared with risk adjustment? • Should performance measures and benchmarks be designed and implemented to account for health disparities and SDOH? If so, how? • What are risk score caps? Do they unfairly penalize certain participants? If so, how? 	<p>CMS websites, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • Risk adjustment population-based total cost of care • Risk adjustment Alternative Payment Model • Risk adjustment Maryland Total Cost of Care Model
Section XVI. Consumer Assessment of Healthcare Providers and Systems Survey	
<ul style="list-style-type: none"> • Which models include patient-reported outcomes or measures of patient experience with care (e.g., CAHPS)? What measures do these models include? • How are patient-reported outcomes currently being used in payment models? • What have been some lessons learned? • What are potential implications and options for incorporating patient-reported outcomes in future PB-TCOC models? 	<p>AHRQ and CMS CAHPS websites</p>
Section XVII. Best Practices for Measuring Performance	
<ul style="list-style-type: none"> • What are some best practices for managing the life cycle of a performance measure? When is it appropriate to remove performance measures from model implementation? • What are best practices to ensure that performance measures are implementable and feasible? 	<p>CMS websites, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • TCOC performance measures • Performance measurement best practices • Performance measurement best practices for PB-TCOC

Research Questions	Search Terms
<ul style="list-style-type: none"> • How should our goals for performance measurement and the measures for PB-TCOC models differ from what we do for FFS payment systems? • How are the goals incorporated in the design and implementation of performance measures for PB-TCOC models? • What are best practices for designing performance measurement methodologies for population-based TCOC models? • What are the best practices in developing new measures? Are there opportunities for this process to be improved? • What are best practices for selecting and implementing performance measurement methodologies for population-based TCOC models? • What are some of the innovative approaches for measuring and incentivizing value-based care transformations? • What are some best practices for incorporating patient-centered outcome measures, such as patient preferences and quality of life, into PB-TCOC model implementation? • How can artificial intelligence tools be leveraged to predict health outcomes that can then be incorporated into the development of performance measures? • How can performance measures incorporate social determinants of health (SDOH), including access to housing, nutrition, and transportation, that can influence health care use and outcomes for vulnerable populations? • How can PB-TCOC models leverage performance measures to improve prescription drug spending, utilization, and equity? How can payment incentives be tied to performance measures that encourage appropriate and equitable use of prescription drugs? • To address the lengthy process of endorsing new measures, what are best practices for moving measures from the development process through the endorsement process? • What are best practices for implementing performance measures? • What are the best approaches for assessing the effectiveness of performance measures within PB-TCOC models? • Based on the six aims for care improvement from the Institute of Medicine (IOM) Quality Chasm Report in 2001 (safe, timely, effective, efficient, equitable, and 	<ul style="list-style-type: none"> • Best practices in designing quality measures • Best practices in designing performance measures for APMs • Design performance measures • Measure TCOC • What APMs can learn from Maryland TCOC Model • Developing new performance measures • Best practices for developing new quality measures • Best practices for developing new performance measures • How measure development could be improved • Potential improvements in measure development • Potential improvements in measure design • Performance measure development • Artificial intelligence and performance measures • Artificial intelligence to predict health outcomes • Quality measures artificial intelligence • Quality measures machine learning • Performance measures to improve prescription drug spending • Performance measures to improve prescription drug utilization • Performance measures to improve equity • APM performance measures and prescription drug spending • APM performance measures and prescription drug utilization • APM performance measures and equity • Social determinants of health and quality measures • Performance measures social determinants • Population-based models social determinants

Research Questions	Search Terms
<p>patient-centered), have we achieved the goals, or moved closer to achieving the goals, these six aims set out to achieve? If so, how have we moved closer to achieving these goals? If not, what are the obstacles to achieving these goals?</p> <ul style="list-style-type: none"> • How should APMs balance monitoring performance across a comprehensive set of performance measures with mitigating administrative burden of performance reporting on participants? 	<ul style="list-style-type: none"> • Quality measures social determinants • Quality measures social risks • Performance measures for specific provider types • Population-based models performance by provider type • Transitioning process measures to outcome measures • Process measures versus outcome measures • Best practices for health outcome measures • Best practices for implementing outcome measures • Total cost of care models and patient-centered outcome measures • Patient-centered outcome measures in APM • Population-based models patient experience measures • Population-based models quality of life measures • Best practices for patient experience measures • Patient-centered outcome measures • Patient-reported outcome measures APM • Incorporating patient experience into performance measurement • Data sources for patient experience measures • Broad versus narrow performance measures • Use of organization-wide performance measures • Effectiveness of performance measures • Relationship between performance measures and quality of care • Performance measures versus care quality
Section XVIII. Opportunities to Improve Performance Measures	
<ul style="list-style-type: none"> • What are the major gaps in design and implementation of performance measures for population-based TCOC models? 	<p>CMS websites, PubMed and Google Scholar searches:</p> <ul style="list-style-type: none"> • Gaps in APM performance measures

Research Questions	Search Terms
<ul style="list-style-type: none"> • What has contributed to cost savings in PB-TCOC models? Have efforts to improve quality or access to care been achieved while also maintaining costs? Have efforts to reduce costs been achieved while also maintaining quality of care? • How does the implementation of the performance measures under the model create opportunities for participants to “game” the system? How can performance measures be designed and implemented to mitigate the risk of model participants following perverse incentives? • To what extent should performance measures be standardized across payers and models? • How does benchmark ratcheting reduce incentives for providers to participate in APMs? What are some strategies to address the ratchet effect? • How can payers improve the transparency of performance targets and benchmarks 	<ul style="list-style-type: none"> • Gaps in performance measures for population-based models • Performance measures for chronic conditions • Gaps in performance measures for chronic conditions • New performance measures needed for chronic conditions • Performance measures episodes of care • Gaps in performance measures episodes of care • How to measure systems change in health care • Measure sets to measure systems change • Measure sets to assess how organizations provide care • How to balance outcome, patient experience, and process measures • Accuracy of patient experience measures • Patient experience measures versus objective measures

Appendix C. Definitions of Various Types of Performance Measures

Performance measures, including quality, outcome, cost, and utilization measures, assess all aspects of participants' performance in models. Performance measures are intended to help participants monitor their performance and help those accountable for model oversight assess model impact, for example, improving quality without increasing spending, or decreasing spending without reducing quality.^{445,446,447,448,449}

Several types of performance measures have been developed to evaluate health care quality, including those related to the structures, processes, and outcomes of care. Specifically, they may include a focus on health care utilization, cost, patient-reported experience, access to care, and measures specific to health equity. The following are examples of how these different types of performance measures have been defined.

Quality Measures. Several HHS agencies and organizations offer definitions pertaining to quality measures. The physician Avedis Donabedian developed the Donabedian model of measuring health care quality in 1966, which includes structure, process, and outcome measures.⁴⁵⁰ Several federal agencies incorporate Donabedian's model into their measurement framework, including AHRQ and CMS.^{451,452}

According to the National Academy of Medicine, quality measures assess performance on the following:

- Effectiveness;
- Safety;
- Timeliness;
- Patient- and family-centeredness;
- Access; and
- Efficiency.⁴⁵³

CMS describes quality measures as standards for measuring performance and the improvement of the patient population, as well as health plans, providers, and other clinicians. CMS further explains that "Quality measures are tools that help us measure or quantify health care processes, outcomes, patient perceptions, and organizational structure and/or systems that are associated with the ability to provide high-quality health care and/or that relate to one or more quality goals for health care."⁴⁵⁴ AHRQ states that "quality measurements typically focus on structures or processes of care that have a demonstrated relationship to positive health outcomes and are under the control of the health care system."⁴⁵⁵

Structure Measures. Donabedian described structures as including but not limited to the health care provider's physical setting, organizational policies, and tools and resources, such as staffing and funding.⁴⁵⁶ Donabedian noted that structure "is relevant to quality in that it increases or decreases the probability of good performance." CMS defines structural measures as those that "assess features of a health care organization or clinician relevant to its capacity to provide good health care."⁴⁵⁷

Process Measures. AHRQ states that process measures indicate what a provider does to "maintain or improve health, either for healthy people or for those diagnosed with a health care condition. These measures typically reflect generally accepted recommendations for clinical practice. Process measures can inform consumers about medical care they may expect to receive for a given condition or disease

and can contribute toward improving health outcomes.”⁴⁵⁸ CMS defines process measures as measures that provide adequate care and could increase the likelihood of attaining desired outcomes.⁴⁵⁹

Outcome Measures. Donabedian defined an outcome as “a change in a patient’s current and future health status that can be attributed to antecedent health care,” including structures and processes. The American Academy of Family Physicians (AAFP) defines an outcome measure as one that “assesses the results of health care, such as clinical events, recovery, and health status.” AHRQ defines outcome measures as those that “reflect the impact of the health care service or intervention on the health status of patients;” outcomes may be considered the ultimate goal of performance improvement; however, an outcome is the result of numerous factors, many beyond providers’ control. CMS defines outcome measures as those that focus on the health status of a patient (or change in health status) resulting from health care—desirable or adverse. More specifically, **Patient-Reported Outcome Measures (PROMs)** are a “report of the status of a health condition that comes directly from the patient, without interpretation of the’ response by a clinician or anyone else.”⁴⁶⁰

Utilization Measures. The American Hospital Association (AHA) defines utilization as “all related or covered services delivered during an episode of care or period of coverage, regardless of provider or setting.”⁴⁶¹ The National Academies of Science and Medicine define utilization as the “use of health-care services to diagnose, cure, or ameliorate disease or injury; to improve or maintain function; or to obtain information about their health status and prognosis.”⁴⁶²

Cost Measures. CMS defines a cost measure as a measure that “addresses health care spending, by payer or consumer, for a health care service or group of health care services, associated with a specified patient population, time period, and clinically accountable entities. A cost measure can assess cost in various ways, such as assessing total cost of care or assessing a specific set of costs.”⁴⁶³ AHRQ defines a cost measure as “a financial measure of cost, charge, reimbursement, payment, or out-of-pocket expenses associated with a visit to a health care provider.”⁴⁶⁴

Patient-Reported Measures. Patient-reported performance measures fall into two primary groups: Patient-Reported Outcome Measures (PROMs) and Patient-Reported Experience Measures (PREMs).

PROMs focus on patient outcome experiences like symptoms, functional status, and quality of life as directly reported by the patient.⁴⁶⁵ PROMs are a “report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else.”⁴⁶⁶

PREMs focus on patients’ experiences while receiving care; PREMs include CAHPS measures and often include measures of satisfaction with care, wait times, and interactions with providers and staff.⁴⁶⁷ AHRQ states that:

“Patient experience includes several aspects of healthcare delivery that patients value highly when they seek and receive care, such as getting timely appointments, easy access to information, and good communication with health care providers. The terms ‘patient satisfaction’ and ‘patient experience’ are often used interchangeably, but they are not the same thing. To assess patient experience, one must find out from patients whether something that should happen in a health care setting (such as clear communication with a provider) actually happened or how often it happened. Satisfaction, on the other hand, is about whether a patient’s expectations about a health encounter were met.”⁴⁶⁸

PROMs and PREMs are also discussed in more detail in Section XVI.

Access Measures: CMS states that access measures “provide information on a patient or enrollee’s timely and appropriate access to healthcare. This may include: access to several components of the healthcare system, such as health insurance, usual source of care, and mental health or substance abuse; structural barriers impacting access such as transportation, waiting times, or dealing with care (e.g., appointment times, waits, etc.); and the ability of a provider to address needs, such as patient-physician communication and relationship, cultural competency, and health information.”⁴⁶⁹

Health Equity Measures: RAND, at the direction of ASPE, convened a technical expert panel in 2021 and defined health equity measures as “illustrating or summarizing the extent to which the quality of health care provided by an organization contributes to reducing disparities in health and health care at the population level for those patients with greater social risk factor burden by improving the care and health of those patients.”⁴⁷⁰

Appendix D. Summary of Current Performance Measures for Selected CMMI Models and Medicare Payment Programs

A performance measure-level analysis for 24 selected Medicare payment programs/models was conducted using the CMS Measures Inventory Tool (CMIT). The analysis includes information about 17 payment programs (nine CMS value-based care programs and eight CMS pay-for-reporting programs) and seven CMMI models. The following is a list of the 24 Medicare programs/models that are included in this analysis.

- Accountable Care Organization (ACO) Realizing Equity, Access, and Community Health (REACH) Model
- Ambulatory Surgical Center (ASC) Quality Reporting Program (QRP)
- Bundled Payment for Care Improvement Advanced (BPCI-A) Model
- End-Stage Renal Disease (ESRD) Quality Incentive Program (QIP)
- Home Health Quality Reporting (QR)
- Home Health Value-Based Purchasing (HHVBP) (original)
- Hospice Quality Reporting Program (HQRP)
- Hospital Acquired Condition (HAC) Reduction Program
- Hospital Outpatient Quality Reporting (OQR) Program
- Hospital Readmission Reduction Program (HRRP)
- Hospital Value-Based Purchasing (VBP)
- Independence at Home (IAH) Demonstration
- Inpatient Psychiatric Facility (IPF) Quality Reporting Program (QRP)
- Inpatient Rehabilitation Facility (IRF) Quality Reporting Program (QRP)
- Kidney Care Choices (KCC) Model
- Long-Term Care Hospital (LTCH) Quality Reporting Program (QRP)
- Medicare Advantage (MA) Star Ratings Program
- Medicare Shared Savings Program (MSSP)
- Merit-Based Incentive Payment System (MIPS) Program
- Oncology Care Model (OCM)
- Primary Care First (PCF) Model
- Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (CHQR) Program
- Skilled Nursing Facility (SNF) Quality Reporting Program (QRP)
- Skilled Nursing Facility (SNF) Value-Based Purchasing (VBP)

The analysis provides the following descriptive information about the current performance measures included in these programs/models:

- Total performance measures
- Distinct performance measures
- Measures focused on similar aspects of care
- Types of performance measures
- Sources of performance measures
- Measure reporting level
- Performance measure endorsement status

- Whether performance measures are tied to payment^{xxxv}

The CMIT is a repository of performance measure information that includes 46 federal CMS value-based care programs or models (as of October 2023 when CMIT data were pulled).^{xxxvi} For each measure, the CMIT includes program/model name, measure name, measure definition, measure type, and measure source.

The results of this analysis of CMIT data can be found in the *Overview of Current Performance Measures Included in Selected Medicare Payment Programs* and a supplemental Excel file (2023 Performance Measure Data for 24 CMS Models and Programs) that can be accessed on the [PTAC Resources webpage](#).

The supplemental Excel file contains three tabs. The first tab provides a description of the data included in the second tab of the supplemental Excel file. The second tab provides performance measure-level information obtained for the 24 programs/models pulled from the CMIT or from CMS/CMMI websites; the third tab provides a data dictionary that includes the column name, data sources, column name from the CMIT, definition, and whether the column is a CMIT required field.

See Supplemental Excel File “2023 Performance Measure Data for 24 CMS Models and Programs.”

^{xxxv} The CMIT variable related to identifying if performance on a given measure is tied to payment for the applicable program/model is not a required field, and CMIT does not specify if performance is tied to payment for 60% (n=373) of the 618 performance measures identified in this analysis. This is a limitation of CMIT. Therefore the 24 selected programs/models were categorized as pay-for-performance, pay-for-reporting, or not related to payment based on information from the CMS website or the CMMI Innovation Models webpage.

^{xxxvi} Centers for Medicare & Medicaid Services Measures Inventory Tool, <https://cmit.cms.gov/cmit/#/MeasureInventory>

Appendix E. Summary of Selected CMMI Models

The first table provides specific details on CMMI model characteristics (i.e., clinical focus, providers, setting, and patient population); components relevant to performance measurement (i.e., utilization measures, quality measures, spending measures, and patient experience measures); technical issues related to performance measurement (i.e., how payment is adjusted for performance, requirements, volume, risk stratification or adjustment, and benchmarking); how measurement is used to determine success (i.e., measures used for implementation, measures used for monitoring, how achievement is measured, and how improvement is measured); and gaps related to current performance measures, as applicable.^{xxxvii} The second table describes lessons learned related to performance measurement for models that have been evaluated. The selected CMMI models are presented in alphabetical order by CMMI model name.

Overview of Methodology Used to Review the Selected CMMI Models

The available information on each of the 14 selected CMMI models' summary pages on the Innovation Center website was reviewed. This included an overview of the model, financial operating and performance measurement methodologies, informational webinars, evaluation reports and findings (as applicable), summaries, fact sheets, and press releases. Information found in these materials was used to summarize the models' main themes related to performance measurement and other administrative and payment characteristics. The categorizations were based on the key information highlighted in these documents and are not exhaustive. Models included in the tables are those that include at least one quality measure and one utilization and/or spending measure in implementation and/or monitoring; are ongoing, under development, or completed within the last five years; and are operational in more than one state. The selected models may have elements that fall into additional categories of context, objective, functions, and payment models.

^{xxxvii} For additional information about current performance measures used in many of these models, please refer to *Overview of Current Performance Measures Included in Selected Medicare Payment Programs*, that can be accessed on the [PTAC Resources webpage](#).

Exhibit E1. Model Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Selected CMMI Models

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
<p>Bundled Payments for Care Improvement Advanced (BPCI-A)</p> <p><i>Ongoing</i></p> <p>Years active: 2018-present</p>	<p>Clinical Focus: Cross-clinical focus</p> <p>Providers: Acute care hospitals, physician group practices, Medicare-enrolled providers, ACOs</p> <p>Setting: Inpatient and outpatient services</p> <p>Patient Population: Medicare beneficiaries with certain clinical episodes (29 inpatient, three outpatient)</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Administrative Quality Measures Set: All-Cause Unplanned Hospital Readmission; Advance Care Plan; CMS Patient Safety and Adverse Events Composite (CMS PSI 90); Excess Days in Acute Care after Hospitalization for Acute Myocardial Infarction; Hospital-Level Risk-Standardized Complication Rate Following Elective Primary Total Hip Arthroplasty; Risk-Standardized Mortality Rate Following Coronary Artery Bypass Graft Surgery (CABG)</p> <p>Alternate Quality Measures Set: All-Cause Unplanned Hospital Readmission; Advance Care Plan; Hospital-Level Risk-Standardized Complication Rate Following Elective Primary Total Hip Arthroplasty; 3-Item Care Transition Measure (CTM-3); Atrial Fibrillation and Atrial Flutter; Bariatric Surgery Standards for Successful Programs Measure; Cardiac Rehabilitation Patient Referral from an Inpatient Setting; Defect Free Care for Acute Myocardial Infarction; Discharge Medications in Eligible Implantable Cardioverter-Defibrillator/Cardiac Resynchronization Therapy Defibrillators Implant Patients; Heart Failure: ACE Inhibitor or ARB or ARNI Therapy for Left Ventricular Systolic Dysfunction (LVSD); Heart Failure: Beta-Blocker Therapy for LVSD; Risk-Standardized Mortality Rate Following Pneumonia Hospitalization; Hospital Risk-Standardized Complication Rate following Implantation of ICD Composite Measure; In-Person Evaluation Following Implantation of a</p>	<p>How payment is adjusted for performance: Participants, called Episode Initiators (EIs), receive a retrospective bundled payment or are required to pay a Repayment Amount based on reconciliation against the benchmark/target price. Further, EIs receive a Composite Quality Score (CQS) based on selected quality measures, and payment is adjusted by up to 10% for positive reconciliation amounts (where EI receives a payment) or negative reconciliation amounts (where EI is required to pay back).</p> <p>Requirements: Participants are required to select the Administrative Quality Measures Set or the Alternate Quality Measures Set at the beginning of each Model Year. Two measures (All-Cause Hospital Readmission and Advance Care Plan) apply to both Measures Sets and are required; participants must select two additional quality measures from the Administrative Quality Measures Set or three additional measures from the Alternate Quality Measures Set to be scored on. Scores result in a CQS which is tied to payment.</p> <p>Attribution: CMS attributes a Clinical Episode to one EI. Clinical Episodes are attributed to EIs based on the Physician Group Practice (PGP) that submits a claim that includes the physician's National Provider Identifier (NPI) or the Acute Care Hospital (ACH) where the services that triggered the Clinical Episode were rendered. A Clinical Episode begins at the start of an inpatient admission or outpatient procedure (MS-DRG codes are used to identify qualifying inpatient admissions, and Healthcare Common Procedure Coding System (HCPCS) codes are used to identify qualifying procedures) and ends 90 days</p>	<p>Measures used for implementation: Administrative or Alternate Quality Measures Sets based on the EI selections</p> <p>Measures used for monitoring: All listed measures, as applicable</p> <p>How achievement is measured: Payment is adjusted by up to 10% based on the CQS</p> <p>How improvement is measured: Each participant's performance is scored based on the participant's placement in the performance distribution from the baseline year.</p>	<p>N/A</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
		<p>Cardiovascular Implantable Electronic Device; Patient-Centered Surgical Risk Assessment and Communication; Preventive Care and Screening: Tobacco Use Screening and Cessation Intervention; Risk Standardized Bleeding for Patients Undergoing Percutaneous Coronary Intervention; Severe Sepsis and Septic Shock: Management Bundle Measure; Discharged on Statin Medication; STS Coronary Artery Bypass Graft Composite Score; Substance Use Screening and Intervention Composite; Therapy with Aspirin, P2Y Inhibitor, and Statin at Discharge Following Percutaneous Coronary Intervention in Eligible Patients; Time to Intravenous Thrombolytic Therapy; Volume Weighted Aortic Valve Replacement and Aortic Valve Replacement + CABG Composite Measures; Volume Weighted Mitral Valve Repair and Replacement and Mitral Valve Repair and Replacement + CABG Composite Measures</p> <p>Spending measure(s): All costs of care provided to a Medicare beneficiary during a clinical episode</p> <p>Patient experience measure(s): Patient-reported care experiences and satisfaction with care</p>	<p>after the day of discharge from hospital or completion of outpatient procedure.</p> <p>Volume: An entity must have a minimum of 10 attributed Clinical Episodes to generate a quality score.</p> <p>Risk stratification or adjustment: CMS employs a risk adjustment model that adjusts target prices based on hierarchical condition categories (HCC), HCC interactions, HCC severity, recent resource use, demographics, long-term institutional care, dementia, MS-DRGS/APCs, clinical episode category specific adjustments, and COVID-19 infection rate.</p> <p>Benchmarking: CMS calculates a Benchmark Price based on historical data to account for variation in costs.</p>		
<p>Comprehensive ESRD Care (CEC) Model</p> <p><i>No longer active</i></p> <p>Years active: 2015-2021</p>	<p>Clinical Focus: End-stage renal disease (ESRD)</p> <p>Providers: Nephrologists; ESRD Seamless</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): diabetes care (eye exam & food exam), advance care plan, medication reconciliation post-discharge, influenza immunization for the ESRD population, pneumococcal vaccination status, screening for clinical depression and follow-up plan, tobacco use: screening and cessation, falls: screening, risk assessment</p>	<p>How payment is adjusted for performance: The CEC Operations Contractor calculates the Shared Savings or Shared Losses at the end of each performance year. If the ESCO met or exceeds the total performance score (TPS) minimum levels of attainment and the total quality score (TQS) minimum level of attainment (in PY1) or the TQS minimum performance threshold (in PY2 onward), CMS multiplies the total Medicare savings or losses</p>	<p>Measures used for implementation: CEC TQS and ESRD QIP TPS</p> <p>Measures used for monitoring: Same as for implementation</p>	<p>Performance measure benchmarks for future models should be reexamined to increase likelihood of net savings.</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
	<p>Care Organizations (ESCOs)^{xxxviii}</p> <p>Setting: Nephrology clinics</p> <p>Patient Population: Medicare beneficiaries with ESRD</p>	<p>and plan of care to prevent future falls, bloodstream infection in hemodialysis outpatients, hemodialysis adequacy, proportion of patients with hypercalcemia, peritoneal dialysis adequacy: delivered dose of peritoneal dialysis above minimum, hemodialysis vascular access: maximizing placement of arterial venous fistula, hemodialysis vascular access: minimizing use of catheters as chronic dialysis access, standardized hospitalization ratio for admissions, standardized readmission ratio, standardized mortality ratio, standardized first kidney transplant waitlist ratio for incident dialysis patients, percentage of prevalent patients waitlisted</p> <p>Spending measure(s): Total Medicare Parts A and B spending</p> <p>Patient experience measure(s): In-Center Hemodialysis; CAHPS (ICH CAHPS) score based on six sub-measures: nephrologists' communication and care, quality of dialysis center care and operations, providing information to patients, rating of kidney doctors, rating of dialysis center staff, and rating of dialysis Center; Kidney disease quality of life (KDQOL) survey</p>	<p>by the ESCO TQS to determine the preliminary shared savings or preliminary shared losses payments.</p> <p>Requirements: ESCOs must successfully report CEC measure set with satisfactory reporting for data quality measures.</p> <p>Attribution: Based on first dialysis utilization encounter with a participating facility; conducted quarterly</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Organizations are categorized as either large dialysis organizations (LDOs; with 200+ dialysis facilities) and small dialysis organizations.</p> <p>Benchmarking: Based on historical Medicare Parts A and B expenditures for beneficiaries who would have been aligned to the ESCO in each of the three years prior to the start of the first PY, trended forward using national data</p>	<p>How achievement is measured: Using a sliding scale, ESCO performance is ranked, and points are awarded if being in the 30-99% percentile of national performance,</p> <p>How improvement is measured: The improvement scale connects the percentile of national performance, quality points earned, and the percentage change from the prior year.</p>	
<p>Comprehensive Primary Care Plus (CPC+)</p> <p><i>No longer active</i></p> <p>Years active: 2017-2021</p>	<p>Clinical Focus: Primary care</p> <p>Providers: Primary care providers (PCPs)</p> <p>Setting: Primary care practice</p>	<p>Utilization measure(s): Emergency Department Utilization, Acute Hospital Utilization</p> <p>Quality measure(s): Diabetes: Hemoglobin A1c (HbA1c) Poor Control (>9%); and Controlling High Blood Pressure. Both quality measures are reported as electronic clinical quality measures (eCQMs).</p>	<p>How payment is adjusted for performance: Practices receive performance-based incentive payments (PBIPs) based on patient experience, clinical quality, and utilization; practices retain all or a portion of the PBIP based on performance. The PBIP is paid prospectively for the entire subsequent year based on the prior year's performance. Practices that do not meet the annual performance thresholds for clinical quality/patient experience or</p>	<p>Measures used for implementation: Emergency Department Utilization; Acute Hospital Utilization; Diabetes: Hemoglobin A1c (HbA1c) Poor Control (>9%); Controlling High Blood Pressure; Medicare Parts A and B spending; Patient Experience of Care Survey</p>	N/A

^{xxxviii} ESCOs comprise nephrologists, dialysis facilities, and other providers.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
	<p>Patient Population: All Medicare and Medicaid beneficiaries in participating regions</p>	<p>Spending measure(s): Total Medicare Parts A and B spending</p> <p>Patient experience measure(s): Patient Experience of Care Survey, measured by response to questions from the Consumer Assessment of Healthcare Providers and Systems Clinician and Group (CG-CAHPS) Survey and the Patient-Centered Medical Home Survey Supplement</p>	<p>utilization are “at risk” for repaying all or a portion of the PBIP.</p> <p>Requirements: To retain quality and utilization components of the PBIP, CPC+ practices must:</p> <ul style="list-style-type: none"> -Report two eQMs and receive a Patient Experience Survey score; practices failing to meet this requirement are not eligible to keep the quality or utilization component of the PBIP. -Meet the 30th percentile on one out of three quality measures; practices failing to meet this requirement are not eligible to keep the quality or utilization component of the PBIP. -Meet the 30th percentile on two out of three quality measures (utilization gate); practices failing to meet this requirement are not eligible to keep the utilization component of the PBIP, and the percent of the quality component retained is the combined dollar amount based on individual performance that meets or exceeds the 30th percentile for the quality measures. -Meet the 70th percentile on two out of three quality measures and meet the 30th percentile on all three quality measures. For practices failing to meet this requirement, the percent of the quality component retained is the combined dollar amount based on individual performance that meets or exceeds the 30th percentile for the quality measures, and the percent of the utilization component retained is the combined dollar amount based on the individual performance for each of the two utilization measures. Practices that meet this requirement retain 100% of the quality component, and the percent of the utilization component retained is the combined dollar amount based on the individual performance for each of the two utilization measures. 	<p>Measures used for monitoring: Same as for implementation</p> <p>How achievement is measured: The PBIP retained is calculated by comparing a CPC+ practice’s performance with benchmark thresholds derived using a reference population. CPC+ practices may set goals by comparing their performance with benchmark performance thresholds on measures of utilization, spending, and quality of care.</p> <p>How improvement is measured: Practices may use these benchmarks to track their performance over time.</p>	

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			<p>Attribution: Beneficiaries are prospectively attributed to CPC+ practice sites, rather than individual practitioners, under both voluntary alignment and claims-based attribution for each payment quarter.</p> <p><u>Voluntary:</u> Process by which beneficiaries specify the health care practitioner and practice that they consider to be responsible for providing and coordinating their health care. Attestations made by the end of the lookback period are used for attribution to pay practices prospectively (using historical data to perform attribution before each payment quarter).</p> <p><u>Claims-based:</u> Prospective, using a two-year “look back” period to identify eligible primary care visits (e.g., evaluation and management [E&M], welcome and annual wellness visits, advance care planning, collaborative care model, cognition and functional assessment for patients with cognitive impairment, outpatient clinic visit for assessment and management [CAHs only], transitional care management [TCM], chronic care management [CCM], complex CCM, assessment/care planning for patients requiring CCM services, and care management services for behavioral health attribution) based on HCPCS codes. Beneficiaries are attributed to practices based on CCM-related services first, followed by annual wellness visits or welcome to Medicare visits, and then using the plurality of eligible primary care visits.</p> <p>Volume: At least 20 reporting practitioners or groups must meet the MIPS eligible clinician criteria for contributing to MIPS benchmarks for a benchmark to</p>		

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			<p>be created. These practitioners or groups must also each have a minimum case size of 20 beneficiaries.</p> <p>Risk stratification or adjustment: All Medicare FFS beneficiaries attributed to a CPC+ practice are assigned to one of four risk tiers for Track 1 or one of five risk tiers for Track 2. Risk score tier thresholds are defined separately for each CPC+ region. Each risk tier corresponds to a specific monthly care management fee (CMF) payment. Higher risk tiers are associated with higher beneficiary risk and higher CMFs. Beneficiary risk is generally determined by the CMS-HCC risk adjustment model. For Track 2 beneficiaries, risk tier is also based in part on a diagnosis of dementia.</p> <p>Benchmarking: The PBIP retained is calculated by comparing a CPC+ practice's performance with benchmark performance thresholds derived using a reference population. The benchmarks establish the minimum thresholds that CPC+ practices must reach to retain a portion of the incentive payment and the maximum thresholds that practices must achieve to retain the full incentive payment. Benchmarks are based on quality (patient experience of care and clinical quality) and utilization (Medicare acute care utilization).</p>		
<p>Enhancing Oncology Model (EOM)</p> <p><i>Ongoing</i></p> <p>Years active: 2022-present</p>	<p>Clinical Focus: Oncology</p> <p>Providers: Oncologists</p> <p>Setting: Oncology practices</p>	<p>Utilization measure(s): Admissions and Emergency Department Visits for Patients Receiving Outpatient Chemotherapy</p> <p>Quality measure(s): Proportion of Patients who Died who Were Admitted to Hospice for 3 Days or More; Percentage of Patients who Died from Cancer Receiving Chemotherapy in the Last 14 Days of Life; Pain Assessment and Management Set: a)</p>	<p>How payment is adjusted for performance: Retrospective performance-based payment (PBP) or performance-based recoupment (PBR) based on quality and savings during the performance period (i.e., six-month episodes of care).</p> <p>Requirements: Participants report participant-level quality measure data, beneficiary-level clinical and staging data, and beneficiary-level sociodemographic</p>	<p>Measures used for implementation: All listed measures</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: To calculate quality performance,</p>	N/A

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
	<p>Patient Population: Medicare beneficiaries with cancer</p>	<p>Oncology: Medical and Radiation - Pain Intensity Quantified, b) Oncology: Medical and Radiation - Plan of Care for Pain; Preventative Care and Screening: Screening for Depression and a Follow-Up Plan</p> <p>Spending measure(s): Total cost of care</p> <p>Patient experience measure(s): Patient-Reported Experience of Care Survey based on the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for cancer drug therapy and other validated surveys to assess end-of-life and hospice care (e.g., CAHPS Hospice Survey)</p>	<p>data to CMS no more than once a performance period. The PBP or PBR is calculated as a percentage of the benchmark amount for an attributed episode. An EOM Participant could earn a PBP if actual expenditures for attributed episodes are below the target amount.</p> <p>Attribution: CMS attributes an episode to an eligible EOM participant if the participant provided the first qualifying E&M service during the episode if that participant provided at least 25% of all qualifying E&M services for the episode. Otherwise, CMS attributes the episode based on plurality of qualifying E&M services, and the episode is attributed to the participant providing the largest proportion of qualifying E&M services during the episode.</p> <p>Volume: If a participant does not have enough episodes to meet the minimum denominator for a measure, the measure is excluded from the calculation of the Aggregate Quality Score (AQS) for the performance period.</p> <p>Risk stratification or adjustment: Cost benchmarks/target amounts are adjusted based on cancer type, dual-eligible status, and Low-Income Subsidy eligibility.</p> <p>Benchmarking: The benchmark amount is the total projected costs of attributed episodes. CMS creates a separate price prediction model for each cancer type. After using these models to establish predicted expenditures for each EOM episode, CMS applies a series of adjustments to obtain the benchmark price for each episode.</p>	<p>CMS will compare participant's performances on each quality measure to the measure benchmark and calculate the participant's aggregate quality score (AQS). CMS will crosswalk the participant's AQS to the PBP or PBR performance multiplier to arrive at the payment amount.</p> <p>How improvement is measured: Not specified</p>	

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
<p>ESRD Treatment Choices (ETC) Model</p> <p><i>Ongoing</i></p> <p>Years active: 2021-present</p>	<p>Clinical Focus: Home dialysis and kidney transplants for patient with ESRD</p> <p>Providers: Nephrologists</p> <p>Setting: ESRD facilities, transplant centers, large donor hospitals, patient home</p> <p>Patient Population: Patients with ESRD</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Home Dialysis Rate; Transplant Rate</p> <p>Spending measure(s): Per treatment payments for dialysis</p> <p>Patient experience measure(s): N/A</p>	<p>How payment is adjusted for performance: Participants receive a home dialysis payment adjustment (HDPa) and a performance payment adjustment (PPA). Medicare claim payments are increased for facilities and clinicians supporting dialysis at home, and PPAs are either increased or decreased based on the rate of home dialysis and transplant rate, calculated as the sum of the transplant waitlist rate and the living donor transplant rate.</p> <p>Requirements: All participants receive positive adjustments for home dialysis during the first three years of the model. The direction and magnitude of the PPA adjustment is determined based on the Modality Performance Score (MPS), which is assigned by CMS based on performance on the home dialysis rate and transplant rate.</p> <p>Attribution: A beneficiary is attributed to the ESRD facility with the most dialysis claims during the month, and the Managing Clinician billing the first monthly capitated payment for the month. Attribution occurs on a month-by-month basis</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Benchmarks are stratified by dual-eligible or Low Income Subsidy (LIS) beneficiaries. Further, the transplant waitlist rate is risk-adjusted based on beneficiary age</p> <p>Benchmarking: Achievement – CMS calculates benchmark year home dialysis rate and transplant rate and then calculates the 30th, 50th, 75th, and 90th percentiles of the distributions. Improvement – CMS assesses the home dialysis rate and transplant</p>	<p>Measures used for implementation: Home Dialysis Rate, Transplant Rate</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: CMS compares participant's rates to the percentile-based benchmarks and assigns points using an achievement score scale ranging from 0 to 2 points.</p> <p>How improvement is measured: CMS calculates percent improvement as the difference between the model and base year rates divided by the base year rate, multiplied by 100. Points are assigned using an improvement score scale ranging from 0 to 1.5. Participants are evaluated according to whether their percent improvement falls within >10%, >5%, >0%, and 0 or <0% improvement. Further, there is a Health Equity Incentive to the improvement scoring methodology: participants who show improvement in the home dialysis rate or transplant rate for their attributed dual-eligible or Low Income Subsidy (LIS) beneficiaries can earn additional improvement points.</p>	<p>N/A</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			rate against the base year rates to calculate an improvement score.		
<p>Expanded Home Health Value-Based Purchasing Model (Expanded HHVBP)</p> <p><i>Ongoing</i></p> <p>Years active: January 2022-present</p>	<p>Clinical Focus: Home health care</p> <p>Providers: Medicare-certified Home Health Agencies (HHAs)</p> <p>Setting: Home health setting</p> <p>Patient Population: Medicare beneficiaries requiring home health services</p>	<p>Utilization measure(s): Acute Care Hospitalization During the First 60 Days of Home Health Use; ED Use without Hospitalization During the First 60 Days of Home Health/ED Use</p> <p>Quality measure(s): Improvement in Dyspnea; Discharged to Community; Improvement in Management of Oral Medications; Total Normalized Composite Change in Self-Care; Total Normalized Composite Change in Mobility</p> <p>Spending measure(s): Home health Medicare claims payments</p> <p>Patient experience measure(s): HHCAHPS Care of Patients/Professional Care; Communication between Providers and Patients; Specific Care Issues/Team Discussion; Overall Rating of Home Health Care; Willingness to Recommend the Agency</p>	<p>How payment is adjusted for performance: Home health agencies receive adjustments to their FFS payments based on their Total Performance Score (TPS), a composite score of an agency's quality measures, relative to peers' performance. Performance on quality measures impacts payment adjustments in a later year.</p> <p>Requirements: CMS will apply a payment adjustment of a maximum of 5% upward or downward in the performance year based on an HHA's performance in the baseline year. When calculating the TPS, quality and utilization measures are weighted 35%, and patient experience measures are weighted 30%.</p> <p>Attribution: N/A</p> <p>Volume: To qualify for payment adjustments, an HHA must have 20 home health quality episodes, 20 home health stays, and 40 completed HHCAHPS surveys, along with sufficient data to calculate at least five of the 12 quality measures in the baseline and performance years</p> <p>Risk stratification or adjustment: Cohorts are determined by the HHA's beneficiary count in the prior calendar year: large-volume (HHAs with 60+ beneficiaries in the calendar year prior to the performance year) or small-volume (HHAs with <60 beneficiaries); cohorts group HHAs of similar size and likelihood to receive scores on the same set of measures.</p> <p>Benchmarking: Achievement thresholds and benchmarks are calculated as the 50th and 90th</p>	<p>Measures used for implementation: All listed measures</p> <p>Measures used for monitoring: Same as for implementation</p> <p>How achievement is measured: TPS compared with the achievement threshold, which is the 50th percentile of all HHAs' performance scores for that quality measure within the cohort during the baseline year, with achievement points being assigned for values that are greater than the threshold.</p> <p>How improvement is measured: Measure values are compared with the improvement threshold, with achievement points being assigned for values that are greater than the threshold.</p>	<p>N/A (Model has not been evaluated yet)</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			percentiles for each quality measure by cohort, based on all HHAs' performance data in the designated baseline year.		
<p>Global and Professional Direct Contracting (GPDC)/Accountable Care Organization Realizing Equity, Access, and Community Health (ACO REACH)</p> <p><i>Participants Announced</i></p> <p>Years active: 2021-present^{xxxix}</p>	<p>Clinical Focus: Primary and specialty care</p> <p>Providers: Direct Contracting Entities (DCEs) under GPDC, ACOs under ACO REACH; Participating and Preferred Providers</p> <p>Setting: Broad applicability</p> <p>Patient Population: Medicare FFS beneficiaries; patients with complex chronic diseases and serious illnesses</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Days at home for patients with complex, chronic Conditions (for high-needs DCEs only); timely follow-up (TFU) after acute exacerbations of chronic conditions (for Standard and New Entrant DCEs only); risk-standardized all condition readmission (ACR) measure; all-cause unplanned admissions for patients with multiple chronic conditions (UAMCC)</p> <p>Spending measure(s): Total cost of care</p> <p>Patient experience measure(s): Patient experience measures from CAHPS, patient/caregiver experience</p>	<p>How payment is adjusted for performance: CMS calculates the total cost of care at the end of the performance year. If the payments and additional FFS Medicare expenditures exceed the performance year benchmark, the DCE/ACO repays CMS the shared losses according to its risk sharing arrangement; otherwise, CMS pays shared savings to the DCE/ACO. Advanced Payment Option (APO) payments are also reconciled in a similar manner.</p> <p>Requirements: To earn shared savings, participants must exceed the benchmark.</p> <p>Attribution: Prospective voluntary; Prospective Plus voluntary; Prospective, claims-based, primary care providers; Prospective, claims-based, non-primary care^{xl}</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Participants are organized into three types: Standard DCEs/ACOs, New Entrant DCEs/ACOs, and High Needs DCEs/ACOs. Risk adjustment is used to adjust expenditures for beneficiary health risk and establish performance year benchmarks.</p> <p>Benchmarking: Based on historical baseline expenditures and/or ACO REACH/KCC rate book or a blend of historical and regional expenditures or</p>	<p>Measures used for implementation: Initial quality score (IQS), a percentage determined by the four quality measures (ACR, UAMCC, TFU, CAHPS) out of a total of 40 points</p> <p>Measures used for monitoring: Same as for implementation</p> <p>How achievement is measured: Performance on each of the four quality measures (ACR, UAMCC, TFU, CAHPS) on a scale of 10 compared with the relevant benchmark</p> <p>How improvement is measured: CMS determines whether REACH ACOs exhibit improvement, no change, or a decline in performance on measure scores for each quality measure</p>	Not specified

^{xxxix} The transition from the GPDC Model to the ACO REACH Model was announced on February 24, 2022. The ACO REACH Model began on January 1, 2023.

^{xl} For additional details on attribution, refer to *Environmental Scan on Improving Management of Care Transitions in Population-Based Models*, available at <https://aspe.hhs.gov/sites/default/files/documents/61e603e1beb3f5eb4d528b1e91fadf12/PTAC-Jun-12-Escan.pdf>.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			regional expenditures, depending on DCE/ACO type and alignment		
Home Health Value-Based Purchasing (HHVBP) Model <i>No Longer Active</i> Years active: 2016-2021	Clinical Focus: Home health care Providers: Medicare-certified Home Health Agencies (HHAs) Setting: Home health setting Patient Population: Medicare beneficiaries requiring home health services	Utilization measure(s): Acute care hospitalization during the first 60 days of home health; ED use without hospitalization during the first 60 days of home health Quality measure(s): Discharge to community; total normalized composite change in self-care; total normalized composite change in mobility; improvement in management of oral medications; improvement in dyspnea Spending measure(s): Home health Medicare claims payments Patient experience measure(s): Patient experience measures from CAHPS surveys	How payment is adjusted for performance: Medicare payments were adjusted upward or downward by up to 3%, 5%, 6%, or 7% based on the Total Performance Score, a composite score of an agency's quality achievement/improvement on the measure set and the performance year. Requirements: Payment adjustments were based on prior TPS (i.e., adjustments began in CY 2018 based on 2016 TPS for up to +/- 3% while CY 2021 is based on 2019 TPS for up to +/- 7%) Attribution: N/A ^{xli} Volume: N/A Risk stratification or adjustment: N/A Benchmarking: Achievement thresholds are based on the median measure value for all HHAs in the state during the baseline period. Benchmarks are based on the mean measure value for the best performing decile of all HHAs in the state during the baseline period.	Measures used for implementation: Total performance score Measures used for monitoring: Same as for implementation How achievement is measured: Total performance score compared against the benchmark How improvement is measured: Percent relative change among total performance score and total normalized composite scores	An evaluation report found that the model had performance gaps among agencies within individual states or regions that may have implications for health equity. A larger gap between patients with and without Medicaid emerged over time in HHVBP states compared with non-HHVBP states, which translated to a slightly larger widening in the disparity over time in HHVBP states. Evidence of persisting quality gaps based on Medicaid status, as well as race and ethnicity under the original HHVBP Model suggests a need for more targeted initiatives to reduce these pre-existing inequities among home health patients and to align with CMS' Framework for Health Equity.

^{xli} All Medicare-certified HHAs from participating states are included in the HHVBP Model.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
<p>Independence at Home (IAH) Demonstration</p> <p><i>Ongoing</i></p> <p>Years Active: 2011-present</p>	<p>Clinical Focus: Primary care, chronically ill</p> <p>Providers: Primary care providers</p> <p>Setting: Home-based</p> <p>Patient Population: Medicare beneficiaries with multiple chronic conditions</p>	<p>Utilization measure(s): Hospital admissions for ambulatory care sensitive conditions; ED visits for ambulatory care sensitive conditions</p> <p>Quality measure(s): Follow-up contact within 48 hours of a hospital admission, hospital discharge, or ED visit; Medication reconciliation in the home within 48 hours of a hospital discharge or ED visit; Annual documentation of patient preferences; All-cause hospital readmissions within 30 days</p> <p>Spending measure(s): Medicare Part A and Part B FFS expenditures</p> <p>Patient experience measure(s): N/A</p>	<p>How payment is adjusted for performance: Practices can receive 50% of shared savings for meeting/exceeding performance requirements on three measures, 66.7% of shared savings for four measures, 83.3% for five measures, and 100% for all six measures.</p> <p>Requirements: Practices actual expenditures must be at least 5% below the target, and they must meet or exceed performance requirements on at least three of the six quality measures to share in savings through an incentive payment.</p> <p>Attribution: N/A</p> <p>Volume: Must serve at least 200 eligible beneficiaries</p> <p>Risk stratification or adjustment: Benchmarks are adjusted to reflect the average hierarchical condition category (HCC) risk score, the average frailty score, and a utilization factor for the IAH population in each practice.</p> <p>Benchmarking: Use a revised actuarial methodology that generates practice-specific per beneficiary per month (PBPM) target expenditures based on historical Medicare FFS per capita expenditures for the Medicare FFS population in the same counties as IAH beneficiaries.</p>	<p>Measures used for implementation: All listed measures</p> <p>Measures used for monitoring: All listed measures</p> <p>How achievement is measured: CMS compares practices' performance for each quality measure against the thresholds.</p> <p>How improvement is measured: CMS compares practices' performance on quality measures across years.</p>	<p>Few practices met the threshold for the 48-hour follow-up visit measure. This could be due to a variety of factors, such as whether the practice receives timely notification of a beneficiary having a hospital admission or ED visit and whether the practice had clinicians who made after-hours and weekend visits.</p>
<p>Kidney Care Choices (KCC) Model</p> <p><i>Ongoing</i></p> <p>Years active: 2020-present</p>	<p>Clinical Focus: ESRD</p> <p>Providers: Accountable care/dialysis facilities,</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Gains in Patient Activation Scores at 12 Months; Depression Remission at 12 Months; Optimal ESRD Starts</p>	<p>How payment is adjusted for performance: Kidney Care First (KCF) Option – adjusted capitated payments for managing care of aligned beneficiaries and for those on dialysis based on health outcomes and utilization relative to both the participants' own experience and national standards, and performance on quality measures. KCF Practices can also receive</p>	<p>Measures used for implementation: Relative performance on quality and cost measures</p>	<p>N/A (Model has not been evaluated yet)</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
	<p>nephrologists, and other health care providers form ESRD-focused ACOs (Kidney Contracting Entities^{xliii} [KCEs])</p> <p>Setting: Dialysis facilities</p> <p>Patient Population: Patients with ESRD</p>	<p>Spending measure(s): CKD Cost of Care; ESRD Cost of Care</p> <p>Patient experience measure(s): N/A</p>	<p>bonus payments for successful kidney transplants; Comprehensive Kidney Care Contracting (CKCC) Graduated, Professional and Global Options – Kidney Contracting Entities (KCEs) receive adjusted capitation payments and can receive a portion or all of the Medicare savings they achieve based on their total cost of care compared with the benchmark, as well as their performance on a set of quality measures. CKCC Graduated Option is a one-sided risk track; CKCC Professional Option provides an opportunity to earn 50% of shared saving or be liable for 50% of shared losses; CKCC Global Option participants are eligible to share in 100% of earnings or losses.</p> <p>Requirements: Transplant bonuses are contingent on success; Level 1 KCEs are subject to a minimum savings rate determined by the volume of beneficiaries needed for statistical confidence.</p> <p>Attribution: Alignment based on where beneficiary receives the majority of their kidney care. When an aligned beneficiary receives a kidney transplant, they will remain aligned to participant for the following three years (if successful; otherwise, they could be re-aligned).</p> <p>Volume: KCF Practices must have a minimum of 350 aligned Medicare beneficiaries with CKD Stages 4 or 5 and 200 aligned ESRD beneficiaries. For the CKCC Model Options, KCEs must have a minimum of 750 beneficiaries with CKD Stages 4 or 5 and 350 ESRD beneficiaries.</p>	<p>Measures used for monitoring: Relative and continuous improvement performance on quality and cost measures</p> <p>How achievement is measured: KCEs are subject to quality withholds; KCEs earn back the withhold based on their quality score (compared with a relevant benchmark). KCEs that achieve a 100% quality score earn back the full quality withhold.</p> <p>How improvement is measured: Continuous improvement performance on cost and utilization measures is defined when KCE participants outperform their own past performance.</p>	

^{xliii} Nephrology practices and their nephrologists and nephrology professionals who meet certain eligibility requirements can participate in the Kidney Care First (KCF) Option. KCEs can participate in any of the Comprehensive Kidney Care Contracting (CKCC) Options and are required to include nephrologists or nephrology practices and transplant providers; optional participants in KCEs include dialysis facilities and other suppliers and providers.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			<p>Risk stratification or adjustment: The KCF Option does not risk-adjust. The CKCC Options use the CMS-HCC prospective risk adjustment model for beneficiaries with late-stage CKD and the CMS-HCC ESRD risk adjustment model for beneficiaries with ESRD to establish benchmarks. For KCEs, a KCE-level symmetric cap on risk score growth is applied.</p> <p>Benchmarking: Based on historical baseline expenditures, prospectively trended forward each performance year (PY) using the projected U.S. per capita cost (USPCC).</p>		
<p>Making Care Primary (MCP) Model <i>Accepting Applications</i> Years active: Launching in July 2024</p>	<p>Clinical Focus: Primary care Providers: PCPs Setting: Primary care practices Patient Population: All Medicare beneficiaries in participating regions</p>	<p>Utilization measure(s): Emergency Department Utilization</p> <p>Quality measure(s): Controlling High Blood Pressure; Diabetes Hba1C Poor Control; Colorectal Cancer Screening; Screening for Depression and Follow-Up; Depression Remission within 12 months; Screening for Social Drivers of Health</p> <p>Spending measure(s): Total Per Capita Cost (TPCC), reported as an observed-to-expected ratio for each participant, dividing the observed cost by the expected cost</p> <p>Patient experience measure(s): Person-Centered Primary Care Measure (PCPCM)</p>	<p>How payment is adjusted for performance: Participants are eligible to receive upside-only Performance Incentive Payments (PIPs) that reward participants for improving patient health outcomes and achieving savings.</p> <p>Requirements: Participants must report all quality measures for their Track (Tracks 1, 2, and 3) to receive a PIP adjustment. Once in Tracks 2 and 3, participants must meet or exceed the 30th percentile nationally for TPCC.</p> <p>Attribution: Eligible beneficiaries are prospectively attributed to a participant. Attribution is first determined by CMS based on the beneficiary's chosen alignment to a clinician. Otherwise, CMS will attribute the beneficiary to the participant if one or more of the participant's eligible clinicians provided the plurality of the beneficiary's primary care visits and/or eligible Chronic Care Management (CCM) services, or if one of the participant's eligible clinicians billed the beneficiary's most recent claim for an Annual Wellness Visit or a Welcome to Medicare Visit during the most recently available 24-month period.</p>	<p>Measures used for implementation: TPCC Continuous Improvement (CI) measure for non-Federally Qualified Health Centers (FQHCs) and non-Indian Health Programs; Emergency Department Utilization CI for FQHCs and Indian Health Programs, as well as the quality and patient experience measures listed.</p> <p>Measures used for monitoring: CMS will use self-reported participants care delivery and financial information, MCP Clinician Lists, Specialty Care Partner Lists, claims, utilization, and quality data in its monitoring strategy. This information is collected through annual participant reporting.</p> <p>How achievement is measured: Scoring and benchmarking</p>	<p>Given the length of the model (10.5 years), CMS may update the measure set in future performance years as new and innovative ways to measure quality and performance become available.</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			<p>Volume: N/A</p> <p>Risk stratification or adjustment: Some measures within the MCP performance measure set are risk-adjusted; however, the model does not employ additional adjustments to inform the PIP.</p> <p>Benchmarking: Continuous Improvement Measures assess performance against participants' own historical performance. Other measures use regional or national benchmarks</p>	<p>structure allows participants to achieve tiered levels of success by providing stepped incentives via lower and upper thresholds for receipt of PIPs. Lower thresholds (to receive partial credit) and upper thresholds (to receive full credit) are set for the clinical quality, cost, and utilization measures.</p> <p>How improvement is measured: Measures of participant-level CI carry significant weight in the PIP for Tracks 2 and 3. For the TPCC CI measure, participants are measured against their own TPCC performance in the previous year.</p>	
<p>Next Generation Accountable Care Organization (NGACO)</p> <p><i>No Longer Active</i></p> <p>Years active: 2016-2021</p>	<p>Clinical Focus: Primary and specialty care</p> <p>Providers: Participating PCPs and specialists</p> <p>Setting: Primary and specialty care practices, hospitals, inpatient and outpatient settings</p> <p>Patient Population: Original Medicare FFS beneficiaries</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Risk Standardized, All Condition Readmissions; Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions; Days at Home</p> <p>Spending measure(s): Total Medicare Parts A and B spending</p> <p>Patient experience measure(s): CAHPS measures, including Getting Timely Care, Appointments, and Information; How Well Your Doctors Communicate; Patients' Rating of Doctor; Access to Specialists; Health Promotion and Education; Shared Decision Making; Health Status/Functional</p>	<p>How payment is adjusted for performance: The benchmarking methodology rewards NGACOs for favorable financial performance on spending relative to historical or regional benchmarks. NGACOs participate in shared savings or losses based on performance year expenditures. NGACOs may receive an Earned Quality Bonus for meeting quality requirements. CMS uses a quality "withhold," in which a portion of an ACO's performance year benchmark is held "at-risk," contingent upon the ACO's quality score. An ACO that achieves a 100% quality score (a function of the ACO meeting quality measure benchmarks and reporting requirements) will have the full withhold re-attributed to its performance year benchmark at settlement, while an ACO that achieves less than a 100% quality score will have a proportionate amount discounted from the withhold and re-attributed to its performance year benchmark.</p>	<p>Measures used for implementation: The quality score is based on three measures: hospitalizations for ambulatory care sensitive conditions (ACSCs), 30-day hospital readmissions, and 30-day hospital readmissions from an SNF.</p> <p>Measures used for monitoring: Quality and patient experience measures</p> <p>How achievement is measured: CMS utilized quality scores and measured spending against a</p>	<p>An EHR meaningful use measure was dropped from the model: it was expected that ACOs who were ready and able to take on high levels of risk under the NGACO Model were already using electronic health records (EHRs) and already had robust systems in place.</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
		Status; Stewardship of Patient Resources; Courteous and Helpful Office Staff	<p>Requirements: NGACOs must meet minimum quality requirements to participate in shared savings and the Earned Quality Bonus. If an NGACO does not meet the minimum requirements, it is still required to pay losses.</p> <p>Attribution: Voluntary: Beneficiaries confirm care relationships with participating providers (annual).</p> <p>Prospective, claims-based: Beneficiaries are aligned to the participating provider that provided the majority of that beneficiary’s evaluation and management (E&M) visits (annual).</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Risk scores are used to adjust PY benchmarks, accounting for health status differences between beneficiaries from the baseline period and performance period. The discount that is applied to risk-adjusted baselines is dependent on whether the ACO selects a full or partial risk arrangement. Changes were made to the risk adjustment methodology as a result of COVID-19.</p> <p>Benchmarking: Yes, prospectively set based on historical expenditures and national trends</p>	<p>benchmark based on historical and regional trends.</p> <p>How improvement is measured: The model offers financial arrangements with higher levels of risk and reward, using refined benchmarking methods that reward both attainment of and improvement in cost containment.</p>	
Oncology Care Model (OCM) <i>No Longer Active</i>	Clinical Focus: Cancer care Providers: Oncology providers	Utilization measure(s): N/A Quality measure(s): Screening for Depression and Follow-up Plan; Oncology: Medical and Radiation – Plan of Care for Pain; Oncology: Medical and Radiation –	How payment is adjusted for performance: The amount of the performance-based payment is adjusted based on the participant’s achievement on a range of quality measures. OCM quality measure data derived from claims, aggregate measure results reported to the	Measures used for implementation: Risk-adjusted proportion of patients with all-cause hospital admissions within the 6-month episode; Risk-adjusted proportion of patients	Some practices raised concerns related to the lack of specificity in the validated measures for depression. Further,

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
Years active: 2016-2022	<p>Setting: Outpatient</p> <p>Patient Population: Patients with cancer</p>	<p>Pain Intensity Quantified; Proportion of Patients who Died who Were Admitted to Hospice for 3 Days or More; Risk-adjusted Proportion of Patients with All-Cause ED Visits or Observation Stays that did not Result in a Hospital Admission within the 6-month Episode; Risk-adjusted proportion of patients with all-cause hospital admissions within the 6-month episode; Closing the Referral Loop: Receipt of Specialist Report; Documentation of Current Medications in the Medical Record; Prostate Cancer: Adjuvant Hormonal Therapy for High or Very High Risk Prostate Cancer; Adjuvant chemotherapy is recommended or administered within 4 months (120 days) of diagnosis to patients under the age of 80 with AJCC III (lymph node positive) colon cancer; Combination chemotherapy is recommended or administered within 4 months (120 days) of diagnosis for women under 70 with AJCC T1cN0M0, or Stage IB - III hormone receptor negative breast cancer; Trastuzumab administered to patients with AJCC stage I (T1c) - III and human epidermal growth factor receptor 2 (HER2) positive breast cancer who receive adjuvant chemotherapy; Breast Cancer: Hormonal Therapy for Stage I (T1b)-IIIC Estrogen Receptor/Progesterone Receptor (ER/PR) Positive Breast Cancer</p> <p>Spending measure(s): All Medicare Part A and Part B expenditures, certain Part D expenditures</p> <p>Patient experience measure(s): Patient-Reported Experience of Care</p>	<p>OCM Data Registry, and patient experience survey data are utilized to determine the quality score used in calculation of the performance multiplier. Scoring, or the process of assigning quality points to each quality measure, is based on the OCM participants' reporting of quality measure data and/or quality performance relative to set thresholds. Once quality points are assigned, an Aggregate Quality Score (AQS) will be calculated and translated into a performance multiplier. This performance multiplier is used as part of the performance-based payment calculation.</p> <p>Requirements: To be eligible for performance-based payment, practices opted-in for the performance period must have a target amount that exceeds actual episode expenditures of the episodes attributed to the practice. The practice or pool will also need to meet or exceed an AQS minimum threshold of 30%. Further, the practice or pool must report aggregate measure results for practice-reported measures, and implement all Practice Redesign Activities to receive a PBP.</p> <p>Attribution: Each episode is attributed to an individual practice. Visits are credited to the tax identification number (TIN) or OCM ID (for OCM practices) with the majority of the beneficiary's cancer-related E&M visits during the 6-month episode. For pooled participants, episodes are still attributed to individual practices within the pool.</p> <p>Volume: N/A; performance-based payments will be made only for higher-volume cancer types for which it is possible to calculate accurate benchmarks. Benchmarks will not be calculated for lower-volume cancer types.</p>	<p>with all-cause emergency department visits or observation stays that did not result in a hospital admission within the 6-month episode; Proportion of patients who died who were admitted to hospice for 3 days or more; Care Plan; Closing the Referral Loop: Receipt of Specialist Report; Pain Assessment and Management Composite; Preventive Care and Screening: Screening for Depression and Follow-Up Plan; Patient-Reported Experience of Care; Prostate Cancer: Adjuvant Hormonal Therapy for High or Very High Risk Prostate Cancer; Adjuvant chemotherapy is recommended or administered within 4 months (120 days) of diagnosis to patients under the age of 80 with AJCC III (lymph node positive) colon cancer; Combination chemotherapy is recommended or administered within 4 months (120 days) of diagnosis for women under 70 with AJCC T1cN0M0, or Stage IB - III hormone receptor negative breast cancer; Trastuzumab administered to patients with AJCC stage I (T1c) - III and human epidermal growth factor receptor 2 (HER2) positive breast cancer who receive adjuvant chemotherapy; Breast Cancer: Hormonal Therapy for Stage I</p>	<p>practices with multiple EHRs across clinics/sites reported difficulty reporting quality measures at the practice level.</p>

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			<p>Risk stratification or adjustment: Benchmark prices are risk-adjusted for factors that affect episodic expenditures and that are available in Medicare claims data, including age, sex, dual eligibility for Medicaid and Medicare, selected non-cancer comorbidities, receipt of selected cancer-directed surgeries, receipt of bone marrow transplant, receipt of radiation therapy, type of chemotherapy drugs used during episode (for breast, prostate, and bladder cancers only), institutional status, participation in a clinical trial, history of prior chemotherapy use, episode length, and hospital referral region. The risk-adjusted ED visit and observation stay measure is risk-adjusted using a hierarchical logistic regression model that incorporates many of the same risk adjustment variables used in setting episode spending benchmarks, including cancer type, demographics, institutional status, geographic location, and comorbidities.</p> <p>Benchmarking: CMS calculates benchmark episode expenditures for OCM practices based on historical data. Benchmarks are adjusted for risk (e.g., patient sociodemographics, clinical characteristics) and geographic variation, trended to the applicable performance period, and include a novel therapies adjustment. A discount is applied to the benchmark to determine a target price for OCM-FFS episodes. Pay-for-performance measures are assigned quality points based on the practice's or pool's performance as compared with set thresholds, called quality benchmarks.</p>	<p>(T1b)-IIC Estrogen Receptor/Progesterone Receptor (ER/PR) Positive Breast Cancer; Documentation of Current Medications in the Medical Record</p> <p>Measures used for monitoring: Shared Decision-Making composite measure; claims data and practice-reported quality measure and clinical data.</p> <p>How achievement is measured: Practices earn quality points that contribute to an Aggregate Quality Score (AQS) to measure performance. Practice-reported expenditure and quality measures are compared against thresholds.</p> <p>How improvement is measured: N/A</p>	

Model Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
<p>Primary Care First (PCF) Model Options</p> <p><i>Ongoing</i></p> <p>Years active: 2021-present</p>	<p>Clinical Focus: Primary care</p> <p>Providers: PCPs</p> <p>Setting: Primary care practices</p> <p>Patient Population: Medicare patients with serious illness/chronic conditions</p>	<p>Utilization measure(s): Acute hospital utilization (the overall observed-to-expected ratio of acute inpatient and observation stay discharges)</p> <p>Quality measure(s): Diabetes: Hemoglobin A1c (HbA1c) Poor Control (> 9%); Controlling High Blood Pressure; Colorectal Cancer Screening; Advance Care Plan</p> <p>Spending measure(s): Total per capita cost</p> <p>Patient experience measure(s): Patient Experience of Care Survey (Consumer Assessment of Healthcare Providers and Systems [CAHPS])</p>	<p>How payment is adjusted for performance: A practice's payment amount depends on its performance compared with peer practices and its degree of improvement compared with its historical performance. Performance-based payment can be up to a 50% increase or a 10% decrease in total primary care payment revenue.</p> <p>Requirements: To be eligible for a performance-based payment, practices must meet minimum performance thresholds for quality measures, and their practice performance must be above 50th percentile of the National Benchmark. Performance thresholds and measures vary by risk group.</p> <p>Attribution: CMS identifies eligible beneficiaries using Medicare administrative data and conducts voluntary alignment. CMS also uses a claims-based attribution approach to identify other eligible beneficiaries.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: CMS assigns practices to four risk groups using average CMS-HCC risk scores of their attributed beneficiaries.</p> <p>Benchmarking: Benchmarks are based on the 30th percentile and use either the 2022 MIPS benchmark population or a CMS-derived 2021 national benchmark population for performance year 2023.</p>	<p>Measures used for implementation: Acute hospital utilization, HbA1c poor control, controlling high blood pressure, colorectal cancer screening, advance care plan, total per capita cost, and CAHPS</p> <p>Measures used for monitoring: Same as for implementation</p> <p>How achievement is measured: The practice's utilization measure (for risk groups 1 and 2) or spending measure (for risk groups 3 and 4) performance is compared with that of its peer region group, leading to a rating of 1-7 that corresponds to its performance-based payment percentage adjustment. To pass the Quality Gateway, risk group 1 and 2 practices must exceed the 30th percentile for all quality and patient experience measures. Risk group 3 and 4 practices must exceed the 30th percentile for the advance care plan and CAHPS measures to pass the Quality Gateway.</p> <p>How improvement is measured: Practices can also receive an improvement bonus for improving their performance over time (percent improvement).</p>	<p>Among the 13 PCF participants assessed in the first evaluation report, only five of the 13 payer partners offered financial incentives tied to outcome measures, using cost and utilization metrics. Several practices expressed concerns about the timeliness and quality of beneficiary data provided through PCF.</p>

Exhibit E2. Lessons Learned Related to Current Performance Measures for Selected CMMI Models

Model Name	Clinical Focus, Providers, Setting, Patient Population	Lessons Learned Related to Performance Measurement
<p>Bundled Payments for Care Improvement Advanced (BPCI-A) <i>Ongoing</i> Years active: 2018-present</p>	<p>Clinical Focus: Cross-clinical focus Providers: Acute care hospitals, physician group practices, Medicare-enrolled providers Setting: Inpatient and outpatient services Patient Population: Medicare beneficiaries with certain clinical episodes (29 inpatient, three outpatient)</p>	<p>The model reduced total episode payments, institutional post-acute care (PAC) payments, discharges to institutional PAC settings, and the number of skilled nursing facility (SNF) days among patients who received SNF care relative to the comparison group.</p>
<p>Comprehensive ESRD Care (CEC) Model <i>No longer active</i> Years active: 2015-2021</p>	<p>Clinical Focus: End-stage renal disease (ESRD) Providers: Nephrologists; ESRD Seamless Care Organizations (ESCOs)^{xliii} Setting: Nephrology clinics Patient Population: Medicare beneficiaries with ESRD</p>	<p>An evaluation after performance year 5 (evaluation through December 31, 2020)⁴⁷¹ found there was a 3% decrease in the number of hospitalizations and a 0.4% increase in the number of outpatient dialysis sessions. Additionally, CEC beneficiaries experienced 5% fewer hospitalizations from ESRD complications and were 5% less likely to use a catheter compared with non-CEC beneficiaries. The model also improved phosphate binder adherence by 9%. From PY1-PY5, the model reduced Medicare spending by \$217 million (1.3%) prior to shared savings payments; however, most of this was attributed to Wave 1 ESCOs.</p>
<p>Comprehensive Primary Care Plus (CPC+) <i>No longer active</i> Years active: 2017-2021</p>	<p>Clinical Focus: Primary care Providers: Primary care providers (PCPs) Setting: Primary care practice Patient Population: All Medicare and Medicaid beneficiaries in participating regions</p>	<p>The fourth annual report⁴⁷² found a reduction in acute care utilization and improvement in some quality-of-care measures over the first four years of implementation. However, CPC+ did not reduce Medicare expenditures without enhanced payments; expenditures including enhanced payments increased by 1.5% in Track 1 and 2.6% in Track 2. There was a 1.5% reduction in total Medicare expenditures in PY4 for Track 1 Shared Savings Program practices. Although 95% of practices reported an improvement in care quality, there was little evidence that CPC+ improved continuity, fragmentation, comprehensiveness of care, 30-day unplanned readmissions, or mortality; average rates of emergency department (ED) and hospital follow-up increased from 2017 to 2019. CPC+ practices engaged in transformation activities to improve quality of care. Transformation activities may be in domains of access and continuity (e.g., 24/7 patient access), care management (e.g., care plans for high-risk chronic disease patients), comprehensive and coordinated care (e.g., behavioral health integration), patient and caregiver engagement (e.g., convening a patient and family advisory council), and data-driven population health management (e.g., weekly care team review of population health data). Other contextual factors, such as SDOH and patient preferences, could limit the degree that patients engage with improved primary care and therefore alter their behavior and outcomes.</p>
<p>Enhancing Oncology Model (EOM) <i>Ongoing</i> Years active: 2022-present</p>	<p>Clinical Focus: Oncology Providers: Oncologists Setting: Oncology practices Patient Population: Medicare beneficiaries with cancer</p>	<p>EOM builds on lessons learned from the Oncology Care Model (OCM) and shares certain features with the OCM, including a similar quality measure set. The EOM Model performance period began in July 2023. Model evaluations have not been completed yet. Thus, there are no lessons learned related to performance measurement.</p>

^{xliii} ESCOs comprise nephrologists, dialysis facilities, and other providers.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Lessons Learned Related to Performance Measurement
<p>ESRD Treatment Choices (ETC) Model</p> <p><i>Ongoing</i></p> <p>Years active: 2021-present</p>	<p>Clinical Focus: Home dialysis and kidney transplants for patient with ESRD</p> <p>Providers: Nephrologists</p> <p>Setting: ESRD facilities, transplant centers, large donor hospitals, patient home</p> <p>Patient Population: Patients with ESRD</p>	<p>Year 2 results⁴⁷³ showed 66% of ESRD facilities and 78% of managing clinicians had higher home dialysis rates, and 48% of facilities and 51% of managing clinicians had higher transplant rates than the respective benchmarks (defined as the 50th percentile). Additionally, 80% of facilities and 70% of managing clinicians improved their home dialysis rate, and 61% of facilities and 62% of managing clinicians improved their transplant rates from benchmark year 2 (January 1, 2020 – December 31, 2020) to measurement year (July 1, 2021 – June 30, 2022). Approximately half of ESRD facilities (46%) and managing clinicians (54%) received a positive performance payment adjustment.</p>
<p>Expanded Home Health Value-Based Purchasing Model (Expanded HHVBP)</p> <p><i>Ongoing</i></p> <p>Years active: January 2022-present</p>	<p>Clinical Focus: Home health care</p> <p>Providers: Medicare-certified Home Health Agencies (HHAs)</p> <p>Setting: Home health setting</p> <p>Patient Population: Medicare beneficiaries requiring home health services</p>	<p>Model evaluations have not been completed yet.</p>
<p>Global and Professional Direct Contracting (GPDC)/Accountable Care Organization</p> <p>Realizing Equity, Access, and Community Health (ACO REACH)</p> <p><i>Participants Announced</i></p> <p>Years active: 2021-present^{xliv}</p>	<p>Clinical Focus: Primary and specialty care</p> <p>Providers: Direct Contracting Entities (DCEs) under GPDC, ACOs under ACO REACH; Participating and Preferred Providers</p> <p>Setting: Broad applicability</p> <p>Patient Population: Medicare FFS beneficiaries; patients with complex chronic diseases and serious illnesses</p>	<p>A summary of GPDC Model performance of the first performance year and first three-quarters of the second PY⁴⁷⁴, last updated April 2023, compared all-condition readmissions (ACR) and unplanned admissions for patients with multiple chronic conditions (UAMCC) per 100 beneficiaries with multiple chronic conditions per year. The UAMCC score showed GPDC Model participants performed better than practices who did not participate in the model (30.65 versus 32.54), but the ACR scores among participants and non-participants were similar (15.21% versus 15.18%). The total dollars under risk (the sum of the performance year benchmark across the 99 PY2022 DCEs) were consistent with the average per-beneficiary-per-month benchmark of approximately \$1,117. Across the 99 DCEs that participated in PY2022, they saw roughly a 1.6% reduction in Medicare spending compared with their combined benchmarks in PY2022.</p> <p>As of the first evaluation report, DCEs' most highly prioritized strategies for population health management focused on avoidable utilization (90%), complex or population-specific care management (90%), and investments in primary care (63%). While there was no significant impact on gross or net expenditures for Standard or New Entrant DCEs in PY2021, Standard DCEs significantly reduced acute care hospitalizations and skilled nursing facility days, and both Standard and New Entrant DCEs significantly reduced ED visits. Standard DCEs also reduced hospitalizations for ambulatory care sensitive conditions.</p>

^{xliv} The transition from the GPDC Model to the ACO REACH Model was announced on February 24, 2022. The ACO REACH Model began on January 1, 2023.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Lessons Learned Related to Performance Measurement
Home Health Value-Based Purchasing (HHVBP) Model <i>No Longer Active</i> Years active: 2016-2021	Clinical Focus: Home health care Providers: Medicare-certified Home Health Agencies (HHA) Setting: Home health setting Patient Population: Medicare beneficiaries requiring home health services	The seventh annual report found HHVBP led to decreases in unplanned hospitalizations (-1.2%), skilled nursing facility use (-8.2%), and ED use followed by inpatient admissions (-1.5%), but an increase in outpatient ED visits (2.1%). ⁴⁷⁵ In all model years (2016-2021), home health agencies in participating states received higher TPS scores than agencies in non-HHVBP states and throughout the four payment years. The model also resulted in small improvements in patients' mobility, management of oral medicine, and self-care, and a greater proportion of patients were discharged to the community rather than institutional care. However, patient experience with professional care, communication, and discussion of care decreased during the model. The HHVBP model led to savings in Part A and Part B spending in all model years for a total of \$1.38 billion in savings. Reductions were attributed to reduced spending on skilled nursing facility services, inpatient hospital stays, and home health spending but increased spending for outpatient ED and observational stays.
Independence at Home (IAH) Demonstration <i>Ongoing</i> Years Active: 2011-present	Clinical Focus: Primary care, chronically ill Providers: Primary care providers Setting: Home-based Patient Population: Medicare beneficiaries with multiple chronic conditions	An evaluation of performance years 1-7 ⁴⁷⁶ found the majority of practices reported they were already providing care that was largely consistent with the demonstration requirements and quality measures before participating in the program. Several practices also reported these guidelines became their standard of care for all patients. Nearly all 10 practices met the required threshold for claims-based measures (hospital admissions, ED visits, all-cause hospital readmissions), making them eligible for incentive payments, but most fell short of the threshold for site-reported measures (documenting patient preferences, follow-up within 48 hours, medication reconciliation within 48 hours). There was some improvement in practices documenting patient preferences between years 6 and 7; however, performance remained low for the two measures requiring timely follow-up.
Kidney Care Choices (KCC) Model <i>Ongoing</i> Years active: 2020-present	Clinical Focus: ESRD Providers: Accountable care/dialysis facilities, nephrologists, and other health care providers form ESRD-focused ACOs (Kidney Contracting Entities [KCEs]) ^{xlv} Setting: Dialysis facilities Patient Population: Patients with ESRD	Model evaluations have not been completed yet.
Making Care Primary (MCP) Model <i>Accepting Applications</i> Years active: Launching in July 2024	Clinical Focus: Primary care Providers: PCPs Setting: Primary care practices Patient Population: All Medicare beneficiaries in participating regions	MCP is a new, 10.5-year CMMI advanced primary care model that aims to reduce program expenditures and improve key measures of patient outcomes through more coordinated and integrated care. The model is currently accepting applications and does not yet have participants. Model evaluations have not been completed yet. Thus, there are no lessons learned related to performance measurement.

^{xlv} Nephrology practices and their nephrologists and nephrology professionals who meet certain eligibility requirements can participate in the Kidney Care First (KCF) Option. KCEs can participate in any of the Comprehensive Kidney Care Contracting (CKCC) Options and are required to include nephrologists or nephrology practices and transplant providers; optional participants in KCEs include dialysis facilities and other suppliers and providers.

Model Name	Clinical Focus, Providers, Setting, Patient Population	Lessons Learned Related to Performance Measurement
<p>Next Generation Accountable Care Organization (NGACO)</p> <p><i>No Longer Active</i></p> <p>Years active: 2016-2021</p>	<p>Clinical Focus: Primary and specialty care</p> <p>Providers: Participating PCPs and specialists</p> <p>Setting: Primary and specialty care practices, hospitals, inpatient and outpatient settings</p> <p>Patient Population: Original Medicare FFS beneficiaries</p>	<p>An evaluation after the fifth performance year⁴⁷⁷ found that the NGACO Model was associated with \$1.05B in gross savings, or 1.5%, compared with FFS beneficiaries in the comparison group. In PY5, the NGACO Model reduced gross spending by 3.1%. However, cumulative net Medicare spending increased by \$386.5M after taking into consideration shared savings and payments to the NGACOs. Additionally, some NGACOs reported increased resilience and improved response to the COVID-19 Public Health Emergency as a result of infrastructure, partnerships, and resources developed from participation in the NGACO Model. The NGACOs also saw a reduction in hospital spending and utilization, SNF stays and days, and spending in institutional post-acute care (PAC) settings.</p> <p>Larger spending and utilization reductions in PY5 are likely to have been a result of the selection effect of less successful NGACOs exiting the model. The NGACOs continuing in the model earned shared savings throughout the model implementation, resulting in larger payouts and continuous improvement.</p> <p>Annual wellness visits (AWVs) were measured under the evaluation; NGACO leadership reported using AWVs and IT infrastructure to identify and address gaps in care.</p>
<p>Oncology Care Model (OCM)</p> <p><i>No Longer Active</i></p> <p>Years active: 2016-2022</p>	<p>Clinical Focus: Cancer care</p> <p>Providers: Oncology providers</p> <p>Setting: Outpatient</p> <p>Patient Population: Patients with cancer</p>	<p>An evaluation after the ninth performance period found that OCM was associated with a reduction in total episode payment (TEP) by 1.7%. However, net losses for Medicare exceeded \$500M. OCM increased the use of high-value supportive therapies. In terms of health equity, the report did not find evidence of improved care quality for Black, Hispanic, or dual-eligible beneficiaries across the measures included in the analysis. There was no effect on the timeliness of post-chemotherapy surgery, patient adherence to oral cancer regimens, or provision of higher-value palliative radiation.</p>
<p>Primary Care First (PCF) Model Options</p> <p><i>Ongoing</i></p> <p>Years active: 2021-present</p>	<p>Clinical Focus: Primary care</p> <p>Providers: PCPs</p> <p>Setting: Primary care practices</p> <p>Patient Population: Medicare patients with serious illness/chronic conditions</p>	<p>As of the first annual evaluation report,⁴⁷⁸ most practices met benchmarks for diabetes control, high blood pressure control, and colorectal cancer screening. In addition, qualitative work showed that practices found advanced care plans to be burdensome and costly. Some practices noted that their previous success in reducing preventable hospital utilization could make it difficult to achieve further reductions. On average, PCF payments were higher than expected FFS payments.</p>

Appendix F. Summary of Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP

The first table provides specific details on Hospital VBP, MA Star Ratings Program, MIPS, and MSSP characteristics (i.e., clinical focus, providers, setting, and patient population); components relevant to performance measurement (i.e., utilization measures, quality measures, spending measures, and patient experience measures); technical issues related to performance measurement (i.e., how payment is adjusted for performance, requirements, volume, risk stratification or adjustment, and benchmarking); how measurement is used to determine success (i.e., measures used for implementation, measures used for monitoring, how achievement is measured, and how improvement is measured); and gaps related to current performance measures, as applicable.^{xlvi} The second table describes lessons learned related to performance measurement. The programs are presented in alphabetical order.

Overview of Methodology Used to Review Hospital VBP, MA Star Ratings Program, MIPS, and MSSP

The available information on the Hospital VBP, MIPS, and MSSP summary pages on the CMS website was reviewed. This included an overview of the program, financial operating and performance measurement methodologies, evaluation reports and findings, summaries, fact sheets, and press releases. To assess the MA Star Ratings Program, CMS fact sheets and evaluation reports and findings were reviewed. Information found in these materials was used to summarize the programs' main themes related to performance measurement and other administrative and payment characteristics. The categorizations were based on the key information highlighted in these documents and are not exhaustive. The programs may have elements that fall into additional categories of context, objective, functions, and payment models.

^{xlvi} For additional information about current performance measures used in these programs, please refer to *Overview of Current Performance Measures Included in Selected Medicare Payment Programs* that can be accessed on the [PTAC Resources webpage](#).

Exhibit F1. Program Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP

Program Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
<p>Hospital Value-Based Purchasing Program (Hospital VBP)</p> <p><i>Ongoing</i></p> <p>Years Active: 2013 – Present</p>	<p>Clinical Focus: Hospital care</p> <p>Providers: Acute care hospitals</p> <p>Setting: Inpatient services</p> <p>Patient Population: Medicare beneficiaries requiring inpatient hospital services</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Acute Myocardial Infarction 30-Day Mortality Rate; Coronary Artery Bypass Graft Surgery 30-Day Mortality Rate; Chronic Obstructive Pulmonary Disease 30-Day Mortality Rate; Heart Failure 30-Day Mortality Rate; Pneumonia 30-Day Mortality Rate; Total Hip Arthroplasty/ Total Knee Arthroplasty Complication Rate</p> <p>Spending measure(s): Medicare severity diagnosis-related group (MS-DRG) payment amounts for Medicare FFS claims</p> <p>Patient experience measure(s): N/A</p>	<p>How payment is adjusted for performance: Under the Inpatient Prospective Payment System (IPPS), payments are adjusted based on a total performance score that reflects relevant benchmarks, for each performance measure.</p> <p>Requirements: The program withholds participants’ payments by 2% percent; this amount is used to fund value-based incentive payments to participants based on performance.</p> <p>Attribution: N/A</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Participants who treat a high percentage of low-income patients receive a disproportionate share hospital (DSH) adjustment, which is a percentage increase in Medicare payments to hospitals that serve a disproportionate amount of low-income patients. Further, participants who are approved teaching hospitals receive an indirect medical education (IME) adjustment. Lastly, for substantially costly/outlier cases, IPPS payment is increased.</p> <p>Benchmarking: Historical national data</p>	<p>Measures used for implementation: All listed quality measures</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: Participants earn two scores on each performance measure: one for achievement, based on how well they perform compared with all hospitals, and one for improvement, based on how much they improve their own performance compared with their performance during baseline. The final score awarded for each measure is the higher of these two scores.</p> <p>How improvement is measured: See how achievement is measured.</p>	<p>The hospital total performance score (TPS) is skewed and shows a large gap between top-performing hospitals and all others. The highest performing hospitals tend to maintain their position over time, which may discourage lower-performing hospitals from program participation and improvement.⁴⁷⁹</p>
<p>Medicare Advantage Star Ratings Program (MA Star Ratings Program)</p> <p><i>Ongoing</i></p> <p>Years Active: 2009 – Present</p>	<p>Clinical Focus: Total care</p> <p>Providers: Medicare Advantage Health Plan Contracts</p> <p>Setting: Broad</p> <p>Patient Population: Medicare</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Special Needs Plan (SNP) Care Management; Statin Therapy for Patients with Cardiovascular Disease; Medication Reconciliation Post-Discharge; Improving Bladder Control; Reviewing Appeals Decisions; Annual Flu Vaccine; Care Coordination; Plan makes Timely Decision about Appeals; Reducing the Risk of Falling; Call Center- Foreign Language Interpreter and TTY Availability;</p>	<p>How payment is adjusted for performance: Star Ratings (based on performance) are used to determine 1) whether a plan is eligible for a bonus payment; and 2) the percentage increase in payment benchmarks and rebate amounts.</p> <p>Requirements: To be eligible for a performance-based bonus payment, health plan contracts must obtain a 4, 4.5, or 5 Star Rating.</p> <p>Attribution: N/A</p>	<p>Measures used for implementation: Quality and patient experience measures listed</p> <p>Measures used for monitoring: Quality and patient experience measures listed and measures that were removed (not listed)</p>	<p>A report from McKinsey found that investments in patient-experience improvements have been low.⁴⁸⁰</p>

Program Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
	<p>Advantage beneficiaries</p>	<p>Colorectal Cancer Screening; Monitoring Physical Activity; Care for Older Adults – Pain Assessment; Care for Older Adults – Medication Review; Breast Cancer Screening; Diabetes Care – Blood Sugar Controlled; Diabetes Care – Eye Exam; Osteoporosis Management for Women who had a Fracture; Controlling Blood Pressure; Plan All-Cause Readmissions; Transitions of Care; Follow-up after Emergency Department Visit for People with Multiple High-Risk Chronic Conditions; Health Plan Quality Improvement</p> <p>Spending measure(s): Medicare FFS spending</p> <p>Patient experience measure(s): Getting Appointments and Care Quickly; Customer Service; Complaints about the Plan; Rating of Health Plan; Rating of Health Care Quality; Getting Needed Care; Members Choosing to Leave the Plan</p>	<p>Volume: N/A; Plans without ratings due to low enrollment receive a small benchmark increase.</p> <p>Risk stratification or adjustment: CMS includes a coding intensity adjustment that accounts for potential increases in the average risk score of MA beneficiaries. Adjustments to Star Ratings may also be made based on impacts due to public health emergencies (PHEs), such as COVID-19.</p> <p>Benchmarking: Benchmarks are calculated separately for each county and are based on historic fee-for-service Medicare spending. Benchmarks are then adjusted by plan quality (Star Rating). Benchmarks are capped at pre-Affordable Care Act (ACA) levels.</p>	<p>How achievement is measured: MA Health Plan Contracts receive a numeric measure score for each measure (up to 40 measures). Measures come from four sources: Healthcare Effectiveness Data and Information Set (HEDIS), Consumer Assessment of Healthcare Providers and Systems (CAHPS), Health Outcomes Survey (HOS), and CMS administrative data. Measures are weighted to reflect CMS priority in scoring MA plans. Process measures receive a weight of 1; patient experience measures receive a weight of 1.5; outcomes and intermediate outcomes a weight of 3; quality improvement measures a weight of 5. CMS calculates Star Rating scores for each measure, as well as a summary Star Rating score based on their performance in five domains: use of screenings, tests, and vaccines; management of chronic conditions; member experience with plans (CAHPS); member complaints and changes in plan’s performance; and customer service/appeals. The numeric measure scores are converted to a Star Rating (from 1- 5) based on one of two methods: clustering or relative</p>	

Program Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
				<p>distribution and significance testing. CMS previously defined quality threshold values that a plan would need to achieve to receive four stars on certain measures. Because analysis determined that plans showed more improvement on measures without pre-determined thresholds, CMS eliminated pre-defined thresholds beginning in 2016.</p> <p>How improvement is measured: Although not tied to payment, improvement is measured by comparing the health plan contract's current and prior year measure scores.</p>	
<p>Merit-based Incentive Payment System (MIPS) <i>Ongoing</i> Years Active: 2017 – Present</p>	<p>Clinical Focus: Total care Providers: Clinicians, practices Setting: Broad Patient Population: Medicare FFS beneficiaries</p>	<p>Utilization measure(s): Hospital-Wide, 30-Day, All-Cause Unplanned Readmission (HWR) Rate for MIPS Eligible Groups; Clinician and Clinician Group Risk-standardized Hospital Admission Rates for Patients with Multiple Chronic Conditions; Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure for MIPS</p> <p>Quality measure(s): The 2023 MIPS Quality Measure List includes over 200 measures,^{xlvii} including Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA) for MIPS</p>	<p>How payment is adjusted for performance: Payment adjustment applied to Medicare Part B claims based on performance. Performance is measured across four areas: quality, improvement activities, promoting interoperability, and cost. Participants receive a MIPS final score based on the four performance categories, which determines the payment adjustment.</p> <p>Requirements: Participants must submit collected data for at least six quality measures and achieve at least 70% data completeness for each quality measure.</p> <p>Attribution: N/A</p> <p>Volume: Beginning in 2023, reporting a measure that does not meet the case minimum will result in 0 out of 10 points (3 points for small practices).</p>	<p>Measures used for implementation: From the measures listed, participants select six measures or a complete specialty measure set. CAHPS for MIPS Survey can count as one of the six minimum required quality measures.</p> <p>Measures used for monitoring: Not specified.</p> <p>How achievement is measured: Participants submit data for quality, improvement activities, and promoting interoperability; CMS calculates cost for</p>	<p>Few MIPS measures relate to aspects of access, patient experience, or physician interpersonal skills, which could contribute to or worsen health disparities.⁴⁸¹</p>

^{xlvii} For the full 2023 MIPS Quality Measure List, see <https://app-cm-prod-content.s3.amazonaws.com/uploads/2217/2023%20MIPS%20Quality%20Measures%20List.xlsx>.

Program Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
		<p>Spending measure(s): Total per capita cost; Medicare spending per beneficiary clinician; episode-based costs for 23 specific conditions</p> <p>Patient experience measure(s): CAHPS for MIPS Survey</p>	<p>Risk stratification or adjustment: Expected costs for each episode use the CMS hierarchical condition category (HCC) risk adjustment methodology; it adjusts for 12 age categorical variables and individual indicator variables for history of ESRD, long-term care status, comorbidities, and whether beneficiary qualifies for Medicare through disability or age.</p> <p>Benchmarking: For each performance measure collection type, CMS establishes a benchmark based on historical data.</p>	<p>participants. CMS automatically calculates and scores four administrative claims measures. Measures are scored by comparing to measure-specific benchmarks. A maximum of 10 points can be earned for each measure. Scores from 75-100 points will receive a payment adjustment ranging from 0-9% (e.g., MIPS score of 100 receives positive payment adjustment of 9%).</p> <p>How improvement is measured: Participants are eligible to earn up to 10 additional percentage points based on improvement in performance from the previous year.</p>	
<p>Medicare Shared Savings Program (MSSP)</p> <p><i>Ongoing</i></p> <p>Years Active: 2012-Present</p>	<p>Clinical Focus: Total care</p> <p>Providers: Providers and suppliers (e.g., physicians, hospitals, and others involved in patient care) that create an Accountable Care Organization (ACO)</p> <p>Setting: Broad</p> <p>Patient Population: Medicare FFS beneficiaries</p>	<p>Utilization measure(s): Not specified</p> <p>Quality measure(s): Depression remission at 12 months; falls: screening for future fall risk; hospital-wide, 30-day, all-cause unplanned readmission (HWR) rate for the MIPS Groups; preventive care and screening: influenza immunization; breast cancer screening; colorectal cancer screening; clinician and clinician group risk-adjusted hospital admission rates for patients with multiple chronic conditions; statin therapy for the prevention and treatment of cardiovascular disease; preventive care and screening: tobacco use: screening and cessation intervention; diabetes hemoglobin A1c (HbA1c) poor control (>9%); preventive care and screening: screening for</p>	<p>How payment is adjusted for performance: ACOs are subject to an annual spending target (benchmark) and a series of quality thresholds. ACOs that spend less than the benchmark share the savings with CMS. There is a penalty for spending more than the threshold under the enhanced track. ACOs are subject to quality withholds from their shared savings if they do not meet quality benchmarks.</p> <p>Requirements: Participating ACOs must report quality data to CMS after the close of every performance year to be eligible to share in any earned shared savings and to avoid sharing losses at the maximum level.</p> <p>Attribution:</p>	<p>Measures used for implementation: All quality, spending, and patient experience measures</p> <p>Measures used for monitoring: CMS reports ACO-level setting-specific spending and utilization, including but not limited to categories of inpatient (e.g., short-term acute care hospital) and outpatient (e.g., outpatient facility) expenditures, as well as types of inpatient (e.g., hospital discharges) and</p>	<p>In recent years, CMS has proposed changes to the program to promote equity, especially in rural and underserved areas. Access to ACOs appeared inequitable, based on data indicating that Black (or African American), Hispanic, Asian/Pacific Islander, and American Indian/Alaska</p>

Program Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
		<p>depression and follow-up plan; controlling high blood pressure. MSSP ACOs are also given a quality score based on their performance on three quality measures related to care coordination/patient safety, preventive health, and control of diabetes, depression, and hypertension.</p> <p>Spending measure(s): Medicare Parts A and B FFS claims</p> <p>Patient experience measure(s): CAHPS for MIPS Clinician/Group Survey</p>	<p>Voluntary: Beneficiaries confirm care relationships with a primary clinician who is an ACO professional participating in the ACO.</p> <p>Prospective and retrospective claims-based: Based on receiving the plurality of primary care services from primary care physicians, nurse practitioners, physician assistants, clinical nurse specialists, or specialist physicians in the participating ACOs.</p> <p>Volume: Not specified; as of 2020, average number of beneficiaries in an MSSP ACO is 20,700, with the program including approximately 10.6 million attributed beneficiaries, or around 28% of Medicare fee-for-service beneficiaries.</p> <p>Risk stratification or adjustment: When establishing the historical benchmark, CMS uses the hierarchical condition category (HCC) scores to adjust for changes in severity of the population assigned to the ACO between the first and third benchmark years and between the second and third benchmark years. CMS risk-adjusts the county-level expenditures used in calculating the regional component of the national-regional blend growth rate used to trend the first and second benchmark years to the third benchmark year. Under the two-sided risk model, shared losses are adjusted for the percentage of beneficiaries in counties affected by an extreme and uncontrollable circumstance. Individual measures are risk-adjusted, as applicable.^{xlviii}</p> <p>Benchmarking: Yes, based on spending for beneficiaries who would have been assigned to the ACO in the baseline years and the region. When establishing the historical benchmark, CMS uses the</p>	<p>outpatient (e.g., primary care services) utilization.</p> <p>How achievement is measured: ACOs earn quality points on a sliding scale based on level of performance for each measure. The higher the level of performance, the higher the number of quality points earned. The total points earned are summed and divided by the total points available to determine the ACO's quality score. The percentage of shared savings varies based on the ACO's quality score. Additionally, ACOs are assessed based on a combination of both reporting and quality performance requirements.</p> <p>How improvement is measured: Beginning in Performance Year 2015, CMS introduced a Quality Improvement Reward that allows ACOs to earn up to four additional points in each quality domain if they show statistically significant improvement in their performance on quality</p>	<p>Native beneficiaries were less likely to be assigned to a Shared Savings Program ACO than their Non-Hispanic White counterparts. CMS proposed assigning more people who receive care from nurse practitioners, physician assistants, and clinical nurse specialists to ACOs in order to increase the number of people receiving high-quality, accountable care. CMS also proposed changes to the benchmark methodology to encourage ACOs caring for medically complex, high-cost beneficiaries to join the program. CMS previously established advance investment payments for ACOs</p>

^{xlviii} The HWR measure can assign each admission to multiple eligible clinician groups; thus, a two-step approach is used to account for clustering of patients in which five specialty cohort models are used to adjust for case mix differences among providers by risk-adjusting for patients' comorbid conditions. The unplanned hospital admissions for patients with multiple chronic conditions measure uses a hierarchical statistical model that accounts for the clustering of patients within MIPS providers/ACOs and accommodates the varying patient sample sizes of different providers. The model adjusts for 47 demographic and clinical and two social risk factors.

Program Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success	Gaps Related to Current Performance Measures, as applicable
			<p>HCC scores to adjust for changes in severity of the population assigned to the ACO. CMS risk-adjusts the county-level expenditures used in calculating the regional component of the national-regional blend growth rate.</p>	<p>measures from one year to the next.</p>	<p>in rural and underserved communities, made changes to the benchmark methodology, increased the time for ACOs to transition to downside risk, and implemented a health equity adjustment that rewards excellent care delivered to underserved communities.</p>

Exhibit F2. Lessons Learned Related to Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP

Program Name	Clinical Focus, Providers, Setting, Patient Population, and Attribution	Lessons Learned Related to Performance Measurement
<p>Hospital Value-Based Purchasing Program (Hospital VBP)</p> <p>Ongoing</p> <p>Years Active: 2013 – Present</p>	<p>Clinical Focus: Hospital care</p> <p>Providers: Acute care hospitals</p> <p>Setting: Inpatient services</p> <p>Patient Population: Medicare beneficiaries requiring inpatient hospital services</p>	<p>A review of studies that evaluated Hospital VBP between 2013 and 2019 found that Hospital VBP does not lead to meaningful improvements in quality of care or patient outcomes. According to this review, safety-net hospitals’ outcomes may be negatively affected by participation in Hospital VBP.⁴⁸² More recent studies also did not observe associations between Hospital VBP program participation and improvements in health care quality or patient outcomes.^{483,484} An analysis of data from California hospitals participating in the Hospital VBP program suggested that increasing the weight given to outcome measures may increase the quality of care delivered and reduce costs.⁴⁸⁵</p>
<p>Medicare Advantage Star Ratings Program (MA Star Ratings Program)</p> <p>Ongoing</p> <p>Years Active: 2009 – Present</p>	<p>Clinical Focus: Total care</p> <p>Providers: Health plans</p> <p>Setting: Broad</p> <p>Patient Population: Medicare Advantage beneficiaries</p>	<p>A report from McKinsey⁴⁸⁶ found that health plans can improve their Medicare Advantage (MA) Star Ratings scores in two areas:</p> <ul style="list-style-type: none"> - Outcome and process measures: as cut points for measures evolve, health plan contracts will have to continue to make improvements which could require long-term investments. - Patient experience and access scores: health plan contracts will need to increase their focus on patient experience and access measures as scores for these measures have seen less improvement than other domains. <p>Further, improvements in digitization and analytics could help health plans engage beneficiaries and providers and decrease administrative burden.</p> <p>A report from Urban Institute⁴⁸⁷ concluded that the Star Rating system and the quality bonus program (QBP) have many limitations:</p> <ul style="list-style-type: none"> - Score inflation (generous bonuses and no downside risk/penalties for programs with low Star Ratings); - Limitations in underlying data sets, which lead to measures focused on the needs of younger and healthier beneficiaries; - Performance is measured at the contract level and not plan level; and - Many beneficiaries do not use Star Ratings when selecting plans.
<p>Merit-based Incentive Payment System (MIPS)</p> <p>Ongoing</p> <p>Years Active: 2017 – Present</p>	<p>Clinical Focus: Total care</p> <p>Providers: Clinicians, practices</p> <p>Setting: Broad</p> <p>Patient Population: Medicare FFS beneficiaries</p>	<p>Around half of clinicians only partially participated in the program (reporting data for only one or two of the three categories), including one-quarter of clinicians who did not report quality measures. The average performance for each measure was very high, potentially due to selective participation and/or unambitious targets. Approximately 74% of clinicians who partially participated during the first year of MIPS received a positive payment adjustment. There are concerns that the design of MIPS may be too flexible to effectively incentivize clinicians to improve quality.⁴⁸⁸ In addition, participation in MIPS was associated with increased after-hours documentation for physicians, suggesting that physicians may need resources to help reduce the reporting burden.⁴⁸⁹</p>
<p>Medicare Shared Savings Program (MSSP)</p> <p>Ongoing</p> <p>Years Active: 2012- Present</p>	<p>Clinical Focus: Total care</p> <p>Providers: Providers and suppliers (e.g., physicians, hospitals, and others involved in patient care) that create an Accountable Care Organization (ACO)</p> <p>Setting: Broad</p> <p>Patient Population: Medicare FFS beneficiaries</p>	<p>Early performance results from CMS’ Medicare Shared Savings Program Pathways to Success final rule suggest that ACOs with greater financial accountability (e.g., more accurate financial benchmarks, downside risk) are more likely to deliver better coordinated and efficient care for Medicare patients. As of 2022, MSSP had generated overall savings and high-quality performance results for six consecutive years. In the same year, ACOs had a higher average performance on quality measures they are required to report in order to share in savings, including statistically significant higher performance for quality measures related to diabetes and blood pressure control, breast cancer and colorectal cancer screening, tobacco screening and smoking cessation, and depression screening and follow-up, compared with other similarly sized clinician groups not in the program. ACOs that earned more shared savings tended to be low revenue, which may mainly be made up of physicians and may include a small hospital or serve rural areas.</p>

Appendix G. Summary of Model and Performance Measurement Characteristics of Proposals Reviewed by PTAC as of September 2020

Nearly all of the proposals that have been submitted to PTAC included information about proposed performance measures to some degree. The following tables provide specific details on PTAC proposal characteristics (i.e., clinical focus, providers, setting, and patient population); components relevant to performance measurement (i.e., utilization measures, quality measures, spending measures, and patient experience measures); technical issues related to performance measurement (i.e., how payment is adjusted for performance, requirements, volume, risk stratification or adjustment, and benchmarking); and how performance measurement is used to determine success (i.e., measures used for implementation, measures used for monitoring, how achievement is measured, and how improvement is measured) for the 35 proposals that have been submitted to PTAC as of September 2020.

The PTAC proposals are organized into four separate tables:

- Proposals that were found to Meet^{xlix} Criterion 2, Quality and Cost, and Criterion 4, Value over Volume (19 proposals); proposals that were found to Meet Criterion 2 and Not Meet Criterion 4 (1 proposal); proposals that were found to Not Meet Criterion 2 and were found to Meet Criterion 4 (5 proposals); and proposals that were found to Not Meet on both Criterion 2 and Criterion 4 (1 proposal), proposals that were withdrawn (6 proposals), and proposals for which PTAC concluded that the criteria for PFPs established by the Secretary were not applicable (2 proposals). The PTAC proposals in each table are presented in alphabetical order by the proposal submitter's name.

Overview of Methodology Used to Review the Proposals

The following information was reviewed for each submitter's proposal, where available: proposal and related documents, Preliminary Review Team (PRT) Report, and Report to the Secretary (RTS). Information found in these materials was used to summarize the models' main themes related to performance measurement and other administrative and payment characteristics. The categorizations were based on the key information highlighted in these documents and are not exhaustive. Proposals may have elements of their proposed models that fall into additional categories of context, objective, functions, and payment models.

^{xlix} For purposes of this table, proposals that were found to Meet a given criterion includes both proposals that were found to Meet the criterion and proposals that were found to Meet the criterion and Deserve Priority Consideration.

Exhibit G1. Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that Were Found to Meet Criterion 2 and Criterion 4

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>American Academy of Family Physicians (AAFP) <i>(Provider association and specialty society)</i></p> <p>Advanced Primary Care: A Foundational Alternative Payment Model (APC-APM) for Delivering Patient-Centered, Longitudinal, and Coordinated Care</p> <p>Recommended for limited-scale testing, 12/19/2017</p>	<p>Clinical Focus: Primary Care</p> <p>Providers: All physicians with a primary specialty of family medicine, general practice, geriatric medicine, pediatric medicine, or internal medicine</p> <p>Setting: Primary care practices</p> <p>Patient Population: PCPs' patient panels</p>	<p>Utilization measure(s): Inpatient Hospitalization Utilization; Emergency Department Utilization; HEDIS Non-recommended Cervical Cancer Screening in Adolescent Females; Use of Imaging Studies for Low Back Pain</p> <p>Quality measure(s): Diabetes: Hemoglobin A1c (HbA1c) Poor Control (> 9%); Medication Reconciliation Post-Discharge; Colorectal Cancer Screening; Diabetes Eye Exam; Diabetes Medical Attention for Nephropathy; Body Mass Index (BMI) Screening and Follow-Up Plan; Diabetes Foot Exam; Use of Antithrombotic for Ischemic Vascular Disease; Tobacco Use Screening and Intervention; HEDIS Controlling High Blood Pressure; Cervical Cancer Screening; Depression Remission at 12 Months; Persistent Beta-Blocker Treatment After a Heart Attack; Medication Management for People with Asthma; Avoidance of Inappropriate Use of Antibiotic Treatment</p> <p>Spending measure(s): Not specified.</p> <p>Patient experience measure(s): CAHPS</p>	<p>How payment is adjusted for performance: If a provider does not meet performance benchmarks, the provider will have to repay all or part of their incentive payments (depending on level of performance) or may be expelled from the APM and forced to return to traditional FFS.</p> <p>Requirements: To be eligible for performance-based incentives, providers must meet or exceed "agreed upon" benchmarks for performance measures.</p> <p>Attribution: Prospective, hierarchical process based on patient choice, wellness visits, evaluation & management (E&M) visits, and primary care prescription and order events</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Uses the Minnesota Complexity Assessment Method to risk stratify the primary care global payment and the population-based payment on an annual basis.</p> <p>Benchmarking: Benchmarking is based on historical performance and reassessed after two or more years.</p>	<p>Measures used for implementation: APC-APM entities are not expected to track all of the measures in the core measure set. The measures used for implementation are not specified.</p> <p>Measures used for monitoring: Each APC-APM entity will be evaluated based on reporting their choice of six quality, utilization, or patient experience measures. APM entities will have the opportunity to choose a set number of core measures that are meaningful to them. Entities will also be evaluated on their inpatient hospitalization and emergency department utilization.</p> <p>How achievement is measured: Success is measured by assessments of quality and cost-effective care relative to benchmarks. Measurement methodology is not specified.</p> <p>How improvement is measured: Benchmarking is based on historical performance. Measurement methodology is not specified.</p>
<p>American College of Emergency Physicians (ACEP) <i>(Provider association/ specialty society)</i></p> <p>Acute Unscheduled Care Model (AUCM): Enhancing</p>	<p>Clinical Focus: Emergency department (ED) services</p> <p>Providers: ED physicians</p> <p>Setting: ED</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Safe Discharge Assessment (patient engagement); Shared Decision-Making (process of care coordination); Event-Free Post-Discharge Rate (post-discharge outcomes)</p> <p>Spending measure(s): ED costs per episode</p>	<p>How payment is adjusted for performance: If spending for eligible and attributed episodes is less than the bundled payment target price, the participant is eligible for a positive reconciliation payment; if it is more, the participant will have to reimburse CMS. Performance on a set of quality measures determines eligibility for reconciliation payments and the size of discount built into each episode's target price.</p> <p>Requirements: To be eligible for reconciliation payments, providers must be in the acceptable, good, or excellent quality performance category.</p>	<p>Measures used for implementation: All measures listed</p> <p>Measures used for monitoring: All measures listed</p> <p>How achievement is measured: Quality performance is classified as unacceptable, acceptable, good, or excellent, based on ability to meet or surpass the minimum threshold for the</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>Appropriate Admissions</p> <p>Recommended for implementation, 09/06/2018</p>	<p>Patient Population: Patients with qualifying ED visits</p>	<p>Patient experience measure(s): N/A</p>	<p>Attribution: Episodes are attributed to the ED physician based on a qualifying ED visit. All Medicare services (except those identified in BPCI Advanced) that occur in 30-days post ED visit are included.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Uses CMS hierarchical condition category (CMS-HCC) methodology to adjust target prices annually.</p> <p>Benchmarking: Benchmarking is based on participants' historical performance, risk-adjusted for factors that impact the admission decision.</p>	<p>three quality measures. To get a quality performance classification of good or excellent, there are additional requirements regarding achieving a certain percentage of clean cases.</p> <p>How improvement is measured: Quality benchmarks are based on historical performance. Facility-based target prices are based on three years of historical claims.</p>
<p>American College of Physicians-National Committee for Quality Assurance (ACP-NCQA)</p> <p><i>(Provider association and specialty society/other)</i></p> <p>The "Medical Neighborhood" Advanced Alternative Payment Model (AAPM) (Revised Version)</p> <p>Recommended for testing to inform payment model development, 09/15/2020</p>	<p>Clinical Focus: Coordination between specialists and PCPs</p> <p>Providers: Primary Care Practices in Comprehensive Primary Care Plus (CPC+) and Primary Care First (PCF), specialty practices meeting clinical transformation and care coordination criteria for Medicare Access and Children's Health Insurance Program Authorization Act of 2015 (MACRA)-recognized Patient Centered Specialty Practices (PCSPs)</p>	<p>Utilization measure(s): Two readmission measures based on administrative claims data: 1) Hospital Wide All-cause Readmission Rate Measure; 2) SNF 30-day All-cause Readmission Measure</p> <p>Quality measure(s): Core: Revised Care Plan Measure and Revised All-cause unplanned admissions for patients with multiple chronic conditions measure. Three relevant measures from following list: Atrial Fibrillation and Atrial Flutter: Chronic Anticoagulation Therapy; Cardiac Rehabilitation: Patient Referral from Outpatient Setting; Chronic Stable Coronary Artery Disease: Antiplatelet Therapy; CAD: Beta-Blocker Therapy-Prior MI or LVEF<40%; Statin Therapy for the Prevention and Treatment of CVD; Persistence of Beta-Blocker Treatment after Heart Attack; IVD: Use of Aspiring or Another Antiplatelet; HF: ACE or ARB; Therapy for LVSD; HF: LVSD; Use of Imaging for Low Back Pain; Documentation of Signed Opioid Treatment Agreement; Stroke and Stroke Rehab: Discharged on Antithrombotic Therapy; Overuse of Neuroimaging for</p>	<p>How payment is adjusted for performance: Performance-Based Payment Adjustments (PBPA). Retrospective positive or negative payment adjustment made based on actual spending compared with a financial benchmark.</p> <p>Requirements: Meet minimum standards for all quality and utilization performance measures. PBPA retained is continuously adjusted based on how well practice performs (up to 100%).</p> <p>Attribution: Monthly Care Coordination Fees (CCFs) triggered when at least one office visit billed by PCSP and a specialist is designated in Care Coordination Agreement as managing or co-managing a patient's condition, based on services billed under practice's TIN. Patients must be appropriately referred by CPC+ participating primary care clinicians and have an office visit billed through the participating Medical Neighborhood Model (MNM) specialist; attribution is conducted on quarterly basis.</p> <p>Volume: At least 100 patients must be attributed and trigger monthly CCFs over course of year.</p> <p>Risk stratification or adjustment: Based on hierarchical condition categories (HCC) risk scoring methodology. CCF is varied based on risk tier.</p>	<p>Measures used for implementation: Based on financial performance relative to a benchmark based half on practice's historical spending (two calendar years before performance year) and half on regional spending during relevant performance year</p> <p>Measures used for monitoring: An independent third-party evaluator compares patients referred to an MNM Model to control group and assesses whether and how patient participation affects patient experience, health outcomes, resource utilization, and total cost of care.</p> <p>How achievement is measured: Comparison of total spending and rates of hospital admissions and ED visits following referral</p> <p>How improvement is measured: Benchmarking is based on historical performance. Measurement methodology is not specified.</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
	<p>Setting: Primary care and specialty practices</p> <p>Patient Population: Patients with multiple chronic conditions</p>	<p>Patients with Primary Headache and a Normal Neurological Evaluation; Measures Included in NCQA List; Adult MDD: Suicide Risk Assessment; Evaluation or Interview for Risk of Opioid Misuse; Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis; Annual HCV Screening for Patients who are Active Injections Users; One-Time Screening for HCV for Patients at Risk; HIV: Viral Load Suppression, HIV/AIDS: CD4 Cell Count or Percentage Performed; HIV/AIDS Pneumocystis Jiroveci Pneumonia Prophylaxis; Annual Cervical Screening or Follow-up in High-Risk Women; National Healthcare Safety Network Facility-Wide Inpatient Hospital-Onset MRSA Bacteremia Outcome; National Healthcare Safety Network CLBSI Outcome Measure; Adult Sinusitis: CT for Acute Sinusitis</p> <p>Spending measure(s): Medicare spending</p> <p>Patient experience measure(s): CAHPS questions 6, 8, 10-15, 17-18, 20. PCMH3: Provider always informed and up to date about care patient received; PCMH4: Someone in provider's office discussed specific health goals; PCMH5: someone in provider's office asked about things making it hard to take care of health.</p>	<p>Benchmarking: Recalibrated annually; based on practice's historical spending and trended forward based on regional growth rates</p>	

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>The American College of Surgeons (ACS)</p> <p><i>(Provider association/specialty society)</i></p> <p>The ACS–Brandeis Advanced Alternative Payment Model</p> <p>Recommended for limited-scale testing, 4/11/2017</p>	<p>Clinical Focus: Cross-clinical focus</p> <p>Providers: Single/multispecialty practices; groups of small provider practices</p> <p>Setting: Inpatient, outpatient, ambulatory</p> <p>Patient Population: Broad (includes 100+ conditions or procedures)</p>	<p>Utilization measure(s): Unplanned hospital readmission within 30 days of principal procedure</p> <p>Quality measure(s): Surgical Plan and Goals of Care; Identification of Major Comorbid Medical Conditions; Preventive Care and Screening: Tobacco Screening and Cessation Intervention; Preoperative Key Medications Review for Anticoagulation Medication; Patient-Centered Surgical Risk Assessment and Communication; Patient Frailty Evaluation; Perioperative Composite; Intraoperative Timeout Safety Checklist; Intraoperative Surgical Debriefing; Postoperative Care Plan; Postoperative Review of Patient Goals of Care; Postoperative Care Coordination and Follow-up; Postoperative Plan Communication with Patient and Family; Post-Discharge Review of Patient Goals of Care; Resumption Protocol</p> <p>Spending measure(s): Quarterly Expenditures</p> <p>Patient experience measure(s): Patient experience with surgical care based on CAHPS Surgical Care Survey (S-CAHPS)</p>	<p>How payment is adjusted for performance: Payment is adjusted based on quality measures, incorporating two-sided risk. The model incorporates shared savings and losses.</p> <p>Requirements: Entities are assigned cost targets (based on CMS episode grouper); entities share in savings if costs are under the target and are required to repay losses if their costs are over the expected target.</p> <p>Attribution: Algorithms will be used to identify all clinicians who contribute to the care for each patient for each type of episode. Only qualifying providers who are in the APM entity and involved in the patient’s care are included in the payment model.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: The episode grouper risk adjusts each patient based on the patient’s historical claims data related to each episode, as well as selected indicators, such as dual-eligible status, disability status, and rural/urban location.</p> <p>Benchmarking: Authors propose to use an updated CMS episode grouper to generate patient-specific, risk-adjusted cost targets (episode benchmarks) using Medicare Parts A and B claims data. Target prices are compared with actual cost.</p>	<p>Measures used for implementation: All measures listed</p> <p>Measures used for monitoring: All measures listed</p> <p>How achievement is measured: Achievements are measured through the entity’s composite performance, placing them in a quality tier that determines the discount factor for targets, shared savings, or losses</p> <p>How improvement is measured: N/A</p>
<p>Avera Health (Avera Health)</p> <p><i>(Regional/local multispecialty practice or health system)</i></p> <p>Intensive Care Management in Skilled Nursing Facility Alternative</p>	<p>Clinical Focus: Primary care (geriatricians) in skilled nursing facilities (SNFs)</p> <p>Providers: Geriatrician Care Teams (GCTs)</p> <p>Setting: SNFs and NFs</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Percentage of short-stay residents who have had an outpatient emergency department visit; SNF 30-day All-Cause Readmission Measure; percentage of short-stay residents given seasonal influenza vaccine; percentage of short-stay residents given pneumococcal vaccine; percentage of short-stay residents new administered antipsychotic medication; percentage of long stay residents with</p>	<p>How payment is adjusted for performance: Two options: 1) Performance-Based Payment- Based on performance on scored quality metrics, payment adjustments will be made at 100% (receive full payment), 50% (receive half payment), and 0% (receive no payment). 2) Shared Savings Model- Annual financial reconciliation to determine if savings were generated. If necessary, additional shared savings will be given to model participant, but in later years, repayment may be due to CMS when savings are not achieved.</p> <p>Requirements: High Quality Performance (meet performance criteria on 8 or more of 11 metrics); Average Quality Performance (meet performance criteria on 4-8 of 11 metrics);</p>	<p>Measures used for implementation: Scored quality metrics (percentage of short-stay residents who have had an outpatient emergency department visit); SNF 30-day All-Cause Readmission Measure; percentage of short-stay residents assessed and given, appropriately, the seasonal influenza vaccine; percentage of short-stay residents assessed and given, appropriately, the pneumococcal vaccine; percentage of short-stay residents who</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>Payment Model (ICM SNF APM)</p> <p>Recommended for implementation, 3/27/2018</p>	<p>Patient Population: SNF Residents</p>	<p>UTI; percentage of long stay residents administered antipsychotic medications; percentage of long-stay residents with depressive symptoms; percentage of long-stay residents who received antianxiety or hypnotic medication; percentage of long-stay residents given influenza vaccines; percentage of long-stay residents given pneumococcal vaccine</p> <p>Spending measure(s): PBPM payments</p> <p>Patient experience measure(s): Beneficiary satisfaction, but there are no metrics related to outcomes due to there being no standardized CMS measures.</p>	<p>Low Quality Performance (meet performance criteria on 4 or less of 11 metrics).</p> <p>Attribution: For the Performance-Based Payment option, all Medicare FFS beneficiaries that offer model services are used to calculate regular payments. For Shared Savings Model, all Medicare beneficiaries in facilities which offer model services but are not attributed to another shared savings program are attributed. Based on trigger event being the beneficiary's admission to a participating SNF/NF; beneficiaries are aligned to the facility throughout their stay, and the alignment period ends 30 days following facility discharge.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: The Performance-Based Payments option does not require risk adjustment. The Shared Savings Model will use CMS' prospective hierarchical condition category risk score to adjust the Target Bundle Price to reflect underlying risk.</p> <p>Benchmarking: Yes, with measure-specific performance criteria for achievement and improvement; use of the Scored Quality Metrics (SQM) will allow programs to benchmark themselves against the LTC population as a whole.</p>	<p>are newly administered antipsychotic medication; percentage of long stay residents with a urinary tract infection; percentage of long stay residents who are administered antipsychotic medications; percentage of long stay residents who have depressive symptoms; percentage of long stay residents who received an antianxiety or hypnotic medication; percentage of long stay residents assessed and given, appropriately, the seasonal influenza vaccine; percentage of long stay residents assessed and given, appropriately, the pneumococcal vaccine</p> <p>Measures used for monitoring: Percentage of short-stay residents who made improvements in function; percentage of short-stay residents who were successfully discharged to the community; percentage of short-stay residents who self-report moderate to severe pain; percentage of short-stay residents who have pressure ulcers that are new or worsened; percentage of long stay residents whose ability to move independently worsened; percentage of long stay residents whose need for help with daily activities has increased; percentage of long stay residents with pressure ulcers; percentage of long stay residents who have/had a catheter inserted and left in their bladder; percentage of long stay residents who were physically restrained; percentage of long stay residents who self-report moderate to severe pain; percentage of long stay residents experiencing one or more falls with major injury; percentage of long stay residents who lose control of their bowels or bladder; percentage of</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
				<p>long stay residents who lose too much weight.</p> <p>How achievement is measured: Whether or not practices/facilities meet performance criteria of 11 metrics; participants' performance will be evaluated on their ability to achieve above average performance.</p> <p>How improvement is measured: Whether or not practices/facilities meet performance criteria of 11 metrics. Data are taken from CMS quarterly reports, as well as Electronic Health Records. Participants' performance will be evaluated on their ability to demonstrate measurable improvement toward the goal for key metrics directly impacted by the care model.</p>
<p>Coalition to Transform Advanced Care (C-TAC) (Coalition) Advanced Care Model (ACM) Service Delivery and Advanced Alternative Payment Model Recommended for limited-scale testing, 3/26/2018</p>	<p>Clinical Focus: Serious illness and palliative care</p> <p>Providers: ACM care team; other ancillary collaborator organizations</p> <p>Setting: Patient home</p> <p>Patient Population: Patients with serious illness</p>	<p>Utilization measure(s): Not specified</p> <p>Quality measure(s): ACM Team Visit within 48 hours of hospital discharge; timeliness of advance care planning, medication reconciliation post-discharge, proportion of patients who died and who were admitted to the ICU in the last 30 days of life, proportion of patients who died who were admitted to hospice for three days or more; ACM provider attestation that patient's care plan is consistent with preferences</p> <p>Spending measure(s): Total cost of care for the last 12 months of life</p> <p>Patient experience measure(s): Timeliness of care; getting help for</p>	<p>How payment is adjusted for performance: The proposed model provides an upside bonus for quality funded by shared savings and downside risk (beginning in year 3).</p> <p>Requirements: Not specified</p> <p>Attribution: Based on the participating entity's full Medicare population or only those that are ACM-eligible (those with advanced illness¹)</p> <p>Volume: The entity must have a defined network of participating physicians and other eligible professionals with a reasonable projected advanced illness patient volume to operate the ACM services.</p> <p>Risk stratification or adjustment: Episode-based regression analysis modeling; percentile scoring adjusted at the regional level.</p>	<p>Measures used for implementation: Total cost of care for the last 12 months of life</p> <p>Measures used for monitoring: Level of ACM services provided to ACM beneficiaries; proportion of enrolled beneficiaries over projected volume based on historical trend; differences in patient characteristics between enrolled and non-enrolled advanced illness; hospice enrollment and length of stay for ACM beneficiaries; characteristics of hospice vs. non-hospice ACM beneficiaries; proportion of ACM enrollees with more than 12 months of enrollment; differences in patient characteristics between ACM enrollees</p>

¹ Identification of advanced illness is based on *International Classification of Diseases, 10th Revision (ICD-10)* primary diagnosis codes in the diagnosis category that appeared on the majority of a patient's claims in their last 12 months of life.

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
		<p>symptoms (pain, anxiety and sadness, trouble breathing); effective communication composite; care coordination; patient overall satisfaction; patient engagement composite; shared decision-making; caregiver support composite; quality of care transitions from ACM to hospice composite</p>	<p>Benchmarking: Yes, quality performance based on historical trends; financial performance based on risk-adjustedⁱⁱ historical trends, adjusted at the regional level and weighted toward more recent episodes</p>	<p>with more than 12 months of enrollment vs. ACM enrollees with 12 months or less of enrollment; all-cause unplanned admissions for ACM beneficiaries; ambulatory sensitive condition acute composite for ACM beneficiaries</p> <p>How achievement is measured: Meeting minimum quality standard; achieving savings of at least 4% of the spending target</p> <p>How improvement is measured: Performance targets will be set over time (by year 3) to promote ongoing improvement or maintenance of high performance. Low-quality performers have a remediation period and will be required to leave the program if corrective actions do not show positive trends within six months and significant improvement within a year.</p>
<p>Hackensack Meridian Health and Cota, Inc. (HMH/Cota)</p> <p><i>(Regional/ local multispecialty practice or health system; Device/ technology company)</i></p> <p>Oncology Bundled Payment Program Using CNA-Guided Care</p>	<p>Clinical Focus: Oncology</p> <p>Providers: Eligible professionals in HMH health system with attributed Medicare cancer patients</p> <p>Setting: Inpatient and outpatient care</p> <p>Patient Population: Cancer (breast, colon, rectal, and lung)</p>	<p>Utilization measure(s): Visits</p> <p>Quality measure(s): Surgery, oncology, and genetics measures for breast cancer; surgery and oncology measures for colorectal cancer; surgery and oncology for lung cancer; oncology, infection monitoring, COTA analytics, risk management for all disease groups</p> <p>Spending measure(s): Total cost of care</p> <p>Patient experience measure(s): Patient-reported outcomes from Press Ganey, College of Surgeons, OCM, GPRO, and national guidelines concerning pain management and guidelines</p>	<p>How payment is adjusted for performance: The costs of each certified nursing assistant (can) will be aggregated up to the bundle level using a weighted average approach. These average costs would be used to compute a prospective 12-month price for each of the 27 bundles that cover all the CNAs in the four cancer types. HMH will receive these prospective payments and use them to compensate providers and pay for care coordination and other uncovered services. The submitter states that physicians will not assume downside risk. Physicians will receive higher compensation through the bundle if performance metrics are achieved.</p> <p>Requirements: Not specified; the submitter noted the importance of investment in analytics, seamless physician communication.</p> <p>Attribution: N/A</p>	<p>Measures used for implementation: Visit occurrence, quality measures, and related outcomes</p> <p>Measures used for monitoring: In addition to measures used for implementation, total cost of care, finance, and reliability are monitored. Financial monitoring including monthly/quarterly review of all HMH bundle costs, cost analysis for each bundle patient, review of lane assignment for each bundle patient, and review of payer excluded services, and finance prepared monthly/quarterly MD report care; reliability measures for radiation, surgery, pharmacy</p>

ⁱⁱ Risk adjustment factors include clinical risk, prior utilization, and Medicare-Medicaid dual-eligibility.

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>Recommended for limited-scale testing, 9/8/2017</p>			<p>Volume: On average, each oncologist has approximately 36 patients in the program.</p> <p>Risk stratification or adjustment: The CNA's precision is leveraged to adjust the bundle price for relative patient risk.</p> <p>Benchmarking: Yes, based on data-driven classification system for cancer patient risk and treatment pathways</p>	<p>chemotherapy, nursing chemotherapy, and genetic counseling referrals; patient satisfaction from Press Ganey report</p> <p>How achievement is measured: Services provided, meeting agreed upon thresholds for clinical quality and patient satisfaction, total cost of care</p> <p>How improvement is measured: Improvement in relevant clinical cancer outcomes; measurement methodology not specified</p>
<p>Johns Hopkins School of Nursing and the Stanford Clinical Excellence Research Center (Hopkins/Stanford) <i>(Academic institution)</i> CAPABLE Provider Focused Model</p> <p>Recommended for testing as specified in PTAC comments, 9/6/19</p>	<p>Clinical Focus: Home health, functional care for elders</p> <p>Providers: Registered nurses, occupational therapists</p> <p>Setting: Home</p> <p>Patient population: Patients living at home and reporting difficulty in at least one activity of daily living or at least two instrumental activities of daily living, income</p>	<p>Utilization measure(s): Patient-centered visits</p> <p>Quality measure(s): Number of activities of daily living (ADLs) and instrumental ADLs (IADLs) considered "difficult" before and after intervention, depressive symptoms (eight or nine item Patient Health Questionnaire [PHQ-8, PHQ-9]) before and after intervention, home hazard or fall risk</p> <p>Spending measure(s): Not specified</p> <p>Patient experience measure(s): Patient satisfaction via post-intervention qualitative interview of HCAHPS survey</p>	<p>How payment is adjusted for performance: The submitter recommended implementing a partial bundled payment with partial upside risk and moving toward a fully capitated model. Pre-defined quality/cost metrics and achievement/improvement thresholds were not specified.</p> <p>Requirements: N/A</p> <p>Attribution: N/A; intended patients include Medicare beneficiaries with at least two chronic conditions and difficulty with at least one ADL. Ideally, any patient identified as high-risk could be enrolled by a health plan or have a provider "prescribe" CAPABLE services.</p> <p>Volume: Not specified</p> <p>Risk stratification or adjustment: Not specified</p> <p>Benchmarking: Not specified</p>	<p>Measures used for implementation: Activities of daily living (ADLs) and instrumental ADLs (IADLs), depression, home hazard or fall risk</p> <p>Measures used for monitoring: ADL, IADL, and PHQ-8 scores and fall risk assessment</p> <p>How achievement is measured: Avoidable utilization, function, and depression scores</p> <p>How improvement is measured: Not specified</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>Illinois Gastroenterology Group and SonarMD, LLC (IGG/SonarMD)</p> <p><i>(Regional/local single specialty practice; Device/technology company)</i></p> <p>Project Sonar</p> <p>Recommended for limited-scale testing, 4/10/2017</p>	<p>Clinical Focus: Crohn's disease</p> <p>Providers: Gastroenterology practices; community-based physicians; nurse care manager; community-based specialists</p> <p>Setting: Patient home</p> <p>Patient Population: Patients with Crohn's disease</p>	<p>Utilization measure(s): ED visits, inpatient admissions</p> <p>Quality measure(s): Risk assessment tool including 26 biopsychosocial risk metrics in three categories (inflammation risk, disease burden risk, and comorbidity risk); MIPS measures; potentially avoidable complications</p> <p>Spending measure(s): Crohn's related cost of care</p> <p>Patient experience measure(s): SonarMD patient survey</p>	<p>How payment is adjusted for performance: Payment adjustments are based on quality and financial performance. In the case of cost savings, the shared savings component of the payment would be paid to the physician entity. In a situation where the physician's attributed costs overrun the expected target, the physician would be required to repay losses up to the agreed upon limit in its contract with CMS. To protect against catastrophic losses, the model will build in stop-loss provisions and outlier protections.</p> <p>Requirements: Not specified</p> <p>Attribution: Patients with Crohn's Disease; methodology not specified.</p> <p>Volume: N/A; Project Sonar was initially deployed in 2012 by the Illinois Gastroenterology Group (IGG), a 50-physician practice.</p> <p>Risk stratification or adjustment: The PMPM payment is additional to the fee-for-service payment and varies on an annual basis, adjusted based upon mutually agreed upon goals for the clinical and financial performance of the patients enrolled.</p> <p>Benchmarking: Not specified</p>	<p>Measures used for implementation: Risk assessment tool</p> <p>Measures used for monitoring: Clinical and financial performance</p> <p>How achievement is measured: Clinical and cost outcomes; methodology not specified</p> <p>How improvement is measured: Patient quality of life, costs; methodology not specified.</p>
<p>Innovative Oncology Business Solutions, Inc. (IOBS)</p> <p><i>(For-profit corporation)</i></p> <p>Making Accountable Sustainable Oncology Networks (MASON)</p> <p>Referred for further development and Implementation, 12/10/2018</p>	<p>Clinical Focus: Cancer</p> <p>Providers: National Cancer Care Alliance (NCCA) oncology physicians</p> <p>Setting: Patient home</p> <p>Patient Population: NCCA patients with cancer</p>	<p>Utilization measure(s): Hospitalization rates, diversion from ED to office, ED visits</p> <p>Quality measure(s): Compliance with evidence-based pathways, outcomes of chemotherapy regimens, infrastructure certification (e.g., American College of Radiology certification of imaging and radiation therapy)</p> <p>Spending measure(s): Total cost of care</p> <p>Patient experience measure(s): Patient satisfaction via survey</p>	<p>How payment is adjusted for performance: Two percent of the oncology payment category (OPC), which includes all expenses related to cancer care except drugs, is reserved for a quality pool. If quality measures are not met, that money is returned to CMS. At the end of an episode of care, the actual costs are compared with the OPC. If the practice spends less caring for the patient, and all the quality parameters are met, the practice shares in the savings.</p> <p>Requirements: N/A</p> <p>Attribution: N/A</p> <p>Volume: N/A; NCCA currently manages approximately 250,000 cancer patients. Practices will add patients at an average of 300 new patients per oncologist per year. The submitter noted that</p>	<p>Measures used for implementation: Hospitalization rates, diversion from ED to office, ED visits, compliance with evidence-based pathways, outcomes of chemotherapy regimens, infrastructure certifications, total cost of care, patient satisfaction</p> <p>Measures used for monitoring: HCCs, socioeconomic status</p> <p>How achievement is measured: Meeting quality metrics</p> <p>How improvement is measured: Hospitalization rates, length of hospital stay, patient satisfaction, hospital-</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>a major barrier is the need to spend at least a year with large volumes of patients to develop accurate OPCs.</p> <p>Risk stratification or adjustment: MASON target prices in the OPCs are adjusted for comorbidities and the clinical situation of each individual cancer patient.</p> <p>Benchmarking: Yes, based on distribution of expenditures, as opposed to a point estimate</p>	<p>acquired infections, compare cost savings achieved by enrollment in MASON program to traditional Oncology Care Model</p>
<p>Large Urology Group Practice Association (LUGPA) <i>(Provider association and specialty society)</i></p> <p>LUGPA Advanced Payment Model for Initial Therapy of Newly Diagnosed Patients with Organ Confined Prostate Cancer</p> <p>Not recommended, 2/28/18</p>	<p>Clinical Focus: Urology/Oncology (treatment of prostate cancer)</p> <p>Providers: Eligible professionals (including urologists) at large and small urology and multispecialty practices</p> <p>Setting: Large and small urology and multispecialty practice</p> <p>Patient population: Newly diagnosed prostate cancer patients with localized disease</p>	<p>Utilization measure(s): Avoidance of overuse of bone scan for staging low-risk prostate cancer</p> <p>Quality measure(s): Time on active surveillance, biopsy follow-up</p> <p>Spending measure(s): All Medicare Parts A and B payments in initial episodes</p> <p>Patient experience measure(s): Prostate cancer shared decision-making process</p>	<p>How payment is adjusted for performance: Practice performance on all quality measures would be tied to the performance-based payment calculation. The performance-based payment would retrospectively compare actual initial episode spending against a target amount. Beneficiaries who are diagnosed with localized prostate cancer after biopsy would begin 12-month initial total cost of care episodes, including all Parts A and B services starting with the prostate biopsy. The episode would be classified into one of 12 proposed subcategories. Participants earn performance-based payments or owe performance-based repayments based on the number of quality performance targets achieved/exceeded.</p> <p>Requirements: Not specified</p> <p>Attribution: Initial 12-month episodes of care, beginning with prostate biopsy and a diagnosis of prostate cancer, for both beneficiaries receiving active surveillance and those receiving active intervention; subsequent 12-month episodes of care for beneficiaries who remain on active surveillance at the end of an initial 12-month active surveillance episode. All episodes would be attributed prospectively to the TIN that bills the professional claim for the prostate biopsy.</p> <p>Volume: The submitter estimated total episode volume of 62,640 episodes. Because certain episode subcategories have low volume, the submitter proposed pooling episode spending across episode subcategories when blending practice and regional historical episode expenditures. Additionally, the submitter noted that while small or rural practices with low episode volume might expect more variation in episode expenditures, such practices can mitigate their financial risk by</p>	<p>Measures used for implementation: Avoidance of overuse of bone scan for staging low-risk prostate cancer, biopsy follow-up, time on active surveillance, prostate cancer shared decision-making process</p> <p>Measures used for monitoring: Histopathological grade and stage, PSA results, molecular/genetic biomarkers if applicable, and an attestation regarding beneficiary health status for each initial episode</p> <p>How achievement is measured: Participants must achieve or exceed quality performance targets.</p> <p>How improvement is measured: Relative to historical baseline or previously submitted data</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>applying for an alternative risk track with lower stop-loss/stop-gain limits.</p> <p>Risk stratification or adjustment: Risk adjustment for initial episodes will incorporate the CMS-HCC scores of beneficiaries in initial active surveillance episodes, as well as the type of active intervention for beneficiaries receiving active intervention. The submitter noted that time on active surveillance could be risk-adjusted by weighting the distribution of beneficiaries across the low-, medium-, and high-risk active surveillance categories.</p> <p>Benchmarking: Although each category will have a component benchmark price for the performance year, each APM entity would ultimately receive a single composite benchmark price calculated based on 1) practice-specific and regional historical utilization of active surveillance; and 2) practice-specific performance year composition of episodes in subcategories within active surveillance and active intervention episode categories. Benchmarking approach varies for each measure.</p> <p>Avoidance of overuse of bone scan for staging low-risk prostate cancer – 85% target (i.e., at least 85% do not receive bone scan)</p> <p>Biopsy follow-up – 80% target</p> <p>Time on active surveillance – improvement relative to historical baseline</p> <p>Prostate cancer shared decision-making process – improvement relative to previously submitted data</p>	
<p>Icahn School of Medicine at Mount Sinai (Mount Sinai)</p> <p><i>(Academic institution)</i></p> <p>"HaH-Plus" (Hospital at Home-Plus): Provider-Focused Payment Model</p>	<p>Clinical Focus: Inpatient services in home setting</p> <p>Providers: Physicians; HaH Plus providers</p> <p>Setting: Patient home</p>	<p>Utilization measure(s): Post-acute emergency department (ED) visits</p> <p>Quality measure(s): Propose to use existing measures adapted for the home setting: Measures of Care Plan, Documentation of Current Medications in the Medical Record, Medication Reconciliation Post-Discharge, rate of combined adverse events (falls and pressure sores); and the Activity Measure</p>	<p>How payment is adjusted for performance: If APM entity spending is less than benchmark, entity could earn performance-based payment of up to 100% of difference between benchmark and cost up to cap of 10% of benchmark. If APM entity spending is more than benchmark, entity could be liable for up to 100% of losses up to cap of 10% of benchmark</p>	<p>Measures used for implementation: Measures of Care Plan, Documentation of Current Medications in the Medical Record, Medication Reconciliation Post-Discharge, five HCAHPS measures, rate of combined adverse events (falls and pressure sores), AM-PAC Inpatient Basic Mobility Short-Form and Inpatient Daily Activity Short Form</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>Recommended for implementation, 9/17/2017</p>	<p>Patient Population: Eligible patients in one of 44 diagnosis-related groups (DRGs) for acute conditions</p>	<p>for Post-Acute Care (AM-PAC) Inpatient Basic Mobility Short Form and Inpatient Daily Activity Short Form. Plus, the following: Adverse Events (e.g., hospital-acquired infections, complications); Outcomes (e.g., mortality, readmissions); Process Measures of Quality; Clinical Process; Program Process.</p> <p>Spending measure(s): Total spending in HaH-Plus episodes.</p> <p>Patient experience measure(s): Patient-reported outcomes and experiences with care. Use of adapted version of the HCAHPS. Five measures are most applicable: Communication with doctors; communication with nurses; communication about medications; care transitions; and overall rating of care experience.</p>	<p>Requirements: APM entity's spending should be less than the benchmark. Pro-ration available based on number of quality metrics attained, shared savings, or repayment.</p> <p>Attribution: Patients are offered the option of HaH-Plus if they are clinically appropriate and meet home and patient safety criteria. Claims with qualifying diagnosis-related groups (DRGs) are aligned to the furnishing provider.</p> <p>Volume: Not specified; the submitter estimated that 1.7 million discharges could be clinically appropriate nationwide across all payers. The submitter noted that requirements for participation may be beyond the scope of typical solo small independent practices, but that it may be possible to offer modifications to the payment methodology.</p> <p>Risk stratification or adjustment: A comparison group of patients admitted in the same calendar year quarter to non-participating hospitals in the same region will be used to find a spending target for the amount Medicare would have spent without the HaH-Plus program.</p> <p>Benchmarking: Yes, separate achievement thresholds for each of 10 quality metrics linked to payment. HaH-Plus total Medicare spending per beneficiary for the HaH-Plus episode should be compared with that for a weighted sample of FFS discharges from the region with matching DRGs.</p>	<p>Measures used for monitoring: Process measures (Care Plan, Documentation of Current Medication, Medication Reconciliation), length of stay, utilization, escalation rates, and readmissions will be captured in the EHR. Adverse events can be audited against the medical record and through reviewing claims associated with these events. Patient-reported outcomes and experiences with care can be collected by an outside party. Additionally, the submitter noted that CMS could track the ratio of HaH-Plus episodes, HaH-Plus observation episodes, inpatient admissions, and observation stays to total ED visits, compared with the historical ration of inpatient admissions and observation stays to total ED visits.</p> <p>How achievement is measured: The APM entity must meet both financial benchmarks and quality metrics to earn shared savings.</p> <p>How improvement is measured: Financial benchmarks and quality metrics</p>
<p>New York City Department of Health and Mental Hygiene (NYC DOHMH)</p> <p><i>(Public health department)</i></p> <p>Multi-provider, bundled episode of care payment model for treatment of chronic hepatitis C virus (HCV) using care coordination by</p>	<p>Clinical Focus: Hepatitis C virus (HCV)</p> <p>Providers: Primary care physicians (trained by hepatologists/gastroenterologists); specialists; nurse practitioners; physician assistants; and non-clinician staff</p>	<p>Utilization measure(s): ED visit rate</p> <p>Quality measure(s): Risk-adjusted facility-based HCV Sustained Virological Response (SVR) rate</p> <p>Spending measure(s): Part B payments</p> <p>Patient experience measure(s): N/A</p>	<p>How payment is adjusted for performance: The payment model includes the opportunity for shared savings. Providers that meet the HCV SVR benchmark are eligible to receive shared savings and an annual bonus. Providers achieving a score below the benchmark must pay back losses, corresponding to a proportion of the per patient expected savings missed summed across all patients with undocumented SVR.</p> <p>Requirements: All employed physicians treating patients with HCV at a participating facility would be required to participate in the model. To earn shared savings, providers must be designated as "high-performers" (i.e., those that meet or exceed the HCV SVR benchmark).</p>	<p>Measures used for implementation: Risk-adjusted facility-based HCV SVR rate</p> <p>Measures used for monitoring: Risk-adjusted facility-based HCV SVR rate, ED rate, Part B payments, total cost of care</p> <p>How achievement is measured: Achievement is measured by assessment of quality (HCV SVR rate) relative to benchmarks.</p> <p>How improvement is measured: The submitter noted that a facility-based SVR rate catalogue could be established and</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>employed physicians in hospital outpatient clinics</p> <p>Not recommended, 12/18/2018</p>	<p>Setting: Primary care and specialty</p> <p>Patient Population: Patients with chronic condition (HCV)</p>		<p>Attribution: Qualifying episodes are identified using International Classification of Diseases, 10th Revision (ICD-10) codes, Current Procedural Terminology (CPT) codes, and HCPCS codes. Medicare beneficiaries with HCV are eligible to participate in this model.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: SVR rates are adjusted for patient-level influences on SVR likelihood, including disease stage and patient age. Patients with HCV are nested within hospitals, as patients receiving care in the same facility may have similar outcomes. The submitter planned to stabilize rates for clinics in which the number of patients achieving SVR is small by including facility-specific random effects in risk-adjusted SVR rates.</p> <p>Benchmarking: Based on risk-adjusted facility-based HCV SVR rate, compared with other model participants (e.g., compared with the average among all participants)</p>	<p>updated annually to better evaluate quality standards over time. Facilities may also be able to internally calculate their own SVR progress throughout the year (by assessing EHRs of patients being treated by primary care physicians, specialists, or both) and make practice modifications to ensure meeting the annual threshold necessary to achieve shared savings.</p>
<p>Pulmonary Medicine, Infectious Disease and Critical Care Consultants Medical Group (PMA)</p> <p><i>(Regional/local single specialty practice)</i></p> <p>The COPD and Asthma Monitoring Project</p> <p>Not Recommended, 4/11/2017</p>	<p>Clinical Focus: COPD and/or asthma</p> <p>Providers: Pulmonary physicians</p> <p>Setting: Patient home</p> <p>Patient Population: COPD and asthma patients</p>	<p>Utilization measure(s): ED visits, hospital admissions</p> <p>Quality measure(s): Mortality, well-controlled patient conditions, “optimal” COPD and asthma care (including assessment and classification of COPD and asthma control using a validated instrument; stepwise approach to identify treatment options and adjust medication and other therapies, written patient self-management asthma action plan customized to take advantage of real time monitoring and early detection/intervention protocols; stepwise approach to identify treatment options and adjust medication and other therapies; patients over four years of age with flu shot (or flu shot recommendation); smoking cessation and advise where appropriate</p>	<p>How payment is adjusted for performance: The proposed two-sided risk arrangement would permit CMS to recoup up-front costs first, use number of chronic conditions as a risk-adjuster to find the target spending level, and then remaining savings from total Part A and B costs of care above the cost to CMS of the technology and of the PBPM payments would be shared, as well as would losses up to a stop loss percentage amount. The proposal does not specify how quality measures would affect payment.</p> <p>Requirements: N/A; model participation is restricted to physicians board-certified in pulmonary medicine.</p> <p>Attribution: Enrollment offered to all Medicare beneficiaries with a diagnosis of asthma or COPD at participating providers.</p> <p>Volume: N/A; initial target pilot enrollment of 2,000 patients, with the intention of scaling the service locally once the pilot is validated.</p> <p>Risk stratification or adjustment: Total cost of care benchmark is risk-adjusted for chronic conditions. Patients are grouped</p>	<p>Measures used for implementation: Total Medicare costs, ED visits, hospitalizations, and mortality</p> <p>Measures used for monitoring: Program participation, data system function, provider and patient usage of alerts, COPD and asthma care measures, and other clinical outcomes. Other clinical outcomes under consideration include ED utilization rate, hospital inpatient utilization rate, hospital readmission rate, and ambulatory care sensitive readmissions.</p> <p>How achievement is measured: Asthma and COPD outcomes; methodology not specified</p> <p>How improvement is measured: The submitter states that the technology infrastructure will allow participants to track and make real-time performance</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
		<p>Spending measure(s): Medicare Part A and Part B spending</p> <p>Patient experience measure(s): Patient experience survey at the end of each encounter</p>	<p>into three risk categories (low, medium, and high) based on their disease control.</p> <p>Benchmarking: National, chronic condition-adjusted benchmark developed in risk pool of patients with COPD</p>	<p>improvements to achieve asthma and COPD outcomes; methodology to measure such improvements is not specified.</p>
<p>Personalized Recovery Care (PRC) <i>(Regional/local single specialty practice)</i></p> <p>Home Hospitalization: An Alternative Payment Model for Delivering Acute Care in the Home</p> <p>Recommended for implementation, 3/26/2018</p>	<p>Clinical Focus: Inpatient services in home setting</p> <p>Providers: Admitting physician at facility receiving PRC payments; On-Call Physician; Recovery Care Coordinators</p> <p>Setting: Patient home</p> <p>Patient Population: Commercial and Medicare Advantage patients with acute conditions, based on approximately 150 DRGs</p>	<p>Utilization measure(s): Utilization of resources, ED visits</p> <p>Quality measure(s): Percentage of Episodes with Follow-Up PCP Appointment Scheduled Within 7 Days; Percentage of Episodes with Medication Reconciliation; Patient Safety – Percentage of Episodes with Adverse Events (Deep Vein Thrombosis [DVT], Pressure Ulcer, Fall with Injury); Functional Status Assessments (Using the Patient-Reported Outcomes Measurement Information System [PROMIS]) – Percentage of Episodes with Functional Status Assessments Completed for Each Patient; Hospital Readmissions</p> <p>Spending measure(s): Total cost of care</p> <p>Patient experience measure(s): Patient Experience – Percentage of Questions Answered with Top Box Response</p>	<p>How payment is adjusted for performance: The program utilizes a retrospective bundled payment model. The bundled payment consists of two primary components: 1) a risk payment for delivering high-quality care as compared with the target cost of care (i.e., the “Target Bundled Rate”); and 2) a per episode payment made for the care being provided in lieu of an acute care hospitalization (“Home Hospitalization Payment”). If the total related costs are more than the Target Bundled Rate, the PRC Operators will be liable for 100% of the difference, up to 10% of the Benchmark Rate. Each metric for which PRC Operators satisfy requirements will result in the PRC Operators receiving 20% of the savings. If all five metrics are satisfied, the PRC Operators would receive 100% of the savings, whereas meeting none of the metrics would result in zero savings payments received by the PRC Operators in the reporting time frame.</p> <p>Requirements: To be eligible for shared savings, providers must meet or exceed benchmarks for performance measures tracking clinical quality, patient engagement, and program personalization on an episode basis.</p> <p>Attribution: Claims with qualifying DRGs are aligned to the furnishing provider</p> <p>Volume: N/A; a network approach, which involves the PRC Operators contracting with existing ancillary providers in the local market, is intended to increase program viability, especially for smaller practices, and mitigate the need to unnecessarily admit patients who do not qualify in order to meet a volume threshold.</p> <p>Risk stratification or adjustment: Yes, for patient clinical characteristics. The PRC Operators also propose excluding beneficiaries who have the following clinical characteristics:</p>	<p>Measures used for implementation: Quality measures and patient experience measures</p> <p>Measures used for monitoring: Total cost of care, utilization of resources, and adverse events. The PRC Operators collect data on a weekly basis for core operational metrics (including insurance type, number of calls received, number of calls placed to patients, number of patients screened, time to admit, number of readmits, average length of stay, and time to upload Continuity of Care Document) and claims data. Patient satisfaction metrics are collected at the end of the 30-day episode. The submitter noted that measuring outcomes for a period of time beyond the 30-day episode and comparing outcomes to historical benchmarks is another form of monitoring that could be incorporated.</p> <p>How achievement is measured: The target is >90% for Percentage of Episodes with Follow-Up PCP Appointment Scheduled Within 7 Days; Percentage of Episodes with Medication Reconciliation; Functional Status Assessments (Using PROMIS) – Percentage of Episodes with Functional Status Assessments Completed for Each Patient; and Patient Experience – Percentage of Questions Answered with Top Box Response. The target is < 3% for Patient Safety –</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>end-stage renal disease, hospice enrollment, or initial admissions to the intensive care unit.</p> <p>Benchmarking: Based on historical, episodic expenditures for each condition plus a three percent discount to derive target prices</p>	<p>Percentage of Episodes with Adverse Events (Deep Vein Thrombosis [DVT], Pressure Ulcer, Fall with Injury).</p> <p>How improvement is measured: N/A; clinical data from previous operators of this model demonstrate improvements, including a 19% reduction in total cost of care, 33% reduction in mean length of stay, 24% reduction in readmissions, and 20% reduction in mortality.</p>
<p>Renal Physicians Association (RPA) <i>(Provider association and specialty society)</i></p> <p>Incident ESRD Clinical Episode Payment Model</p> <p>Recommended for implementation, 12/18/2017</p>	<p>Clinical Focus: End-stage renal disease (ESRD)</p> <p>Providers: Nephrologists, PCPs</p> <p>Setting: Dialysis centers</p> <p>Patient Population: Patients with chronic condition (incident ESRD)</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Advanced Care Planning; Catheter % for ICHD (90 and 180 day); Optimal start: day 1 of outpatient dialysis with no catheter in place (ICHD/HHD) or initiate dialysis on PD; Fistula rate of all permanent vascular access for ICHD & HDD (180 day); Home dialysis % (PD and HDD); Referral to Transplant; Patient Centeredness: Karnofsky Functionality Score</p> <p>Spending measure(s): Medicare Part A and Part B spending</p> <p>Patient experience measure(s): Patient-Reported Outcomes Measurement Information Systems (PROMIS)</p>	<p>How payment is adjusted for performance: Quality scores determine physician's eligibility and amount of shared savings. Physicians' quality scores based on performance on patient-centered quality measures (0-100) determine the percentage of overall shared savings the physician receives. The higher the quality score, the higher amount of shared savings to the participant. Further, physicians choosing to participate in MIPS APM vs. Advanced APM will determine the total upside shared savings and downside risk. There is also a one-time financial incentive/bonus payment for patient receiving a kidney transplantation.</p> <p>Requirements: Physicians must attain a quality score of at least 30 and meet financial benchmarks to qualify for shared savings.</p> <p>Attribution: Attribution is determined based on the date of first dialysis treatment entered on the CMS Form and subsequent Medicare Part B claims for a patient by a participating provider</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Medicare beneficiary's most recent hierarchical condition category (HCC) risk score normalized so that an average risk patient would have a score of 1; values greater than 1 would indicate comorbidities associated with higher costs of care, values less than 1 would indicate lower costs of care.</p> <p>Benchmarking: Will be set for first six months of dialysis care using the RPA Registry, which gathers data from many practices</p>	<p>Measures used for implementation: All quality measures</p> <p>Measures used for monitoring: Quality and patient experience measures</p> <p>How achievement is measured: Achievement is measured using EHR and claims data. Physicians receive a quality score between 0-100. Physicians receive 15 points for having 100% Advance Care Plans on file; 5 points for >25% relative reduction (RR) from day 0 to day 90 or <60% absolute rate in catheter % for ICHD by day 90; 5 points for >50% relative reduction (RR) from day 0 to day 180 or <40% absolute rate in catheter % for ICHD by day 180; 20 points for having >50% in Optimal Start: Day 1 of outpatient dialysis with no catheter in place (ICHD/HHD) or initiate dialysis on PD; 10 points for >70% for Fistula Rate of all permanent vascular access for ICHD & HDD; 15 points for >20% total patient months on home dialysis modality; 10 points each for reporting Referral to Transplant, Patient Centeredness, and Patient Experience. Some measures include a sliding scale for points.</p> <p>How improvement is measured: N/A</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			and locations. Further, historical Medicare Parts A and B costs (from claims data) for the first six months of incident dialysis patient care divided by the average of normalized HCC scores of these same patients	
<p>University of Chicago Medicine (UChicago)ⁱⁱⁱ</p> <p><i>(Academic Institution)</i></p> <p>The Comprehensive Care Physician Payment Model (CCP-PM)</p> <p>Recommended for limited-scale testing, 9/7/2018</p>	<p>Clinical Focus: Frequently hospitalized patients</p> <p>Providers: Inpatient and Outpatient providers</p> <p>Setting: Home care and rehabilitation</p> <p>Patient Population: Patients at increased risk of hospitalization</p>	<p>Utilization measure(s): Hospitalizations</p> <p>Quality measure(s): Empanelment process, percentage of inpatient and outpatient general medical care provided by the participating clinician</p> <p>Spending measure(s): Total cost of care</p> <p>Patient experience measure(s): HCAHPS, self-rated mental health, patients' rating of providers</p>	<p>How payment is adjusted for performance: Care continuity fee given to providers who meet benchmarks for providing their patients with both inpatient and outpatient care. Providers continue to be subject to financial incentives/penalties under their current model (e.g., MIPS, MSSP). Care continuity fees are not at risk to not penalize providers twice.</p> <p>Requirements: Physicians can receive care continuity fees only if they are able to care for patients in clinic and in the hospital.</p> <p>Attribution: Eligible physicians can enroll a panel of CCP-PM patients for which they intend to provide an increased proportion of inpatient and outpatient general medical care, and eligible patients join the program by enrolling in the CCP-PM panel of a participating physician; alignment can continue for up to six years, with pathways based on whether the patient has had an additional hospitalization.</p> <p>Volume: The submitter noted a maximum panel size of 300 patients, with a typical panel size of 200 patients. This limitation was implemented to encourage a focus on high utilizers. The submitter also noted that they wanted to provide flexibility for individual providers to set a panel size most appropriate for their practice, and that there are challenges in variability with low patient volumes.</p> <p>Risk stratification or adjustment: N/A. The submitter noted that the high-risk population the CCP-PM targets poses significant challenges to appropriately risk-adjusting quality metrics.</p> <p>Benchmarking: Yes, based on percent provision of inpatient care and outpatient general medicine care for their enrolled patients</p>	<p>Measures used for implementation: Patient experience (HCAHPS), self-rated mental health status, hospitalizations, total cost of care</p> <p>Measures used for monitoring: Enrollment in the model</p> <p>How achievement is measured: Percent provision of inpatient and outpatient care for enrolled patients relative to established benchmark. Measurement methodology is not specified.</p> <p>How improvement is measured: N/A</p>

ⁱⁱⁱ Participating physicians continue to be responsible for financial and quality measures under their payment model (e.g., the Merit-based Incentive Payment System [MIPS], the Medicare Shared Savings Program [MSSP]).

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>The University of Massachusetts Medical School (UMass) <i>(Academic institution)</i></p> <p>Eye Care Emergency Department Avoidance</p> <p>Not recommended, 11/8/2019</p>	<p>Clinical Focus: Eye care</p> <p>Providers: Optometrists and ophthalmologists</p> <p>Setting: Practices and other entities employing eye care professionals</p> <p>Patient Population: Patients with non-emergent eye conditions</p>	<p>Utilization measure(s): Number of qualifying ED-avoidable visits (utilization of office-based services vs. utilization of ED for non-emergent eye conditions)</p> <p>Quality measure(s): Office-based eye care seven-day post-visit adverse event rate (including unscheduled ED visits, unscheduled hospital inpatient admissions, blindness/permanent vision damage, or death occurring within seven days of eligible office visit)</p> <p>Spending measure(s): Total cost of care (including payments to EDs, payments to eligible providers, program costs)</p> <p>Patient experience measure(s): Standardized patient survey</p>	<p>How payment is adjusted for performance: Providers who meet or exceed the target number of qualifying ED-avoidable visits and upheld or improved quality performance will receive shared savings payments. If providers do not meet utilization targets or quality performance thresholds, their financial loss will equal the minimum of 8% of payments for qualifying visits during the performance year.</p> <p>Requirements: To be eligible for shared savings and continued participation, providers must meet quality performance thresholds for both patient safety and patient experience and meet provider-specific utilization targets.</p> <p>Attribution: Not specified</p> <p>Volume: N/A; The proposal states that the model was designed based on the experiences of smaller optometry practices (averaging 1.8 clinicians/per practice). Smaller practices demonstrated the ability to implement the model and provide quantifiable value.</p> <p>Risk stratification or adjustment: The proposal states that risk adjustment is not necessary because payment methodology is based on FFS rates, although the proposal notes that adverse event rates could be adjusted for age, gender, and other risk factors.</p> <p>Benchmarking: Medicare and other participating payers will establish target utilization goals for the participating providers based on provider-specific historical volume of ED-avoidable visits. Payers will establish a baseline year and then select a specific percentage increase over base year visits to determine provider-specific target numbers for each participating provider.</p>	<p>Measures used for implementation: Utilization, quality, and patient experience measures</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: Quality thresholds are defined as achieving an adverse event rate less than or equal to the adverse event rate for ED-avoidable eye conditions in the ED setting, as well as receiving a score of 3 points or higher on each question in the standardized patient survey.</p> <p>How improvement is measured: N/A</p>
<p>The University of New Mexico Health Sciences Center (UNMHSC) <i>(Academic institution)</i></p>	<p>Clinical Focus: Cerebral emergent care; telemedicine</p> <p>Providers: Neurologists and neurosurgeons; providers in rural</p>	<p>Utilization measure(s): Number of sites with ready telehealth systems (THS) in place; number of sites with trained telehealth administrator; number of sites with trained health care providers; number of sites with trained radiology technicians; proportion of patients with neuro-emergent conditions enrolled per</p>	<p>How payment is adjusted for performance: Performance measures are not tied to payment</p> <p>Requirements: N/A</p> <p>Attribution: Participating neurologists and neurosurgeons will consult with patients presenting with cerebral neuro-emergent conditions in emergency rooms in underserved regions; there</p>	<p>Measures used for implementation: N/A</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: N/A</p> <p>How improvement is measured: N/A</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>ACCESS Telemedicine: An Alternative Healthcare Delivery Model for Rural Emergencies</p> <p>Recommended for implementation, 9/16/2019</p>	<p>and community systems</p> <p>Setting: Inpatient; outpatient; or emergency department</p> <p>Patient Population: Patients with neurological emergencies</p>	<p>site; number of local follow-ups completed; number of consults completed within 60 minutes</p> <p>Quality measure(s): Imaging results for acute stroke patients within 45 minutes; timeliness of Emergency Medicine Care; Hospital-Wide-All-Cause Unplanned-Readmission measure; time to treatment; proportion of patients transported to UNMHSC or other tertiary referral center; rate of tissue plasminogen activator (tPA) administration</p> <p>Spending measure(s): Total Cost of Care Population-Based PMPM Index adapted to diagnostic codes for neuro-emergent conditions</p> <p>Patient experience measure(s): Patient Experience Questionnaire; Telemedicine Satisfaction Questionnaire</p>	<p>is no attribution of specific patients to providers since it is not tied to payment.</p> <p>Volume: N/A; the proposal does not specify the percentage who visit the ED for traumatic brain injury or their expectations of number in the model.</p> <p>Risk stratification or adjustment: N/A</p> <p>Benchmarking: Neurology consults are \$850 and neurosurgical consults are \$1,200; these were set using a fair market value process that took into account the bundled costs necessary to provide telemedicine services from an academic medical organization. However, actual benchmarking is not done since payment is not adjusted for performance.</p>	

Exhibit G2. Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that were Found to Meet Criterion 2 and were Found to Not Meet Criterion 4

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>American College of Allergy, Asthma & Immunology (ACAAI)</p> <p><i>(Provider association/specialty society)</i></p> <p>Patient-Centered Asthma Care Payment (PCACP)</p> <p>Referred for other attention by HHS, 6/22/2020</p>	<p>Clinical Focus: Asthma care</p> <p>Providers: Allergists; immunologists; pulmonologists; PCPs; other providers</p> <p>Setting: Emergency department</p> <p>Patient Population: Patients with asthma and asthma-like symptoms</p>	<p>Utilization measure(s): Average number of months during which the Diagnosis and Initial Treatment Payments were billed before a diagnosis was assigned</p> <p>Quality measure(s): Percentage of patients with improvement in asthma-like symptoms based on self-reports; percentage of patients with improved spirometry measures; percentage of patients with ED visits or urgent care visits for asthma-like symptoms</p> <p>Spending measure(s): Standardized average total per-patient spending on allergy testing, asthma medications, urgent care visits for asthma-like symptoms, ED visits related to asthma, and hospitalizations related to asthma</p> <p>Patient experience measure(s): percentage of patients rating access to physician as “very good” or “excellent” via patient survey</p>	<p>How payment is adjusted for performance: Asthma Care Teams receive the default payment level for each patient if the team scored “good” on all performance measures; payments are increased or decreased (up to + or - 5% to increase over time to + or - 9%) if team scored “high” or “low” on some performance measures.</p> <p>Requirements: There are three types of bundled payments the Asthma Care Team receives depending on patient status category. The Asthma Care Team is required to meet minimum quality standards in order to bill and receive the Diagnosis and Initial Treatment Payment.</p> <p>Attribution: Patients are not attributed; patient designates the physician(s) as their “Asthma Care Team” and agrees to receive all asthma-related services from this team for a three-month period.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Payment amounts and performance measures would be stratified into subcategories based on severity of symptoms and comorbidities; payment amounts and subcategories differ depending on the patient status category.</p> <p>Benchmarking: Performance on measures is assessed by comparing the Asthma Care Team’s performance to the average performance of all Asthma Care Teams in the model in the current or previous year (depending on the type of measure).</p>	<p>Measures used for implementation: All measures listed</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: Care teams are assessed, and payment is based on performance measure scores.</p> <p>How improvement is measured: N/A</p>

Exhibit G3. Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that were Found to Not Meet Criterion 2 and were Found to Meet Criterion 4

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>American Academy of Hospice and Palliative Medicine (AAHPM)</p> <p><i>(Provider association/specialty society)</i></p> <p>Patient and Caregiver Support for Serious Illness (PACSSI)</p> <p>Recommended for limited-scale testing, 3/26/2018</p>	<p>Clinical Focus: Serious illness and palliative care</p> <p>Providers: Palliative care teams (PCT)</p> <p>Setting: Inpatient; outpatient; other palliative care settings</p> <p>Patient Population: Patients with serious illness</p>	<p>Utilization measure(s): Percentage of patients who died who received hospice care; Percentage of patients who died and were enrolled in hospice more than seven days before death; Percentage of patients who died and did not have any days in an ICU during the 30 days before death</p> <p>Quality measure(s): For years 1 and 2, completion of four applicable palliative care activities within 15 days of PACSSI enrollment: comprehensive assessment; screening for pain, dyspnea, nausea, and constipation; documentation of a discussion regarding emotional needs; documentation of a discussion about advance care planning; Beginning year 3, completion of six applicable palliative care activities within 15 days of PACSSI enrollment: same four listed above, as well as documentation of a discussion of spiritual concerns and completion of a structured assessment of caregiver needs and distress</p> <p>Spending measure(s): Total cost of care for enrolled patients (including PACSSI care management payments)</p> <p>Patient experience measure(s): Measures from patient admission survey: Likelihood of patient recommendation, timeliness of response to urgent needs, adequacy of treatment for pain and symptoms, patient’s percepts regarding quality of communication; Post-death survey for PACSSI enrollees; Hospice CAHPS survey for PACSSI enrollees transferring to hospice and dying within seven days of disenrollment from PACSSI</p>	<p>How payment is adjusted for performance: For the first two years of the model, PCTs are required to report on measures, but payment will not be tied to performance on these measures. Starting in year 3, PCTs receive a composite score that equally weights performance across three categories of quality metrics: patient-reported outcomes; completion of care processes; and utilization of health care services. Track 1 PCTs are subject to positive and negative performance incentives of up to 4% of total care management fees received for a year based on their performance on quality and spending; Track 2 (starting in Year 3) involves shared savings and losses based on total cost of care.</p> <p>Requirements: PCTs must meet minimum quality benchmarks to be eligible for payment participation.</p> <p>Attribution: Patients are not attributed; patients have to designate the PCT as their primary provider; PCTs are responsible for all patients who have chosen them to be their provider.</p> <p>Volume: PCT teams are required to accept any patient living in the service area who meets the eligibility criteria; the model will establish pre-defined capacity limits for number of patients (not further specified). If PCTs fail to meet minimum participation standards, they will be terminated from the model.</p> <p>Risk stratification or adjustment: Patients are assigned to one of two tiers (moderate- and high-complexity) based on criteria including diagnosis of serious illness, function, and health care utilization. Higher payment amounts are provided to PCTs for delivering care to tier 2 patients (\$650 vs. \$400 per beneficiary per month). Further, monthly payments are adjusted based on current Geographic Practice Cost Indices and patient’s primary site of care (home vs. facility).</p> <p>Benchmarking: The model proposes to collect and analyze data during the first two years to establish benchmarks for each performance measure. The proposal does not specify the benchmark that will be used to assess total cost of care.</p>	<p>Measures used for implementation: Beginning year 3, all listed measures are used to adjust payment</p> <p>Measures used for monitoring: All listed measures</p> <p>How achievement is measured: Achievement is measured based on reported outcomes, clinical quality performance measures, and utilization against benchmarks (to be established for each performance measure by year 3).</p> <p>How improvement is measured: Risk-based payment for improvement in patient/caregiver outcomes, care processes, and utilization/cost (not further specified)</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>American Society of Clinical Oncology (ASCO)</p> <p><i>(Provider association/specialty society)</i></p> <p>Patient-Centered Oncology Payment Model (PCOP)</p> <p>Referred for other attention by HHS, 9/15/2020</p>	<p>Clinical Focus: Oncology</p> <p>Providers: Clinicians, including hematologists and oncologists</p> <p>Setting: Oncology specialty practices</p> <p>Patient Population: Oncology practice patients</p>	<p>Utilization measure(s): Unplanned hospital admissions per treatment month; emergency and observation care visits per treatment month; supportive and maintenance care drug costs per treatment month</p> <p>Quality measure(s): Chemotherapy administered to patients with metastatic solid tumor; GCSF administered to patients who received chemotherapy for metastatic cancer; Care Plan; Preventive care and screening for clinical depression; Preventive care and screening for high blood pressure; Trastuzumab received by patient with AJCC stage I-III and HER2 positive breast cancer; KRAS gene mutation testing performed for patients with metastatic colorectal cancer; Patients with Metastatic Colorectal Cancer and KRAS gene mutation; Proportion of patients receiving chemotherapy in the last 14 days of life, Proportion Not Admitted to Hospice, and Bone Density Evaluation for Patients with Prostate Cancer and Receiving Androgen Therapy.</p> <p>Spending measure(s): Total Cost of Care</p> <p>Patient experience measure(s): Patient Satisfaction Surveys, including the Oncology Medical Home Patient Satisfaction Survey</p>	<p>How payment is adjusted for performance: Providers receive three payments: monthly Care Management Payments (CMP), Performance Incentive Payments (PIP), and adjustments to FFS reimbursement. A portion of the CMP will be allocated to a PIP. PIPs will be positively or negatively adjusted based on provider success in adherence to clinical treatment pathways, quality metrics, and cost reduction. There are two tracks: Track 1 participants continue to receive FFS reimbursement in addition to the CMPs; Track 2 participants participate in the Consolidated Payments for Oncology Care (CPOC) where practices can bundle 50% or 100% of the value of specified services. 10% of the amount bundled will be subject to the same performance adjustment as PIPs times a 1.4 multiplier.</p> <p>Requirements: Practices must follow specified care delivery requirements, including providing education on PCOP model, financial counseling services, education on diagnoses, and other patient engagement requirements.</p> <p>Attribution: Episodes are attributed to a provider or practice based on the billing provider for the Cancer Treatment CMP or the billing of specific treatments.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Practices will utilize data from comprehensive patient assessments to identify patients at higher risk for symptoms, complications, and non-adherence to their cancer treatment plan. Risk stratification methods may include an algorithm based on diagnoses or a structured scoring system administered by clinicians.</p> <p>Benchmarking: Established benchmarks; however, the proposal does not specify what will be used to establish benchmarks.</p>	<p>Measures used for implementation: All listed utilization, quality, and spending measures</p> <p>Measures used for monitoring: All listed measures</p> <p>How achievement is measured: Achievement is measured based on adherence and progress in quality of care and targeted cost metrics. Providers are responsible for calculating the aggregate pathway adherence rate, quality metric adherence, and the cost of care and report to the Oncology Steering Committee on a quarterly basis. The Oncology Steering Committee will be responsible for weighting performance categories for calculation of an aggregate performance score. PIP amounts will be adjusted based on the aggregate performance score.</p> <p>How improvement is measured: If providers do not meet minimum expectations, CMP and PIP amounts may be suspended, and providers will need to develop an improvement plan.</p>
<p>Dialyze Direct (Dialyze Direct)</p> <p><i>(Regional/local single specialty practice)</i></p> <p>APM for Improved Quality and Cost in</p>	<p>Clinical Focus: End-stage renal disease (ESRD)</p> <p>Providers: Nephrologists</p>	<p>Utilization measure(s): Hospitalizations; re-hospitalizations; Emergency Department (ED) visits; observation hospital events</p> <p>Quality measure(s): Post-hospital discharge outcomes; complications of transportation (e.g., falls, fractures)</p>	<p>How payment is adjusted for performance: N/A; Bundled payment model with ability to receive shared savings, as well as a one-time additional payment for efforts related to educating patients on the benefits of on-site staff-assisted home dialysis in the nursing home</p> <p>Requirements: N/A</p>	<p>Measures used for implementation: Not specified.</p> <p>Measures used for monitoring: Not specified.</p> <p>How achievement is measured: Not specified.</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>Providing Home Hemodialysis to Geriatric Patients Residing in Skilled Nursing Facilities</p> <p>Recommended for attention, 9/6/2018</p>	<p>Setting: Patient home</p> <p>Patient Population: Geriatric dialysis patients residing in SNFs</p>	<p>Spending measure(s): All Medicare Part A and Part B costs with the exception of those attributable to transplantation</p> <p>Patient experience measure(s): Patient-Reported Outcome Measures (PROMs) such as the In-Center Hemodialysis Survey Consumer Assessment of Healthcare Providers and Systems (ICH-CAHPS) and Kidney Disease Quality of Life 36 (KDOL-36), which will be tracked and compared with patient experiences with conventional dialysis.</p>	<p>Attribution: Attribution of the incident dialysis patient is determined by the date of the patient's admission to the skilled nursing facility (SNF). An episode of care will be the time a patient resides in the SNF.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: N/A</p> <p>Benchmarking: Participating physician's financial benchmark is the dollar value of the average Medicare cost necessary to provide one round-trip to a nephrologist's office. The benchmark used to quantify overall value is the dollar amount of included Medicare Part A and Part B non-randomized cost derived from the comparison of two groups of patients: a prospective cohort of patients residing in a SNF receiving on-site, staff-assisted, mode of dialysis technology (more frequent dialysis [MFD]), and a matched retrospective cohort of patients residing in a SNF receiving conventional, predominantly off-site, in-center home hemodialysis (HD).</p>	<p>How improvement is measured: Not specified</p>
<p>Jean Antonucci, MD (Dr. Antonucci)</p> <p><i>(Independent individual)</i></p> <p>An Innovative Model for Primary Care Office Payment</p> <p>Recommended for limited-scale testing, 9/6/2018</p>	<p>Clinical Focus: Primary care</p> <p>Providers: Primary care providers, nurse practitioners</p> <p>Setting: Primary care practices</p> <p>Patient Population: Medicare beneficiaries</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): N/A</p> <p>Spending measure(s): PBPM payments</p> <p>Patient experience measure(s): How's Your Health (HYH)</p>	<p>How payment is adjusted for performance: \$60 PMPM for low- and medium-risk patients, and \$90 PMPM for high-risk patients. Further, 15% of annual income will be withheld; if participants do not meet quality and cost benchmarks, they may lose this income.</p> <p>Requirements: Participants will employ the HYH survey to all patients, which includes many patient-reported measures. Using HYH results from other users as a benchmark, participants will be assessed on whether they meet, exceed, or are below the benchmark.</p> <p>Attribution: Four-step process (AAFP): 1. Patient selection of primary care physician and team; 2. Primary Care Visit Events: Wellness Visits; 3. Primary Care Visit Events: All other E/M Visits; and 4. Primary Care Prescription and Order Events.</p> <p>Volume: Cap at 1,500 patients per physician.</p> <p>Risk stratification or adjustment: Using the What Matters Index (WMI) part of HYH measures, patients will be grouped into low-, medium-, and high-risk pools. Risk adjustment should occur</p>	<p>Measures used for implementation: HYH measures</p> <p>Measures used for monitoring: Not specified.</p> <p>How achievement is measured: Quality and cost-effective care will be measured against benchmarks (using results from HYH); participants that meet or exceed benchmarks will retain full capitation; failure to meet benchmarks will result in loss of up to 15% of annual income and possible removal from the APM.</p> <p>How improvement is measured: Not specified</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>annually, allowing physicians to review, add, and remove patients from the formal list the payer supplies to them.</p> <p>Benchmarking: Baseline will be a set time period prior to performance year. Benchmarks should be held steady for at least two years. National benchmark data will come from HYH.</p>	
<p>Upstream Rehabilitation (Upstream)</p> <p>(Regional/local single specialty practice)</p> <p>CMS Support of Wound Care in Private Outpatient Therapy Clinics: Measuring the Effectiveness of Physical or Occupational Therapy Intervention as the Primary Means of Managing Wounds in Medicare Recipients</p> <p>Not recommended, 5/11/2019</p>	<p>Clinical Focus: Chronic wound care</p> <p>Providers: All providers in the patient’s care plan</p> <p>Setting: Physical and occupational therapy centers</p> <p>Patient Population: Patients with chronic wounds</p>	<p>Utilization measure(s): Total time in treatment (duration of care)</p> <p>Quality measure(s): Functional outcomes of patients with open wounds utilizing the Bates-Jensen Wound Assessment; Functional independence using the Functional Independence Measure (FIM); Objective measurements (e.g., range-of-motion, strength, or edema); Disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH); Lower Extremity Functional Scale (LEFS); Pain Scale; Oswestry Disability Index</p> <p>Spending measure(s): Total cost of each patient treatment, including wound care supplies; cost savings of utilizing physical and occupational therapists in outpatient, private settings versus traditional outpatient hospital-based wound care centers</p> <p>Patient experience measure(s): Patient satisfaction</p>	<p>How payment is adjusted for performance: Claim refunded to CMS if minimum standards of improvement are not met. Clinicians can receive a 3% savings bonus for achieving average reimbursement costs below risk-adjusted thresholds.</p> <p>Requirements: Participating clinicians that do not meet cost or patient satisfaction targets for two subsequent quarters will be removed from the program.</p> <p>Attribution: N/A</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: Patients are categorized into low-, medium-, and high-complexity, based on comorbidities and complicating factors that would extend the level of care needed to address wound care issues. Cost targets vary based on patient complexity.</p> <p>Benchmarking: A benchmark of 80% is set for patient satisfaction scores across all claims submitted.</p>	<p>Measures used for implementation: Functional outcomes, patient satisfaction, total cost of each patient treatment, duration of care</p> <p>Measures used for monitoring: Functional outcomes, patient satisfaction, total cost of each patient treatment, duration of care</p> <p>How achievement is measured: Patient satisfaction scores must exceed 80% across all claims submitted. Claims must demonstrate that each patient participating in the program achieved a minimal clinically-important difference in one of the reported outcomes (excluding patient satisfaction).</p> <p>How improvement is measured: N/A</p>

Exhibit G4. Proposal Features, Technical Issues, and Potential Gaps Related to Current Performance Measures for Proposals Reviewed by PTAC as of September 2020 that were Found to Not Meet Criterion 2 or Criterion 4, were Withdrawn, or were Determined to be Out of Scope by PTAC

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
<p>American Academy of Neurology (AAN)</p> <p><i>(Provider association and specialty society)</i></p> <p>The Patient-Centered Headache Care Payment (PCHCP)</p> <p>N/A - Withdrawn</p>	<p>Clinical Focus: Neurology</p> <p>Providers: PCPs; neurologists; other physicians with expertise in headache care</p> <p>Setting: Inpatient or outpatient in primary care; patient home</p> <p>Patient Population: Patients with headaches</p>	<p>Utilization measure(s): Axon 22 MIPS 419 Headache- advanced brain imaging NOT ordered; Average per-patient rates of visits to emergency departments for management of headaches; Average per-patient rates of admission and duration of stay to the hospital for management of headaches</p> <p>Quality measure(s): Frequency, severity, and disability of headaches, and changes from the patient’s baseline, using Medical Information Data Analysis System (MIDAS); Axon 13 Medication prescribed for acute migraine attack; Axon 25 Overuse of barbiturate and opioids for primary headache disorders; Axon 15 MIPS 435 Quality of life assessment for patients with primary headache disorder; MIPS 431 Preventive Care Screening: Unhealthy Alcohol Use: Screening and Brief Counseling; MIPS 134 Preventive Care and Screening: Screening for Clinical Depression and Follow-up Plan</p> <p>Spending measure(s): Average per-patient spending on headache-related medications; Average per-patient per-month total spending on (1) E/M visits related to headaches, (2) non-face-to-face billing codes, and (3) emergency department visits and urgent care center visits for headache</p> <p>Patient experience measure(s): Percent of patients rating access to providers and experience of care as “excellent”</p>	<p>How payment is adjusted for performance: Flexible fixed payments or add-on payments adjusted based on performance on measures. Participating physicians receiving quality designation of “good” on all performance measures would receive the default payment level; payment increased (+4%) if some measures were “high”; payment decreased (-4%) if some measures were “low.”</p> <p>Requirements: Participants must achieve quality scores of good/high to receive the default or increased payment. Participants that receive “low” scores for some measures still receive payment, but it is decreased in amount.</p> <p>Attribution: Patients are not attributed; they must opt-in to the model.</p> <p>Volume: N/A; Small practices could have their performance measured over a longer period of time (e.g., two years) in order to have more reliable measures with smaller numbers of patients.</p> <p>Risk stratification or adjustment: The payment amounts and performance measures for each payment category would be adjusted to reflect the cost of treating more</p>	<p>Measures used for implementation: All performance measures</p> <p>Measures used for monitoring: Not specified</p> <p>How achievement is measured: For each measure, high-quality designation means performing in the 76-100th percentile; good quality 26-75th percentile; low-quality 0-25th percentile. The proposal states that AAN would work with CMS to define high, good, and low thresholds for performance measures.</p> <p>How improvement is measured: During first two years, participants will receive feedback reports from CMS detailing their performance on selected quality measures to help inform quality improvement and make adjustments.</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>complex patients. Payments are adjusted based on patient characteristics, such as frequency, headache severity (measured using MIDAS), select comorbidities, patient demographic information, and resource use. In addition, payment type varies depending on the category of headache patient: one-time payment for category 1 patients (most severe); monthly payment for category 2 patients; and add-on service for category 3 (less severe) patients.</p> <p>Benchmarking: Not specified</p>	
<p>Clearwater Cardiovascular and Interventional Consultants, MD, PA (CCC) <i>(Provider association/specialty society)</i> Bundled PCI Services N/A - Withdrawn</p>	<p>Clinical Focus: Percutaneous Coronary intervention services in lower-cost non-catheter labs Providers: Cardiovascular physicians Setting: Outpatient Patient Population: Patients with stable angina</p>	<p>Utilization measure(s): Resource utilization Quality measure(s): Appropriate Use Criteria (AUC); Clinical quality outcomes will be measured through the National Cardiovascular Data Registry (NCDR) system, such as complications, death, major adverse cardiac events, and other health outcomes Spending measure(s): Anchor procedure cost from claims data; 90-day post procedure cost Patient experience measure(s): Patient satisfaction surveys will be measured and compared with HCAPS patient satisfaction survey results.</p>	<p>How payment is adjusted for performance: Payment is not tied to performance measures; bundled episode-based model replacing FFS, with shared risk of managing 90-day post-procedure cost Requirements: N/A Attribution: N/A Volume: N/A Risk stratification or adjustment: Risk-adjustment would be similar to BPCI-A, which employs a risk adjustment model that adjusts target prices based on hierarchical condition categories (HCC), HCC interactions, HCC severity, recent resource use, demographics, long-term institutional care, dementia, MS-DRGS/APCs, clinical episode category specific</p>	<p>Measures used for implementation: N/A Measures used for monitoring: NCDR CathPCI registry; CMS claims data; resource utilization from internal accounting; patient satisfaction surveys How achievement is measured: Clinical success is based on the reported outcomes, quality metrics, and patient satisfaction; financial success is based on the ability of Bundled PCI Services Providers to manage the 90-day post procedure cost at or below the CMS determine target price. How improvement is measured: Although not tied to payment, monitoring is performed through CCC's participation in the ACC CathPCI Registry and through patient satisfaction surveys.</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>adjustments, and COVID-19 infection rate.</p> <p>Benchmarking: Use an existing data registry that includes benchmarks; metrics are quantifiable through existing national criteria and evaluated against benchmark databases.</p>	
<p>Community Oncology Alliance (COA) <i>(Non-profit organization)</i> Oncology Care Model 2.0 N/A - Withdrawn</p>	<p>Clinical Focus: Cancer</p> <p>Providers: Individuals or groups of medical oncologists providing services to patients</p> <p>Setting: Patient home</p> <p>Patient Population: Cancer patients (noted the MA plans may choose to limit types of cancers covered)</p>	<p>Utilization measure(s): Emergency department (ED) utilization; inpatient utilization, Part A post-acute services; selected Part B outpatient services; Part B and D chemotherapy and drug fills; ED visits and inpatient hospitalizations due to complications from chemotherapy; hospice and end-of-life services utilization</p> <p>Quality measure(s): Seven measures, including comprehensive care plan provided to the patient; adherence to recognized pathway and treatment guidelines; screening for clinical depression and follow-up plan; a survivorship care plan provided to the patient; Pneumococcal vaccination provided to older adults; proportion of patients with cancer receiving chemotherapy in the last 14 days of life; proportion of patients with cancer that died but without being admitted to hospice</p> <p>Spending measure(s): Total cost of care; Part A costs for inpatient care and post-acute and long-term care; institutional and non-institutional Part B costs; Part B and D costs for cancer-related services and drugs; total beneficiary deductible and coinsurance costs for Parts A, B, and D</p> <p>Patient experience measure(s): Reported through the Oncology Medical Home (OMH) patient survey</p>	<p>How payment is adjusted for performance: Episode-based payment with shared risk; participants retain a percentage of shared savings depending on their benchmarked quality scores and their total cost of care compared with the benchmark. Participants receive care management fee payments (\$160 per member per month [PMPM]) and initial trigger amounts (~\$150 per patient upon submission of a G code).</p> <p>Requirements: The average of the seven quality measures must be at or above the payer's measures for the grouped practices for that state. The percentage of savings is based on how teams compare to the average of other OCM 2.0 participants.</p> <p>Attribution: Patients identified through the submission of a G-code on claims, corresponding to the recommended cancer treatment plan, would be attributed to the participating team. Episodes are six months in length.</p> <p>Volume: N/A</p>	<p>Measures used for implementation: Seven quality measures listed</p> <p>Measures used for monitoring: OMH patient survey</p> <p>How achievement is measured: Participants that receive a quality score that is above the average of other OCM 2.0 participants receive their full allotment of shared savings; participants that are at 49% of the average would receive 49% of their allotted savings; and participants that are at 5% of the average would receive 5% of the shared savings. Exact percentages (and not quartiles) are used when calculating shared savings.</p> <p>How improvement is measured: Not specified.</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>Risk stratification or adjustment: Shape a mutually acceptable risk methodology by engaging a group of cancer care providers and CMMI staff; this team would improve the CMMI 12-step process for risk methodology targets and hierarchical condition categories (HCCs).</p> <p>Benchmarking: The total cost of care for the entire episode period is compared with other cancer care teams within that state.</p>	
<p>Digestive Health Network, Inc. (DHN) (Provider association/specialty society)</p> <p>Comprehensive Colonoscopy Advanced Alternative Payment Model for Colorectal Cancer Screening, Diagnosis and Surveillance</p> <p>N/A - Withdrawn</p>	<p>Clinical Focus: Gastrointestinal (GI)/digestive health; colorectal cancer (CRC)</p> <p>Providers: All providers within the care team</p> <p>Setting: Outpatient</p> <p>Patient Population: Patients undergoing colorectal cancer screening, diagnosis, and surveillance using colonoscopy</p>	<p>Utilization measure(s): Rates of site-of-service utilization (Ambulatory Surgical Center [ASC] vs. hospital outpatient department); utilization of pathology; repeat procedures; incomplete procedures; Seven-Day Risk-Standardized Hospital Visit Rate after Colonoscopy; ED utilization rate; surveillance intervals</p> <p>Quality measure(s): MIPS quality measures: Colonoscopy Interval for Patients with a History of Adenomatous Polyps – Avoidance of Inappropriate Use; Appropriate Follow-Up Interval for Normal Colonoscopy in Average Risk Patients; Age Appropriate Screening Colonoscopy; Preventive Care and Screening: BMI Screening and Follow-Up Plan; Documentation of Current Medications in the Medical Record; Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention; Screening Colonoscopy Adenoma Detection Rate; Photo documentation of Cecal Intubation; Preventive Care and Screening: Unhealthy Alcohol Use: Screening and Brief Counseling; PHQ-2 screen for depression; Method of sedation; Adequacy of colonoscopy preparation</p> <p>Spending measure(s): Episode cost of care</p>	<p>How payment is adjusted for performance: Prospective episode-based model with retrospective reconciliation; adjust payments based on the quality of care delivered</p> <p>Requirements: Payment is tied to reducing repeat procedures, increasing ASC utilization, and conducting follow-up at appropriate intervals; yearly retrospective reconciliation with downside payment adjustment for providers who fail to meet the re-do target for the CY, fail to meet the ASC utilization target for the CY, and/or fail to meet the quality criteria for surveillance follow-up intervals based on endoscopic and pathology findings.</p> <p>Attribution: Patients are attributed based on ICD-10 codes for screening, surveillance, and diagnostic colonoscopy</p>	<p>Measures used for implementation: Rate of repeat procedures; incomplete procedures; adequacy of colonoscopy preparation; rate of site-of-service utilization (ASC vs. hospital outpatient department); Seven-Day Risk-Standardized Hospital Visit Rate after Colonoscopy; ED utilization rate; surveillance intervals; patient experience measures; could be other measures not specified</p> <p>Measures used for monitoring: Rates on MIPS measures; patients satisfaction measures; rates of site-of-service utilization; method of sedation; incomplete procedures; adequacy of colonoscopy preparation; utilization of pathology; surveillance intervals; episode cost of care</p> <p>How achievement is measured: Participants have a target of</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
		<p>Patient experience measure(s): Overall satisfaction with experience; Professionalism of the non-medical office staff; Communication with physician being easy to understand; Included in decision-making related to care; Information provided on what to expect during and after the colonoscopy procedure; Preparation for the colonoscopy procedure; Likelihood to recommend the physician, practice, and the site of service where the procedure was performed</p>	<p>procedures. Episodes are one year in length.</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: The proposal states that risk adjustment is achieved through limiting participation to outpatients who are American Society of Anesthesiologists (ASA) class I-III.</p> <p>Benchmarking: N/A</p>	<p>reducing poor prep/incomplete procedure rate initially to 7%, with a subsequent reduction of 1% per year until a 4% repeat procedure rate is achieved while keeping complications (ASC-12: Facility Seven-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy and OP-32: Facility Seven-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy) minimal. If the repeat rate is less than the target, then savings are distributed to the endoscopist and the anesthesia professional; if repeat rate is higher than the target, then up to 4% of payment is withheld from the endoscopist and anesthesia professional. Further, participants have a target of a 60% ASC utilization rate. If the ASC utilization is greater than the target, the endoscopist receives a 5% incentive bonus for all procedures performed in the ASC setting; if the ASC utilization is under the target, then up to 4% of payment is withheld from the endoscopist. However, if ED utilization is greater than 0.1%, this reconciliation amount is withheld from the endoscopist's incentive payment. Finally, less than 90% adherence to recommended intervals for surveillance will result in a 2% withhold.</p> <p>How improvement is measured: Reduction in surveillance</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
				procedures performed at an inappropriately early interval
<p>Zhou Yang, PhD, MHP (Dr. Yang) <i>(Individual)</i> Medicare 3 Year Value Based Payment Plan (Medicare 3VBPP) N/A, 6/29/2017</p>	<p>Clinical Focus: Broad Providers: Not specified Setting: Broad Patient Population: Broad</p>	<p>Utilization measure(s): Per member per year (PMPY) ED visits, hospital nights Quality measure(s): Preventive screening and wellness care utilization, annual mortality rate Spending measure(s): PMPY Medicare Contribution/Expenditures, Out of Pocket Expenditures, Medicare prescription drugs cost Patient experience measure(s): Patient survey regarding getting needed care, getting care quickly, how well doctors communicate, plan’s customer choice, coordinated care, perceived value of care</p>	<p>How payment is adjusted for performance: N/A Requirements: N/A Attribution: N/A; voluntary participation into Medicare 3VBPP among community-dwelling beneficiaries (excluding nursing home residents) age 85 or lower without cognitive disability or severe mental illness Volume: Not specified Risk stratification or adjustment: Three-year budget constraints adjusted for inflation, age, demographics, geographic areas, and existing conditions Benchmarking: Not specified</p>	<p>Measures used for implementation: Cost and clinical care utilization, preventive service utilization, health outcomes, and patient satisfaction Measures used for monitoring: Cost and clinical care utilization, preventive service utilization, health outcomes, and patient satisfaction How achievement is measured: Assessment of quality and costs; methodology not specified How improvement is measured: Assessment of longitudinal per capita costs; methodology not specified</p>
<p>Minnesota Birth Center (MBC) <i>(Regional/local single specialty practice)</i> A Single Bundled Payment for Comprehensive Low-Risk Maternity and Newborn Care Provided by Independent Midwife Led Birth Center Practices that Are Clinically Integrated with Physician and Hospital Services N/A - Withdrawn</p>	<p>Clinical Focus: Maternity/newborn care Providers: Certified nurse midwives (CNMs), registered nurses (RNs), and licensed practical nurses (LPNs) Setting: Birth centers Patient Population: Pregnant people and newborns</p>	<p>Utilization measure(s): Comparison of cesarean section rates (NTSV cesarean rate); rate of ultrasound imaging Quality measure(s): Mention use of “outcome-level maternity quality measures” but do not provide detail on what these measures are Spending measure(s): N/A Patient experience measure(s): Printed postpartum survey (but it is not cataloged)</p>	<p>How payment is adjusted for performance: There are not formal performance measures provided in this proposal, and measures are not tied to payment. MBC asks for PTAC assistance in design of the payment methodology but suggests an upfront partial payment at 20 weeks gestation, then a final retrospective bundled payment shortly after completion of the episode; providers could also take on additional risk by assuming responsibility for some multiple of the agreed upon bundled price. Requirements: Not specified</p>	<p>Measures used for implementation: N/A Measures used for monitoring: N/A How achievement is measured: N/A How improvement is measured: N/A</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>Attribution: N/A</p> <p>Volume: The model is designed for cohorts of 250-300 low-risk pregnant mothers per year led by five-member certified nurse midwife (CNM) teams that coordinate with obstetrics, pediatric, and neonatal physicians.</p> <p>Risk stratification or adjustment: Adjustments to cost could be provided via an outlier payment if the costs for the mother or her baby exceed a certain amount.</p> <p>Benchmarking: MBC asks for PTAC assistance in determining the appropriate amount of the bundled payment; authors mention using historical rates as a benchmark; however, Medicaid payments in Minnesota (and many other states) have historically been low.</p>	
<p>Mercy Accountable Care Organization (Mercy ACO)</p> <p><i>(Regional/local multispecialty practice or health system)</i></p> <p>Annual Wellness Visit Billing at Rural Health Clinics (RHCs)</p> <p>The criteria for PFPs established by the Secretary are not applicable to this proposal, 12/18/2017</p>	<p>Clinical Focus: Primary/preventive care</p> <p>Providers: Rural health clinic (RHC) providers</p> <p>Setting: Outpatient</p> <p>Patient Population: Medicare beneficiaries</p>	<p>Utilization measure(s): N/A</p> <p>Quality measure(s): Colorectal cancer screening; tobacco screening & cessation; depression screening and follow-up; screening for fall risk; breast cancer screening; annual wellness visit completion rate</p> <p>Spending measure(s): Total cost of care (PMPM expense)</p> <p>Patient experience measure(s): Patient Satisfaction standards</p>	<p>How payment is adjusted for performance: While this proposal was submitted by an ACO that is part of the MIPS APM, this proposal does not propose a new model. It solely addresses the ability to provide annual wellness visits in rural health clinics (RHCs); there is a 10% withhold in provider compensation if less than 40% of attributed Medicare patients complete an annual wellness visit or if patient satisfaction standards are not met.</p>	<p>Measures used for implementation: Annual wellness visit completion rate, Patient Satisfaction standards</p> <p>Measures used for monitoring: N/A</p> <p>How achievement is measured: Annual wellness visit completion will be tracked through the CMS Group Reporting Option web interface. At least 40% of patients must complete an annual wellness visit to avoid the 10% payment withhold. The proposal does not specify how</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
			<p>Requirements: Providers must complete their documentation within a week, and providers in a Track 3 MSSP must achieve savings; providers in a Track 1 MSSP must not incur a penalty.</p> <p>Attribution: N/A</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: N/A</p> <p>Benchmarking: Not specified</p>	<p>patient satisfaction standards are measured.</p> <p>How improvement is measured: Results from previous years will be used as a baseline; Mercy ACO also created a disease registry to monitor clinical quality metrics and resource utilization.</p>
<p>Seha Medical and Wound Care (Seha) <i>(Individual provider)</i></p> <p>Bundled Payment for All Inclusive Outpatient Wound Care Services in Non Hospital Based Setting</p> <p>Not recommended, 5/17/2019</p>	<p>Clinical Focus: Acute and/or chronic wound care</p> <p>Providers: Independent office-based wound care physician or provider and home care providers</p> <p>Setting: Outpatient</p> <p>Patient Population: Patients with an acute or chronic wound</p>	<p>Utilization measure(s): The proposal suggests number of visits to heal different wounds like diabetic and venous leg ulcers; number of prescriptions filled for proper offloading devices and footwear (for example diabetic footwear), prescriptions for compression garments for patients with venous ulcers</p> <p>Quality measure(s): The proposal suggests measurement of a patient's improvement in quality of life; improvement in pain scale/control; physical and psychological improvements; blood monitoring of A1c; signing of patient contracts to encourage adherence to the plan of care</p> <p>Spending measure(s): Total cost of care per episode</p> <p>Patient experience measure(s): N/A</p>	<p>How payment is adjusted for performance: Medicare will pay a \$400 per visit bundled payment that will include all services provided to an office-based or independent wound care provider or clinic. The proposal mentions that the model will create incentives for best outcomes but does not specify details.</p> <p>Requirements: Not specified</p> <p>Attribution: Not specified</p> <p>Volume: N/A</p> <p>Risk stratification or adjustment: N/A</p> <p>Benchmarking: Can compare to national averages if Medicare makes available its data on overall wound care expenditures</p>	<p>Measures used for implementation: None, not tied to payment</p> <p>Measures used for monitoring: Utilization and quality measures listed</p> <p>How achievement is measured: Not specified</p> <p>How improvement is measured: N/A</p>
<p>Dr. Sobel (Sobel) <i>(Individual)</i></p>	<p>Clinical Focus: Broad/not specified</p> <p>Providers: Regional Referral Centers (specialists)</p>	<p>Utilization measure(s): Avoidable care (e.g., specialist utilization, admissions, tests)</p> <p>Quality measure(s): Not specified</p> <p>Spending measure(s): Cost of care</p>	<p>How payment is adjusted for performance: Quality, outcomes, and patient satisfaction affect fee schedule and future participation</p>	<p>Measures used for implementation: Quality, outcomes, and patient satisfaction</p>

Proposal Name	Clinical Focus, Providers, Setting, Patient Population	Components Relevant to Performance Measurement	Technical Issues Related to Performance Measurement	How Measurement is Used to Determine Success
Remote specialists and experts on demand improving care and saving costs (Revised version) N/A - Withdrawn	Setting: Not specified Patient Population: Not specified	Patient experience measure(s): Patient satisfaction with care	in the remote specialists' and experts' program Requirements: N/A Attribution: N/A Volume: N/A Risk stratification or adjustment: N/A Benchmarking: N/A	Measures used for monitoring: Admissions How achievement is measured: Not specified How improvement is measured: Not specified

Appendix H. Summary of Responses from Discussions with Subject Matter Experts

As part of gathering information for this environmental scan, NORC conducted discussions with three subject matter experts (SMEs) from measure development and endorsement organizations. The purpose of these discussions was to better understand how performance measures are changing in response to the changing health care landscape, notably the shift to population-based total cost of care (PB-TCOC) models. Topics addressed during the discussions included types of measures in PB-TCOC models, implementing measures in PB-TCOC models, and linking measures to payment in PB-TCOC models.

Subject Matter Experts

1. Eric Schneider – National Committee for Quality Assurance (NCQA)
2. Jeffrey Geppert – Battelle / Partnership for Quality Measurement (PQM)
3. Dana Gelb Safran – National Quality Forum (NQF)

SMEs act as discussants representing their own expertise and opinions, not those of their respective organizations or affiliations.

Below is a summary of the subject matter experts' responses in the aggregate. Responses have been intentionally anonymized.

Challenges Related to Developing and Implementing Performance Measures

- Providers face substantial burden when collecting and reporting performance data. The time required to collect and report data should be decreased as much as possible from clinicians' workstreams.
- Providers tend to rely on using claims data for reporting quality measures because it is generally lower burden than using data from the electronic health record (EHR) or registries.
- Providers are burdened by the use of different measures and definitions for a given area of measurement. This could be alleviated if payers use the same measures and measure sets across models.
- PB-TCOC measures should align with the measures included in CMS' Universal Foundation.
- There are three areas where TCOC quality measure development is currently challenged. First, EHRs and tools used by behavioral health providers do not have the same capabilities as the EHRs and tools used by medical care providers. Second, community-based organizations (CBOs) addressing social needs do not have usable or standardized digital platforms. Third, digitizing, reporting, and exchanging public health data is behind the rest of the health care system.
- Because developing measures and obtaining endorsement can be a lengthy process, there is a need to accelerate the development, testing, and adoption process of measures that fill high priority gaps, such as outcomes of depression and anxiety among children and adolescents as well as maternal health outcomes. It is important to incorporate stakeholders' needs, insights, and expertise throughout the measure development process. Considering clinical objections and operational barriers to a measure early in the development process can help to ensure developed measures are ready for broad adoption and use.

- There is a lack of guidance on how to implement measures, particularly measures that do not solely rely on data extracted from claims. To support the feasibility of developed measures, measure developers should explicitly describe what is required to implement the measures (e.g., identify the technology, changes to workflow, changes to processes, and special skill sets required to implement the measures). Components from the electronic clinical quality measure (eCQM) feasibility scorecard should be applied to all measure sets.
- There is also a lack of standardization in measure implementation. For example, the way EHRs were implemented in the Health Information Technology for Economic and Clinical Health (HITECH) Act has led to measure implementation challenges. The HITECH Act prioritized putting paper records into digital form over focusing on the electronic exchangeability of key data elements. Interoperability standards for measure implementation should be reflected in the United States Core Data for Interoperability (USCDI) and the Creating Access to Real-time Information Now through Consumer-Directed Exchange (CARIN) for Blue Button.

Opportunities for Alternative Payment Models (APMs) and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures

- Artificial intelligence (AI) methods, including natural language processing (NLP), can be used to reduce provider burden and support measure development. AI methods can reduce burden associated with reporting quality measures by extracting data from narrative fields without requiring any changes to the clinical workflow. AI tools can also support measure development by allowing measure developers to incorporate data from multiple sources to create clinical constructs that are not immediately apparent in discrete data.
- Interoperability and Fast Healthcare Interoperability Resources (FHIR) standards can reduce provider burden when extracting data from the clinical record and reporting the data for quality measurement purposes.
- CMS and other quality measure developers are creating a health data ecosystem that is broader than claims data and EHR data. CMS' eCQMs were designed to use EHRs as the primary data source.
- The integration of patient-reported outcome measures (PROMs) into the EHR will reduce provider burden and allow providers to more easily use patient-reported measures in population-based models. However, there are feasibility-related challenges with integrating patient-reported outcomes into EHRs.
- The high cost of quality reporting is, in part, due to a lack of standardization. A lack of standardization necessitates manual labor, audits, and other workarounds. Although overcoming the lack of standardization in collecting and reporting data will increase costs exponentially, costs will decrease once the system achieves standardization. Cost savings will occur when data can be retrieved through open Application Programming Interface (API) mechanisms.
- There is a need for transparency in the measurement of clinical and financial risk of populations. There are privacy concerns in how patient data will be used. The public will need to trust that the data will be used to improve health outcomes and reduce costs.

Linking Performance Measures with Payment

- Attributes of PB-TCOC measures that would be most appropriate to link to payment include the following: measures that are outcome-oriented, reliable, valid, clinically important, patient-centered, parsimonious, digital, explicit about what quality means, low burden, and high value.
- Best practices for linking performance to payment include the following:
 - Align the set of measures used in a program for every provider in that program.
 - Ensure every provider faces the same benchmark, or set of benchmarks, as every other provider in their “cohort” (i.e., same start date for the provider in the program).
 - Set benchmarks in absolute terms, not relative terms, so that any provider who can achieve the benchmark is rewarded.
 - For every measure, use a continuum of performance targets, such that providers are rewarded for both performance and improvement. The farther along a provider gets on the continuum, the more reward there should be.
 - Set the measures and benchmarks for a three- to five-year period of the contract to allow providers to plan their improvement journey.
- PB-TCOC models are generally associated with a higher performance-based earning potential than fee-for-service (FFS) models. In FFS models, there is typically a small number of measures and the earning potential for performance on the measures is generally small. For PB-TCOC models, organizations are typically held accountable for a broader set of measures, reflecting provider accountability for the full end-to-end continuum of population health.
- Incentivizing high quality care cannot take a one-size-fits-all approach. Incentivizing quality care should be an organization-specific exercise because labor markets are different across the country. APMs should be tailored to the way organizations manage their resources so that they can take financial risk for populations with different needs.
- Value-based dollars should be focused on helping providers who are systematically and persistently performing poorly rather than continuing to pay providers for what they already do well or to pay providers for performance that does not last or translate.
- Benefits can emerge from separating clinical and financial risk from quality performance. For example, the Dutch system of managed competition reduces the incentive to compete on population health risk selection and increases the incentive to deliver higher quality care. There are financial rewards for the quality measures and financial reallocations for risk-based differences between populations.
- Risk selection is one unintended consequence of linking performance to payment. Risk selection allows providers to optimize the enrolled population under the TCOC model to reduce costs without having to change the clinical management of that population (i.e., “cherry picking”). Risk selection is an important factor to consider when designing quality measures for use in a TCOC context, particularly health outcome-related quality measures. Measures will need to be designed to detect risk selection and provide mitigation strategies to protect against risk selection.

Best Practices for Measuring Performance

- SMEs offered different perspectives on the importance of structure measures in PB-TCOC models. Whereas one SME suggested structure measures are emerging in PB-TCOC models and can help providers be intentional about what information is collected and reported, another SME suggested structure measures are not ideal for PB-TCOC models because there is rarely

strong evidence showing a link between structure and outcome. Measuring structure does not guarantee good outcomes for patients.

- Process measures are necessary in PB-TCOC models. Specifically, process measures are required to understand whether appropriate care is delivered.
- Health outcomes and access measures should not be relied upon to ensure care is delivered equitably.
- Population-based models should prioritize the use of outcome measures, including both patient-reported information about functional status and well-being and clinically sourced data on biometric information. Creating accountability for outcomes transforms health care by allowing providers to decide which care processes will achieve the best outcomes for their patients, as opposed to being prescriptive on process.
- Patient-reported data, including patient-reported outcome-based performance measures (PRO-PMs), should be used in PB-TCOC models. Understanding whether the patients' goals were met can be more informative than understanding whether the patients were satisfied with their care. In addition, a major backstop to stinting on care in a TCOC model is through the use of a patient experience set of measures. Patient-reported data can be particularly useful for risk assessment.
- Access measures should also be used in PB-TCOC models. Organizations with a strong TCOC financial incentive can reduce cost by reducing access, contradicting the purpose of the incentive.

Opportunities to Modify Existing Measures and Areas Where New Performance Measures May be Needed

- Five broad clinical areas that have few endorsed outcome measures include the following: oncology, obstetrics, mental health, musculoskeletal care, and cardiovascular procedures. These five clinical areas represent more than 50 percent of medical spending.
- There are gaps in the available outcome measures that use patient-reported health information for population-based models. Specifically, there are currently few measures that use longitudinal patient-specific information to assess outcomes, or PRO-PMs. Development of PRO-PMs will enhance the ability of PB-TCOC models to evaluate the health outcomes being achieved by these types of models.
- The development of well thought out structural measures is another area in which measure development for PB-TCOC models could be improved. Meaningfully measuring outcomes and costs at the clinician-level or group-level can be difficult because of small sample sizes.
- There are gaps in measures on (1) wellness and prevention and (2) recognition and management (e.g., diagnosis or identifying symptoms early before they manifest into disease). Early diagnosis can have a major impact on the total cost of care.
- A lot of the cost savings that could be achieved reside in a relatively small portion of the population. This population tends to have complex care needs, chronic health conditions, and behavioral health issues. An investment should be made in digital quality measure development to focus on high-risk populations.

Appendix I. Areas for Future Exploration and Research

The following are areas where additional information is needed. Please note that some of the items listed below may be better addressed through stakeholder input (e.g., through a Request for Input (RFI), subject matter expert (SME) panel discussions or listening sessions) or another research approach. They are captured here for further exploration.

- What are the main goals of performance measurement for TCOC organizations (for example, to drive change through financial incentives, to provide actionable information for providers, or to inform beneficiary choices)?
- How should the goals for performance measurement and the measures for PB-TCOC models differ from the goals and performance measures for FFS payment systems?
- If PB-TCOC models want to directly measure systems change, what should be the mixture of outcome, patient experience, and process measures?
- If PB-TCOC models want to directly measure how organizations provide care, what should be the mixture of outcome, patient experience, and process measures?
- To what extent can patient/caregiver experience measures accurately reflect the provision of patient-centered, coordinated care relative to direct measures of those processes?
- Are certain PB-TCOC models' performance measures better suited for specific provider types?
 - If so, which measures are best matched with which provider types?
- Should some of the current process-related quality measures transition to related quality outcome measures over time?
 - If so, what are best practices related to developing and implementing those related quality outcome measures?
- Should performance-based financial incentives (e.g., performance-based payments) be determined using individual measures or a set of performance measures?
- In which contexts does it make sense to have organization-wide versus specialty-specific or setting-specific performance measures?
- What approaches to design of performance-based payment incentives for PFPMs are most likely to facilitate improvement while minimizing unintended consequences?
- Should assessment of achievement and improvement differ by type of measure and, if so, how?
- Should the approach to performance-based payment for quality measures vary based on type of measure/measure selection (e.g., evidence-based process measures versus outcome measures or patient-reported outcome measures)?
- Should payment incentives be phased in for newer measures or certain types of measures? If so, how?
- What are best practices for PB-TCOC models to progress toward incentives that increase participants' financial accountability (e.g., transitioning from pay-for-reporting to pay-for-performance)?
- What role can balancing measures (e.g., measures intended to reduce harm) have in PB-TCOC models?
- To what extent should performance measures be standardized across payers and models?

- To what extent can current measures be modified to meet the goals of PB-TCOC models, as opposed to creating a new measure?
- Are key performance measures missing from the current set of performance measures that should be developed?
- Could existing performance measures be modified to address these gaps?

Appendix J. Annotated Bibliography

Abt Associates. *Evaluation of the Oncology Care Model*. Abt Associates; 2021.

<https://www.cms.gov/priorities/innovation/data-and-reports/2021/ocm-ar4-eval-part-persp-report>

Subtopic(s): Key Highlights; Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To evaluate participant perspectives of the Oncology Care Model (OCM), including participants' reasons for participating in the model and what they learned through the first four years of the model.

Main Findings: OCM led to improved patient-centered care through redesigned care delivery. Participants reported improvements in patient-centered care such as documenting Care Plans, conducting depression screening, and improving patient navigation. Challenges reported by participants included having difficulty with analyzing the claims data.

Strengths/Limitations: The large sample size and efforts to visit practices of different sizes, geography, ownership, and patient population characteristics helped ensure that the reported results generalize to all OCM Model participants.

Generalizability to Medicare Population: Strong; the report evaluates the OCM, an APM which provides enhanced services for eligible Medicare beneficiaries.

Methods: In-person case studies were conducted with 47 practices participating in OCM during the first four model years. Semi-structured protocols were used to interview over 900 people working in OCM practices. Qualitative coding techniques were used to identify key themes emerging from the interviews.

Abt Associates. *Evaluation of the Oncology Care Model: Executive Summary*. Abt Associates; 2023:10

<https://www.cms.gov/priorities/innovation/data-and-reports/2023/ocm-evaluation-pp1-9-exec-sum>

Subtopic(s): Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To evaluate the Oncology Care Model (OCM).

Main Findings: The study found that OCM reduced total episode payments by 1.7 percent, which was driven by savings in higher-risk episodes, given the lack of change in payments for lower-risk episodes. Payment reductions were largely attributable to reductions in spending on non-chemotherapy drug payments, which accounted for approximately half of the overall relative reductions generated by OCM. The OCM resulted in cumulative net losses, but greater payment reductions in more recent performance periods are beginning to generate savings sufficient to cover Monthly Enhanced Oncology Services (MEOS) payments.

Strengths/Limitations: The difference-in-differences design allows for more robust assertions with respect to program causal effects.

Generalizability to Medicare Population: Strong; the report evaluates the OCM, an APM which provides enhanced services for eligible Medicare beneficiaries.

Methods: The study uses a difference-in-differences evaluation approach to measure any changes over the course of the model in the comparison group or the OCM group and incorporated primary data collection, including patient surveys and case study interviews, to evaluate OCM's impact on quality of care, patient satisfaction, and perceptions of clinical changes and quality that resulted from the model.

Afendulis CC, Fendrick AM, Song Z, et al. The Impact of Global Budgets on Pharmaceutical Spending and Utilization. *Inquiry*. 2014;51:0046958014558716. doi:10.1177/0046958014558716

Subtopic(s): Consumer Assessment of Healthcare Providers and Systems Survey

Type of Source: Journal article

Objective: To evaluate the Alternative Care Quality (ACQ) payment system.

Main Findings: ACQ participants did not experience any change regarding beneficiaries' use of drugs, and they had a lower probability of positive utilization compared with control groups.

Strengths/Limitations: The study population was young and included only BCBS, Health Maintenance Organization (HMO), or Point of Service (POS) plan enrollees, which limits its generalizability to other populations. Additionally, the analyses covered only the first two years of the AQC, which set to save money over a five-year period. The AQC targets were set on the basis of actuarial projections to save money over the course of the five-year contract.

Generalizability to Medicare Population: Moderate; the study population consisted of young beneficiaries.

Methods: Difference-in-differences model comparing participants 2006 through 2008 to 2009 through 2010. Participants are all enrollees with primary care providers physician organizations that started in the program in 2009.

Ahluwalia SC, Damberg CL, Haas A, Shekelle PG. How are medical groups identified as high-performing? The effect of different approaches to classification of performance. *BMC Health Serv Res*. 2019;19(1):500. doi:10.1186/s12913-019-4293-9

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To examine how the differing classification systems influence which providers are deemed high-performing.

Main Findings: No medical groups were identified as high-performing in the top 25 percent or 35 percent across all four performance domains (quality, total cost of care, access, and patient experience) when evaluating them using a relative value threshold. One medical group was identified in the top 40 percent, and a second was identified in the top 50 percent. Two medical groups were identified as high-performing across all four domains using an absolute value threshold.

Strengths/Limitations: Datasets have only certain measures and an inability to evaluate how differing population mix may influence provider's performance evaluation.

Generalizability to Medicare Population: Moderate; this article is generalizable to providers who care for Medicare beneficiaries.

Methods: Analysis of medical groups' performance data from the 2014 Minnesota Community Measurement (MNCM) Health Care Quality Report.

Anderson AC, O'Rourke E, Chin MH, Ponce NA, Bernheim SM, Burstin H. Promoting Health Equity And Eliminating Disparities Through Performance Measurement And Payment. *Health Aff (Millwood)*. 2018;37(3):371-377. doi:10.1377/hlthaff.2017.1301

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To apply National Quality Forum's roadmap on promoting equity care through measurement toward reducing racial disparities in hypertension among African Americans.

Main Findings: Tying stratified performance measures to payment incentives can lead to health care settings' leadership investing in the resources needed to achieve equitable outcomes.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; findings from this article are generalizable to organizations that work with Medicare beneficiaries.

Methods: Apply National Quality Forum's roadmap to organizations caring for African American patients with hypertension.

Arbor Research. *Evaluation of the Home Health Value-Based Purchasing (HHVBP) Model*. Arbor Research; 2023:252. <https://www.cms.gov/priorities/innovation/data-and-reports/2023/hhvbp-sixth-ann-rpt-app>

Subtopic(s): Key Highlights; Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To evaluate the HHVBP Model to better understand how the shift in financial incentives may influence agency behavior and, in turn, aspects of home health care.

Main Findings: The model reduced Medicare spending for Parts A and B services (predominantly inpatient and SNF services), made modest declines in some aspects of utilization, and made small improvements in most quality measures. The model's impacts on lower unplanned hospitalizations and greater improvements in functioning were not observed among Medicaid patients.

Strengths/Limitations: The study employed a difference-in-differences design that allows the study to make more robust causal claims about program efficacy. One limitation to study findings, however, is the potential variation in how Medicaid status is coded across states.

Generalizability to Medicare Population: Strong; this is an evaluation of a Medicare value-based purchasing model.

Methods: Multivariate linear regression in a difference-in-differences framework.

Arbor Research. *HHVBP Evaluation Final Report*. Arbor Research; 2023.

<https://www.cms.gov/priorities/innovation/data-and-reports/2023/hhvbp-seventh-ann-rpt>

Subtopic(s): Key Highlights; Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To evaluate the Home Health Value-Based Purchasing (HHVBP) Model and understand how the financial incentives influenced agency behavior, health outcomes, and Medicare spending.

Main Findings: The Total Performance Score (TPS) was higher among agencies in HHVBP states compared with agencies in non-HHVBP states. The HHVBP Model had only a modest impact on unplanned hospitalizations, an increase in outpatient emergency department use, improvement in nearly all Outcome and Assessment Information Set (OASIS)-based quality measures, and a modest impact on patient experience. The cumulative financial impact of the HHVBP Model translated to \$1.38 billion in Medicare savings.

Strengths/Limitations: The final report assessed the cumulative effect of the entire HHVBP Model from 2016 to 2021.

Generalizability to Medicare Population: Strong; this report is an evaluation of a Medicare value-based purchasing model.

Methods: A mixed methods framework, utilizing both quantitative and qualitative data, was used to understand the impact of the HHVBP Model.

Assistant Secretary for Planning and Evaluation. Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. *ASPE*; 2020. Accessed November 29, 2023. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>

Subtopic(s): Key Highlights; Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To explore trends with providers who are addressing beneficiaries' social risk factors through connecting them to social services and community-based organizations.

Main Findings: Beneficiary social risk information is not systematically collected across health care settings. Beneficiaries' dual enrollment status is a strong predictor of poor outcomes on some quality and resource use measures in Medicare's VBP programs. Functional status is also a strong predictor of poor outcomes on some measures but is not always included in measure risk adjustment. It remains unclear which interventions are effective, replicable, and scalable for addressing social risk.

Strengths/Limitations: The report's analysis could not use potentially valuable data sources given their limited social and functional risk factors (such as CAHPS, Health and Retirement Study [HRS], and National Health and Aging Trends Study [NHATS]). The report's use of Medicare Current Beneficiary Survey data limits its ability to have an accurate provider-level analyses because of its small sample size. These limitations are to some extent balanced out by including analyses from American Community Survey, which has a larger sample size.

Generalizability to Medicare Population: Strong; this report focuses on Medicare value-based programs.

Methods: Raw performance on claims-based measures was first calculated for beneficiaries having the risk factor of interest versus beneficiaries not having that risk factor (e.g., living alone versus not living alone, frail versus non-frail). Statistical models were then used to adjust the raw difference in performance for medical risk factors currently included in the measure's risk adjustment, as well as social risk factors, such as dual enrollment, not currently included. Additional models then examined the impact of new measures of social and medical risk.

Austin E, LeRouge C, Hartzler AL, Segal C, Lavallee DC. Capturing the patient voice: Implementing patient-reported outcomes across the health system. *Qual Life Res.* 2020;29(2):347-355. doi:10.1007/s11136-019-02320-8

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To understand how to improve the incorporation of patient voice into health care.

Main Findings: Approximately half of patient-reported outcomes capture shared clinical domains of depression, anxiety, pain, and quality of life. Establishing a system-wide measurement strategy for patient voice can assist health systems with addressing inefficiencies.

Strengths/Limitations: The study's focus on the needs of clinical teams and health systems does not reflect the patient perspective. The study's snowball sampling approach limits the ability to understand the breadth of outreach. It is possible not all patient-reported outcome measures were captured in the analysis.

Generalizability to Medicare Population: Moderate; findings from this article are generalizable to the Medicare population as Medicare programs also have patient-reported outcomes.

Methods: Participatory action research approach which involved engaging stakeholders through electronic questionnaires and in-person interviews.

Basch E, Spertus J, Adams Dudley R, et al. Methods for Developing Patient-Reported Outcome-Based Performance Measures (PRO-PMs). *Value in Health*. 2015;18(4):493-504. doi:10.1016/j.jval.2015.02.018

Subtopic(s): Key Highlights; Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To recommend methods for assessing quality of care using patient-reported outcomes measures.

Main Findings: The best practices for using patient-reported outcomes measures include 1) provide a rationale for measuring the outcome and for using a PRO-PM; 2) describe the context of use; 3) select a measure that is meaningful to patients with adequate psychometric properties; 4) provide evidence of the measure's sensitivity to differences in care; 5) address missing data and risk adjustment; and 6) provide a framework for implementation, interpretation, dissemination, and continuous refinement.

Strengths/Limitations: No mention of inclusion of patients in the Technical Expert Panel, which may have limited the authors' understanding of patient-reported outcomes' best practices.

Generalizability to Medicare Population: Moderate; this article applies to the Medicare population because Medicare programs have patient-reported outcomes.

Methods: A Technical Expert Panel comprised of experts in performance measurement, PROs, clinical research, health services research, as well as clinical practitioners and patient representatives and an environmental scan and structured literature review.

Battista P, Salvatore C, Berlingeri M, Cerasa A, Castiglioni I. Artificial intelligence and neuropsychological measures: The case of Alzheimer's disease. *Neurosci Biobehav Rev*. 2020;114:211-228. doi:10.1016/j.neubiorev.2020.04.026

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To assess the use of machine learning (ML) and neuropsychological measures for classifying mild cognitive impairment (MCI) among patients and the conversion to Alzheimer's disease (AD).

Main Findings: Neuropsychological measures can lead to a successful automatic classification of prodromal AD phenotypes regardless of the employment of different ML algorithms. ML algorithms are able to extract relevant categories of neuropsychological tests that maximize the classification accuracy.

Strengths/Limitations: Using neuropsychological measures to diagnose AD has limited reliability, especially in the early stages of the disease. The papers reviewed for the meta-analysis have high heterogeneity, which limits the findings' generalizability.

Generalizability to Medicare Population: Moderate; findings are generalizable as some Medicare patients have Alzheimer's disease.

Methods: Meta-analysis of existing papers.

Beauvis B, Whitaker Z, Kim F, Anderson B. Is the Hospital Value-Based Purchasing Program associated with reduced hospital readmissions? *Journal of Multidisciplinary Healthcare*. 2022;15:1089-1099. doi: 10.2147/JMDH.S358733

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To determine if hospitals eligible for the Hospital Value-Based Purchasing (HVBP) program are associated with a reduction in a core set of Hospital Readmission Reduction Program (HRRP) readmission rates compared with hospitals that are ineligible for the HVBP program.

Main Findings: Controlling for organizational and environmental factors, there was a positive association between HVBP participation and every measure of hospital readmission in the HRRP program. Results could indicate that the HVBP and HRRP programs may not improve patient care and lower health care costs.

Strengths/Limitations: One limitation of this study is that the authors did not control for every possible influence on readmission rates, including patients' conditions at the point of admission and discharge. The analyses also did not consider patient demographics or socioeconomic status.

Generalizability to Medicare Population: Strong; the article examined readmission rates of hospitals eligible and ineligible for the HVBP program, which is a Medicare-based program designed to improve quality of care.

Methods: Short-term acute care hospital data from multiple public data sources (e.g., the Healthcare Cost Report Information System, Hospital Compare, the United States Census Bureau, CMS Hospital Compare, CMS hospital HVBP program, CMS HRRP) were evaluated to understand the association between readmission rates and eligibility for the HVBP program.

Benabbas R, Shan G, Akindutire O, Mehta N, Sinert R. The Effect of Pay-for-Performance Compensation Model Implementation on Vaccination Rate: A Systematic Review. *Quality Management in Healthcare*. 2019;28(3):155. doi:10.1097/QMH.0000000000000219

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To evaluate the impact of financial incentives on vaccination rate.

Main Findings: Implementing financial incentives has a statistically significant impact on vaccination rate.

Strengths/Limitations: Most of the studies included were for pediatric patients. The variation among studies included prevents the authors from correlating incentive amount with the program's success.

Generalizability to Medicare Population: Moderate; Medicare patients receive certain vaccinations; however, the article's inclusion of studies predominantly evaluating pediatric patients and no specific Medicare program limits the generalizability of the findings to the Medicare population.

Methods: This study used a meta-analytic approach.

Berdahl CT, Easterlin MC, Ryan G, Needleman J, Nuckols TK. Primary Care Physicians in the Merit-Based Incentive Payment System (MIPS): A Qualitative Investigation of Participants' Experiences, Self-Reported Practice Changes, and Suggestions for Program Administrators. *J GEN INTERN MED*. 2019;34(10):2275-2281. doi:10.1007/s11606-019-05207-z

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To describe primary care provider (PCP) perspectives and self-reported practice changes related to MIPS quality measurement and disseminate PCP recommendations for MIPS improvement.

Main Findings: PCPs mentioned that one positive aspect of MIPS is its creation of practice-level systems for quality improvement. They reported its negative aspects include administrative burden and the possibility of penalization of vulnerable populations. PCPs recommend simplifying the program to reduce administrative burden, implementing protections for practices serving vulnerable populations to prevent penalization from occurring, and improving communication between program administrators and PCPs.

Strengths/Limitations: The study's small sample size and focus on PCPs may limit the study's generalizability to the broader MIPS participant population. There may be investigator bias with participants. The respondents' input may have been directed at quality measurement and pay-for-performance programs as a whole, and not specifically for MIPS.

Generalizability to Medicare Population: Strong; MIPS is for Medicare Part B FFS reimbursement.

Methods: Semi-structured interviews with PCPs reporting under MIPS.

Berenson RA, Shartz A, Pham HH. Beyond demonstrations: Implementing a primary care hybrid payment model in Medicare. *Health Affairs Scholar*. 2023;1(2):qxad024. doi:10.1093/haschl/qxad024

Subtopic(s): Risk Adjustment Methods Used for PB-TCOC Models; Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To provide approaches to implement hybrid payments (mix of fee-for-service and population-based payments) in Medicare programs.

Main Findings: Incremental changes to the Medicare Physician Fee Schedule do not achieve primary care transformation or address shortages in the primary care workforce. A quality measurement approach that involves measures used internally by practices instead of publicly or for payment may support clinicians better and motivate them further to take responsibility for care, as opposed to the existing approaches implemented. The authors suggest using primary care-specific measures for implementing hybrid payment models. The MSSP program should adopt hybrid payments, and CMS should develop hybrid payments that focus on quality improvement through focusing on four to six meaningful outcomes concurrently.

Strengths/Limitations: The paper focused on Medicare programs, and it would be helpful to discuss approaches for other payers to implement hybrid payments.

Generalizability to Medicare Population: Strong; the article discusses payment models for Medicare programs.

Methods: A meta-analytic approach was used.

Berenson RA, Upadhyay D, Delbanco SF, Murray R. Payment Methods: How They Work. *Urban Institute*; 2016. Accessed December 1, 2023. <https://www.urban.org/research/publication/payment-methods-how-they-work>

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To explore how established and proposed payment methods and benefit design options work on their own and together.

Main Findings: Discusses health professionals' fee schedules, primary care capitation, per diem payment to hospitals for inpatient stays, diagnosis-related groups-based payment to hospitals for inpatient stays, global budgets for hospitals, bundled episode payment, global capitation for an organization, shared savings, and pay-for-performance approaches to paying for health care.

Strengths/Limitations: In-depth analysis of the different payment approaches.

Generalizability to Medicare Population: Moderate; the report includes payment approaches for differing Medicare programs and services.

Methods: Technical expert panel of payment experts and a meta-analysis of peer-reviewed literature.

Bond AM, Schpero WL, Casalino LP, Zhang M, Khullar D. Association Between Individual Primary Care Physician Merit-based Incentive Payment System Score and Measures of Process and Patient Outcomes. *JAMA*. 2022;328(21):2136-2146. doi:10.1001/jama.2022.20619

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To analyze the link between the Medicare Merit-based Incentive Payment System (MIPS) scores and primary care physicians' performance on clinical and patient outcome measures.

Main Findings: MIPS scores are inconsistently associated with physicians' performance. Physicians with low MIPS scores had worse performance on diabetes and mammogram screenings and better performance on flu vaccines and tobacco screenings compared with physicians with high MIPS scores. Additionally, physicians with low MIPS scores but higher outcomes cared for more high-risk patients compared with physicians with low MIPS scores and poor outcomes.

Strengths/Limitations: The study used measures from Medicare claims data and thus was somewhat limited in the outcomes it analyzed; the study's focus on a single year limits its generalizability and longitudinal evaluation of MIPS and provider performance; although the study controlled for patient and physician characteristics, confounding may still have occurred; and the outcomes measures analyzed may have differing importance to patients and clinicians that their weighting cannot account for.

Generalizability to Medicare Population: Strong; MIPS is tied to Medicare Part B payments and services.

Methods: Cross-sectional observational study of 80,000 primary care physicians using the 2019 Physician Compare data, 2019 Medicare Data on Provider Practice and Specialty file, and a sample of 2018-2019 Medicare FFS claims.

Burke LG, Orav EJ, Zheng J, Jha AK. Healthy Days at Home: A novel population-based outcome measure. *Healthcare*. 2020;8(1):100378. doi:10.1016/j.hjdsi.2019.100378

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To evaluate how the measure Healthy Days at Home (HDAH) varies across traditional Medicare beneficiary populations and which characteristics have the greatest association with HDAH, how HDAH varies across health care markets, and how market characteristics are associated with mean marked HDAH.

Main Findings: Beneficiaries age 65+ had an average of 347 HDAH; those 80+ had an average of 325 HDAH, while those with three+ chronic conditions had an average of 334 HDAH. Beneficiaries' mortality, home health, SNF utilization, and inpatient care rates had the largest reduction in HDAH.

Strengths/Limitations: HDAH and its components are associated with beneficiaries' socioeconomic status, and the measure may have limitations with severity-adjustment using claims data.

Generalizability to Medicare Population: Strong; findings were generalizable as the study focused on traditional Medicare beneficiaries.

Methods: Linear regression of traditional Medicare beneficiaries with HDAH as the outcome and marked fixed effects, as well as beneficiary characteristics as covariates.

Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine*. 2022;37(2):359. doi:10.1007/s11606-021-06617-8

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To evaluate the impact of physician financial incentives on primary care quality.

Main Findings: None of the study's compensation incentives were associated with high- or low-value care.

Strengths/Limitations: Causation cannot be inferred from this study as the findings are only an association; the authors were unable to identify the impact of each individual incentive on quality of care; the measures are skewed toward those that can be commonly met in a visit instead of those that may be more impactful for clinically significant; and the measures were created based on availability in National Ambulatory Medicare Care Survey data and may be limited in their comprehensiveness of a treatment plan.

Generalizability to Medicare Population: Moderate; 31 percent of the patients analyzed had Medicare insurance.

Methods: Chi-square tests and logistic regression models of individual measures using National Ambulatory Medicare Care Survey dataset from 2012-2016.

Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To discuss the time physicians spent on administrative activities related to quality measures.

Main Findings: Per physician, physicians and staff spent an average of 785 hours annually and 15 hours weekly on administrative activities related to external quality measures. This time spent for each physician's quality measurement equates to an average cost of \$40,069 annually. 81 percent of physicians reported spending more time on external quality measures than three

years ago, and 27 percent believe that the current measures are representative of their quality of care.

Strengths/Limitations: The sample was limited to members in the Medical Group Management Association; response bias may have occurred given practices that have stronger negative feelings about quality measures are more likely to respond; the cost estimates did not include costs to practices of information technology or office space used for dealing with quality measures; and the estimates came from one representative from each practice which may result in inaccuracies with time and cost estimations.

Generalizability to Medicare Population: Moderate; some of the physicians in the study likely care for Medicare beneficiaries.

Methods: 1,000 randomly selected practices from the Medical Group Management Association database participated in a web-based survey.

Centers for Medicare & Medicaid Services. *2024 National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report*. Baltimore, MD: U.S. Department of Health and Human Services; 2024. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-AssessmentInstruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports>

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures; How Performance Measures are Linked with Payment in Other Programs; Best Practices for Measuring Performance

Type of Source: Report

Objective: To summarize the quality and efficiency impacts of endorsed measures used in 26 CMS quality and value-based incentive payment programs.

Main Findings: The analysis included 371 measures with three or more years of reliable data from 2016 to 2019. For select CMS programs, improvements in measure performance were associated with patient impacts and avoided costs. The improvements were greatest prior to the COVID-19 public health emergency (PHE). During the PHE, a large proportion of the measures had worse than expected performance. The majority of measures showed gaps health equity for historically disadvantaged groups. Findings from focus groups suggested a need to develop equity measures that address topics including bias in care delivery, cultural competency, health-related social needs, access, and health literacy.

Strengths/Limitations: Limited data availability during the COVID-19 PHE is one limitation of the analysis.

Generalizability to Medicare Population: Strong; the report summarized measures used in CMS programs.

Methods: The report used quantitative methods including regression models to understand the quality and efficiency impacts of the measures used in CMS programs. Focus groups were also conducted.

Centers for Medicare & Medicaid Services. *Digital Quality Strategic Roadmap*. CMS; 2022. Accessed November 3, 2023. https://ecqi.healthit.gov/sites/default/files/CMSdQMStrategicRoadmap_032822.pdf

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Report

Objective: To understand how to use EHRs to improve CMS' quality measurement systems.

Main Findings: Improvements in interoperability have enabled the development of measure calculation tools for digital quality measures that use EHR data. The roadmap provides guidelines for transitioning providers and settings to digital quality measurement.

Strengths/Limitations: Dependent on organizations having robust EHR systems.

Generalizability to Medicare Population: Moderate; the report discusses quality measurement for Medicare programs.

Methods: N/A

Centers for Medicare & Medicaid Services. *Synthesis of Evaluation Results across 21 Medicare Models, 2012-2020*. Centers for Medicare & Medicaid Services; 2022:49.

<https://www.cms.gov/priorities/innovation/data-and-reports/2022/wp-eval-synthesis-21models>

Subtopic(s): Key Highlights; Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To synthesize the results of Center for Medicare and Medicaid Innovation Center (CMMI) models occurring between 2012 and 2020 with at least two years of impact estimates.

Main Findings: Over half of the models analyzed experienced gross savings to Medicare. Among the models that offered financial incentives, six had net savings, six incurred net losses, and six had no discernable effects on net spending. Beneficiary or caregiver self-reported experience of care remained relatively constant among the majority of models. Additionally, mortality rates were largely unchanged for most models, with improvements in mortality in four models.

Models that focused on reducing acute or specialty care or that targeted specific populations such as terminal illness and lower extremity joint replacements were more likely to have gross savings and greater favorable impacts on utilization compared with models focused on primary care and population management.

Strengths/Limitations: Although the model performance periods did not occur during the same time period, the relatively small window for analysis (eight years) decreases the likelihood that the cross-model analysis was biased due to time-variant effects. For some of the later models, it is possible that an insufficient amount of time passed to truly observe model effects.

Generalizability to Medicare Population: Strong; the report focuses on Medicare model evaluations.

Methods: The study identified measures common across studies (e.g., spending, utilization, and quality of care) and then summarized these results.

Chen AH, Lee PV. California's Marketplace Innovations: Driving Health Plan Accountability For Quality And Equity. *Health Affairs Forefront*. 2022. doi: 10.1377/forefront.20220928.429170

Subtopic(s): Performance Measures Used in the Medicare Advantage Star Ratings Program

Type of Source: Journal article

Objective: To describe performance measurement and payment initiatives in California.

Main Findings: Although the quality improvement enterprise has grown, performance on clinical quality and equity outcome measures have not improved. The authors suggest there is an urgent need for purchasers to make improving quality and equity of health care a business imperative for health plans. In addition, the authors suggest that Covered California's quality initiatives can serve as a model to advance population health.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article discusses performance measurement and payment in the context of Medicare models and programs.

Methods: N/A

Chen MM, Rosenkrantz AB, Nicola GN, Silva, E, McGinty, G, Manchikanti, L, Hirsch, JA. The Qualified Clinical Data Registry: A pathway to success within MACRA. *AJNR Am J Neuroradiol.* 2017;38(7):1292-1296. doi:10.3174/ajnr.A5220

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Journal article

Objective: To describe the history of clinical data registries and quality reporting, as well as the role of Qualified Clinical Data Registries (QCDRs) in Merit-based Incentive Payment (MIPS) reporting.

Main Findings: Clinical data registries can play a role in expanding Medicare coverage for specific medical techniques and procedures. In collaboration with other societies, the American College of Radiology pioneered the use of QCDRs. QCDRs will help radiologists succeed in MIPS. Use of QCDRs is anticipated to increase over time, including the trend of linking reimbursement to registry reporting.

Strengths/Limitations: The article is not an original research article and instead provides a summary of the role of QCDRs in MIPS.

Generalizability to Medicare Population: Strong; the article focuses on the role of QCDRs on MIPS pro Kaufman gram reporting, a program within CMS' Quality Payment Program.

Methods: N/A

Chernew ME, McWilliams JM, Shah SA. The Case for Administrative Benchmarks (and Some Challenges). *NEJM Catalyst.* 2023;4(10). doi:10.1056/CAT.23.0194

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To provide an overview of different methods used for benchmarking.

Main Findings: Whereas empirical benchmarks are tied to actual spending or forecasts based on lagged spending in traditional Medicare, administrative benchmarks are set by taking a base rate and expanding it by an administrative factor reflecting goals, anticipated volume, and intensity growth. Using administrative benchmarks may allow one to avoid the shortcomings associated with using empirical benchmarks, such as the ratchet effect.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article discusses challenges with the way benchmarks are calculated for CMS programs and models, including Accountable Care Organizations (ACOs).

Methods: N/A

CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs). *Centers for Medicare & Medicaid Services*; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures; Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures

Type of Source: Report

Objective: To present a strategic framework for the development of quality measures to support MIPS and advanced APMs.

Main Findings: The Medicare Quality Payment Program's Measure Development Plan (MDP) emphasizes strategic alignment with existing CMS initiatives and stakeholder input. Furthermore, it underscores a commitment to reducing clinician burden through collaboration, harmonization of measures, and a person-centered approach to address gaps and support the transition to value-based payment models.

Strengths/Limitations: The new measures will resolve critical gaps in the measure portfolio; allow for alignment across federal, state, and private programs; and contribute to efficient data collection.

Generalizability to Medicare Population: Strong; the report draws on existing Medicare quality measurement and reporting programs.

Methods: The report follows MACRA guidelines to develop and refine quality measures by incorporating public input.

Committee on Core Metrics for Better Health at Lower Cost, Institute of Medicine. *Vital Signs: Core Metrics for Health and Health Care Progress*. (Blumenthal D, Malphrus E, McGinnis JM, eds.). National Academies Press (US); 2015. Accessed November 28, 2023.

<http://www.ncbi.nlm.nih.gov/books/NBK316120/>

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Report

Objective: To propose a concise set of core measures for assessing and monitoring the nation's health. The goal of these measures is to reduce the burden on clinicians, reduce costs, and improve health outcomes.

Main Findings: Fifteen core measures were put forth by the committee to assess and monitor the nation's health care: life expectancy, well-being, overweight and obesity, addictive behavior, unintended pregnancy, healthy communities, preventive services, care access, patient safety, evidence-based care, care match with patient goals, personal spending burden, population spending burden, individual engagement, and community engagement. The committee also provided 39 priority measures related to the core measures to offer insight into distinct areas of interest.

Strengths/Limitations: The core measure set provides an opportunity to promote nationwide collaboration and investment aimed at improving performance on targeted issues in health care. The publication notes challenges in the implementation of the measure set. Infrastructure challenges include a lack of seamless interfaces and interoperability in the measurement infrastructure, in addition to fragmentation across providers, payers, and patient populations. Limitations also exist in ensuring validity and standardization across diverse data sources. Data sharing is met with barriers including privacy and Health Insurance Portability and Accountability Act (HIPAA) constraints.

Generalizability to Medicare Population: Moderate; the paper explored measures designed for widespread application and noted their relevance beyond the Medicare population.

Methods: A committee comprised of 21 experts sourced feedback from organizational and individual stakeholders and deliberated during four meetings to reach consensus on the measure set.

Committee on Quality Measures for the Healthy People Leading Health Indicators; Board on Population Health and Health Practice; Institute of Medicine. *Toward Quality Measures for Population Health and the Leading Health Indicators*. Washington (DC): National Academies Press (US); 2013 Apr 9. 2, Criteria for Selecting Measures. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK202189/>

Subtopic(s): Background

Type of Source: Report

Objective: To identify quality measures for the 12 Leading Health Indicator (LHI) topics and 26 LHIs in Healthy People 2020 with the overarching goal of developing a national framework for quality that incorporates the multisectoral health system.

Main Findings: The committee proposed the adoption of a logic model, criteria for measure selection, and a system for managing measures. It emphasizes the importance of continually updating a set of measures to account for flexibility and context specificity.

Strengths/Limitations: The committee acknowledges the limitations of its focus on LHIs and recognizes the importance of other health issues, such as disaster preparedness and social determinants of health.

Generalizability to Medicare Population: Moderate; the committee proposed measures designed for widespread application and noted their relevance beyond the Medicare population.

Methods: The committee employed a systematic approach to select quality measures. This involved the use of a logic model, criteria for measure selection, and the development of case studies for LHIs under specific topics. The committee also conducted a literature review to provide a general discussion of potential quality measures and interventions.

Conlin PR, Zhang L, Li D, Nelson RE, Prentice JC, Mohr DC. Association of hemoglobin A1c stability with mortality and diabetes complications in older adults with diabetes. *BMJ Open Diabetes Res Care*. 2023;11(2):e003211. doi:10.1136/bmjdr-2022-003211

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: The study investigates the impact of hemoglobin A1c stability within patient-specific target ranges on mortality, macrovascular, and microvascular complications in veterans with diabetes.

Main Findings: Increased time above and below patient-specific A1c target ranges were associated with mortality and macrovascular complications.

Strengths/Limitations: Outcome measures were assessed during a follow-up period, which minimizes the risk of reverse causation. The study also accounted for numerous clinical and patient characteristics that might act as confounding factors. The sample included older adults who were mostly male, limiting the generalizability of the findings. Additionally, the sample included veterans, who tend to exhibit more comorbidities than non-veterans and also utilize VA health care significantly, which would introduce potential differences in outcomes by treatment setting.

Generalizability to Medicare Population: Strong; the sample included VA-Medicare dual enrollees with a diabetes diagnosis.

Methods: A retrospective observational cohort study from 2004 to 2016 of veterans with diabetes and at least four A1c tests within a three-year baseline period. Patients were grouped into categories based on the percentage of time their baseline A1c levels were within specific

target ranges, and the associations of these categories with mortality and complications were investigated.

Conway A, Satin D. The role of pay-for-performance in reducing healthcare disparities: A narrative literature review. *Prev Med.* 2022;164:107274. doi:10.1016/j.yjmed.2022.107274

Subtopic(s): Key Highlights; Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To conduct a literature review to identify the effect of Pay-for-Performance (P4P) on health care disparities and identify design features that are most likely to reduce these disparities.

Main Findings: Six P4P design features, categorized as direct or indirect, were identified that would help reduce disparities by addressing clinical and socioeconomic characteristics. Combining design features and fitting them into the overall payment system would ensure reduced disparities in health care.

Strengths/Limitations: The type of literature supporting each design feature varied significantly. Studies for Risk/Case-Mix typically had outcome studies, while some design features had studies that discussed them in theoretical context. Another design feature was discussed in a quasi-experimental context in which different stratification methods were used.

Generalizability to Medicare Population: Strong; the report focuses on P4P and VBP models.

Methods: A systematic literature review was conducted.

Conwell L, Powell, R, Machta R, Bogen K, Gilman B, Barterian L, McCall N. *Evaluation of the Primary Care First Model.* Mathematica; 2022:201. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/pcf-first-eval-rpt>

Subtopic(s): Key Highlights; Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To assess whether the Primary Care First (PCF) Model reduces hospitalizations and total Medicare expenditures for Medicare FFS beneficiaries served by PCF practices.

Main Findings: CMS launched the PCF Model as part of its efforts to accelerate innovation in primary care delivery and transition practices toward a value-based payment structure for Medicare beneficiaries. More than 80 percent of practices in Cohort 1 were affiliated with health systems or medical groups. In 2022, 2,228 practices enrolled in Cohort 2, representing a more than 200 percent increase in practice participation. PCF payments contributed to practices' revenue stream and helped practices develop operational models to maintain financial stability as they took on more risk. Future data collection will help refine causal pathways to reflect the specific activities that practices undertake and to describe how practices intend these activities to result in changes to short-term and long-term outcomes.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article evaluates a CMS Alternative Payment Model.

Methods: The authors use hypothesized causal pathways to describe the PCF mechanisms, such as care management strategies, through which it is expected to see changes in outcomes.

Chronic Condition Data Warehouse. CCW White Paper: Medicare Claims Maturity. October 2017. <https://www2.ccwdata.org/documents/10280/19002256/medicare-claims-maturity.pdf>

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures

Type of Source: White paper

Objective: To understand the completeness of Medicare claims at different levels of claims maturity.

Main Findings: Over 99 percent of inpatient claims are identified within eight months of a service, hospice and hospital outpatient reach 99 percent at nine months, and skilled nursing facility, home health, and carrier claims are at 99 percent at 10 months post-service. Final action claims are available once claims have reached 12 months post-service maturity. Mature claims are stable and undergo few modifications. As a result, mature claims are ideal for research, policy, and programmatic purposes.

Strengths/Limitations: Generalizability to Medicare Population: Strong; the white paper analyzed Medicare administrative data.

Methods: Medicare administrative data from the CCW database were analyzed. 100 percent of the transactional claims records were available. CMS Medicare claims were divided into three types of claims: institutional claims, non-institutional claims, and Part D Events.

Cuellar A, Helmchen LA, Gimm G, et al. The CareFirst Patient-Centered Medical Home Program: Cost and Utilization Effects in Its First Three Years. *J Gen Intern Med.* 2016;31(11):1382-1388.

doi:10.1007/s11606-016-3814-z

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To evaluate the spending and utilization impacts of the patient-centered medical home (PCMH) model.

Main Findings: A decrease in spending was observed for annual adjusted total claims payments, which was driven mostly by lower spending on inpatient care, emergency care, and prescriptions. The authors also found a decrease in emergency room visits and inpatient admissions during the implementation. The magnitude of the reduction in costs was highest for individuals with chronic conditions.

Strengths/Limitations: While the reduction in spending was significant, the specific amount spent on information and care coordination to implement the program is unknown.

Generalizability to Medicare Population: Moderate; the article evaluates the PCMH model, which includes performance measures (e.g., spending and utilization metrics) that are applicable in the context of Medicare.

Methods: Difference-in-differences estimate in expenditure and utilization across baseline and in program years.

Damberg CL, Sorbero ME, Lovejoy SL, Martsof GR, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. *Rand Health Q.* 2014;4(3):9.

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures; Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To summarize the current state of information around value-based purchasing (VBP) from published literature, publicly available documentation, and discussions with an expert

panel of VBP program sponsors, health care providers/health systems, and academic researchers.

Main Findings: The article concludes that there is still much to be learned about how to implement VBP programs, even though there has been a rise in studies around performance-based payment models, such as P4P (pay-for-performance) models. There is currently limited evidence on the impact of ACOs and bundled payment programs that include clinical quality measures, and there is a lack of ACO evaluation studies. The authors determine that more publicly available information is needed on the lessons learned from VBP programs.

Strengths/Limitations: This article was published in 2014 and could be slightly outdated.

Generalizability to Medicare Population: Strong; Medicare has begun implementing VBP in many health care settings.

Methods: The authors reviewed findings from an environmental scan, literature review, and expert panel discussions.

DeGross A, Schooley M, Chapel T, Poister TH. Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning*. 2010;33(4):365-372. doi:10.1016/j.evalprogplan.2010.02.003

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To explore challenges in developing and implementing performance measurement systems in federal public health and propose solutions to create robust and practical performance measures.

Main Findings: The paper identifies challenges related to the complexity of public health problems, decentralized program implementation, and lack of reliable and consistent data sources and measurement. To address challenges in complexity, strategies include using logic models to enhance conceptual clarity. For decentralization challenges, the authors propose engaging stakeholders and allowing flexibility at local levels through pilot programs. For measurement, solutions involve developing alternative measures for more abstract outcomes and encouraging data quality through audit programs.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; no specific reference to the Medicare population but addressing challenges in performance measurement is beneficial for Medicare programs.

Methods: N/A

Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q*. 2005;83(4):691-729. doi:10.1111/j.1468-0009.2005.00397.x

Subtopic(s): Background

Type of Source: Journal article

Objective: To review and evaluate methods for assessing the quality of medical care.

Main Findings: The article emphasizes the questionable value and lack of rigor in existing methods as limitations and challenges in evaluating the quality of medical care. The author suggests a shift from evaluating quality to understanding the medical care process through empirical studies exploring values in relevant population groups.

Strengths/Limitations: The paper predominantly focuses on evaluating the medical care process at the level of physician-patient interaction. Therefore, it excludes processes related to effective medical care at the community level.

Generalizability to Medicare Population: Weak; the article is not focused on Medicare beneficiaries.

Methods: Exhaustive review of literature emphasizing key studies related to physician-patient interaction.

Duarte M, Pereira-Rodrigues P, Ferreira-Santos D. The Role of Novel Digital Clinical Tools in the Screening or Diagnosis of Obstructive Sleep Apnea: Systematic Review. *Journal of Medical Internet Research*. 2023;25(1):e47735. doi:10.2196/47735

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Journal article

Objective: To review literature with the goal of highlighting and analyzing the most valid digital tools utilized for screening or diagnosing obstructive sleep apnea (OSA) in adults.

Main Findings: The study identified various digital tools, including smartphones, wearables, and sensors, for OSA screening and diagnosis in adults. The authors highlight that, while not intended to replace polysomnography, these digital tools can help identify patients at increased risk for OSA. Contactless devices such as audio recorders that incorporate machine learning emerge as having the greatest potential for OSA screening and monitoring.

Strengths/Limitations: Limitations in the available literature exist in the form of incomplete reporting and unclear enrollment criteria for OSA testing. The prevalence variation of OSA in tested populations raises concerns about generalizability to the broader population. The authors emphasize the need for further research, especially external validation in home settings. This would ensure their applicability and reliability beyond the controlled laboratory conditions of most studies.

Generalizability to Medicare Population: Weak; the study did not focus on Medicare beneficiaries.

Methods: Systematic review of literature; risk of bias assessment.

Dudgeon D. The Impact of Measuring Patient-Reported Outcome Measures on Quality of and Access to Palliative Care. *J Palliat Med*. 2017;20(Suppl 1):S-76-S-80. doi:10.1089/jpm.2017.0447

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To examine global evidence on how measuring indicators of desired outcomes enhances the quality and accessibility of palliative care, with the goal of adapting the findings to the Canadian context.

Main Findings: Patient-reported outcome measures (PROMs) are gaining prominence, with emerging evidence indicating their positive impact on outcomes valued by patients. In Canada, initiatives like the Canadian Partnership Against Cancer's collaboration aim to create common quality measures. Australia's Palliative Care Outcomes Collaborative showcases the feasibility of measuring patient-centered improvements nationally using point of care PROMs.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; the report did not mention Medicare beneficiaries and focused on global approaches to palliative care.

Methods: A systematic review of literature was conducted.

Dullabh P, Dhopeswarkar R, Leaphart D, et al. Trustworthy Artificial Intelligence (TAI) for Patient-Centered Outcomes Research (PCOR). *Assistant Secretary for Planning and Evaluation*; 2023. Accessed December 1, 2023. <https://aspe.hhs.gov/reports/trustworthy-ai-pcor>

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Report

Objective: To enhance comprehension of the application of the Trustworthy Artificial Intelligence (TAI) Playbook's principles to Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) projects.

Main Findings: The authors present three categories for adherence to the Playbook's six TAI principles. Key informants find the six principles comprehensive for AI in PCOR and noted challenges in implementing some. The report also suggests 15 considerations for TAI principles in OS-PCORTF projects that emphasize patient privacy and iterative examination. It also identifies 14 opportunities to improve data practices and resources for AI-enabled research. The principles outlined in the Playbook are crucial when using AI for PCOR; however, implementation is complex and use case-dependent.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; report not focused on Medicare beneficiaries.

Methods: Environmental scan of gray and peer-reviewed literature; key informant interviews with AI experts.

Duseja R, Andress J, Sandhu AT, et al. Development of Episode-Based Cost Measures for the US Medicare Merit-based Incentive Payment System. *JAMA Health Forum*. 2021;2(5):e210451. doi:10.1001/jamahealthforum.2021.0451

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To describe the creation of episode-based cost measures and their significance in transitioning from a system focused on volume-based transactions to value-based purchasing.

Main Findings: The shift to value-based purchasing necessitates an accurate assessment of clinician impact on quality and costs. Episode-based cost measures are crucial to evaluate clinician influence on health care costs, especially for high-priority conditions and procedures. Merit-based Incentive Payment System (MIPS) Value Pathways, introduced by CMS, will align the episode-based cost measures with corresponding quality measures and provide additional incentives for shifting from fee-for-service to value-based care.

Strengths/Limitations: The authors note that the full impact of MIPS and episode-based cost measures (EBCMs) on cost of care is yet to be seen since MIPS is still in its early stages. This early-stage evaluation suggests that ongoing monitoring and adjustment may be necessary as the program matures.

Generalizability to Medicare Population: Strong; the article focused its discussion on beneficiaries populations under MIPS.

Methods: N/A; the report describes the development of measures.

Dusetzina SB, Tyree S, Meyer AM, et al. Linking Data for Health Services Research: A Framework and Instructional Guide [Internet]. *Rockville (MD): Agency for Healthcare Research and Quality (US)*; 2014 Sep. 1, Background and Purpose. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK253315/>

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Book chapter

Objective: To introduce a conceptual framework and instructional insights that explore the merits and constraints of different approaches to linking registries with other data sources in the context of comparative effectiveness research.

Main Findings: The authors present a set of best practices for conducting linkages, as well as how to evaluate and report the accuracy and precision of linkage procedures. Further, the authors apply their data linkage framework to relevant examples to highlight effectiveness. The importance of data linkages is underscored, specifically between registries and health insurance claims.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; the Medicare population is not mentioned, and the focus of the section centers around enhancing data linkage practices.

Methods: The authors define requirements for robust record-linkage of health registries to other data sources, describe the strengths and weaknesses of each approach, and apply their framework to real-world problems.

Gifford K, Eileen E, Edwards BC, Lashbrook A. *States focus on quality and outcomes amid waiver changes*. Henry J. Kaiser Family Foundation. October 2018.

Subtopic(s): Performance Measures Used in the Medicare Advantage Star Ratings Program

Type of Source: Report

Objective: To examine the changes taking place in Medicaid programs across the country in Fiscal Years (FYs) 2018 and 2019.

Main Findings: An increasing number of states were implementing or planning Section 1115 waivers with policies that had enrollment declines. Risk-based managed care continued to be the predominant delivery system for Medicaid services. States were focused on implementing Alternative Payment Models and improving quality within managed care organizations (MCOs). States were working to address social determinants of health within and outside of MCO requirements. Nearly all states in FYs 2018 and 2019 were employing one or more strategies to expand the number of people served in home and community-based settings. In FYs 2018 and 2019, states made provider rate increases compared with restrictions. Positive economic conditions and state priorities resulted in states increasing benefits, such as mental health and substance use disorder treatment. Finally, states continued to focus on cost containment efforts to address rising prescription drug costs and on pharmacy benefit management strategies to address the opioid crisis.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; the report focused on Medicaid policy.

Methods: Analysis of the 18th annual budget survey of Medicaid officials in all 50 states and the District of Columbia.

Giovannetti ER, Dy SM, Leff B, Weston C, Adams K, Valuck T, Pittman AT, Blaum CS, McCann BA, Boyd CM. Performance measurement for people with multiple chronic conditions: Conceptual model. *Am J Manag Care*. 2013;19(10):e359-e366.

Subtopic(s): Trends in Existing Performance Measures

Type of Source: Journal article

Objective: To summarize the development of a conceptual model for Performance Measurement for People with Multiple Chronic Conditions (PM-MCC).

Main Findings: The PM-MCC Model can be used by measure developers, researchers, policy makers, and health plans to implement measurement sets that evaluate and improve health care for patients with multiple chronic conditions. The model focuses on the patient and their family's preferences for care. The model considers preferences for care within the context of multiple care sites (e.g., home-based primary and skilled nursing care), multiple providers, the types of care delivered (e.g., screening, prevention, treatment), and the domains of measurement that apply across sites and types of care (e.g., health and well-being, patient safety, affordable care).

Strengths/Limitations: One strength of the PM-MCC Model is its suitability to guide Accountable Care Organization performance measure development and prioritization because the model cuts across conditions, sites of care, and types of care. One limitation of the model includes the lack of data sources designed to track patients' goals and preferences for care and a lack of data sources designed to support performance measurement for people with multiple chronic conditions.

Generalizability to Medicare Population: Moderate; the article does not reference Medicare specifically; however, the conceptual model can be applied to Medicare.

Methods: The measurement model was developed using reviews of existing performance measurement frameworks, reviews of the literature on multiple chronic conditions, feedback from subject matter experts, and public comment.

Grants NRC (US) P on PM and D for PHPP, Perrin EB, Durch JS, Skillman SM. *Data and Information Systems: Issues for Performance Measurement*. In: Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network. National Academies Press (US); 1999. Accessed November 30, 2023. <https://www.ncbi.nlm.nih.gov/books/NBK231008/>

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Book chapter

Objective: To identify challenges related to using health data resources and provide solutions to strengthen data systems that are used to support performance measurement.

Main Findings: The chapter defines and reviews health data resources. It then explores analytic and operational obstacles associated with using health data, including quality assurance, developing and implementing standards, information technology, and privacy. The authors then outline steps in order to create robust data systems to enhance performance measurement by utilizing a collaborative approach.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; the chapter does not focus on Medicare beneficiaries.

Methods: N/A

Greenhalgh J, Dalkin S, Gooding K, et al. Functionality and Feedback: A Realist Synthesis of the Collation, Interpretation and Utilisation of Patient-Reported Outcome Measures Data to Improve Patient Care. *NIHR Journals Library*; 2017. Accessed November 30, 2023. <http://www.ncbi.nlm.nih.gov/books/NBK409450/>

Subtopic(s): Key Highlights; Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To identify and assess utilization of evidence of PROMs data for improving patient care.

Main Findings: Providers are more inclined to enhance patient care if they perceive PROMs and performance data to be credible and timely with clear indication of the problem source. The authors note that significant and sustained improvement necessitate systemic approaches. Both standardized and individualized PROMs are valued by patients; however, careful consideration is needed to determine which patients benefit.

Strengths/Limitations: The authors note a lack of research focusing on the feedback of aggregate PROMs data to providers.

Generalizability to Medicare Population: Weak; the article does not focus on Medicare beneficiaries.

Methods: Systematic review of peer-reviewed literature; realist synthesis to identify, test, and refine theories about feedback on PROMs in patient care.

Hamilton FL, Greaves F, Majeed A, Millett C. Effectiveness of providing financial incentives to healthcare professionals for smoking cessation activities: Systematic review. *Tob Control*. 2013;22(1):3-8. doi:10.1136/tobaccocontrol-2011-050048

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To review literature and examine evidence related to financial incentives for smoking cessation interventions by health care providers.

Main Findings: Financial incentives seem to improve the documentation of smoking status and the delivery of advice on cessation, along with increasing referrals to smoking cessation services. The authors note, however, that there is a paucity of evidence showing that financial incentives cause reductions in smoking rates.

Strengths/Limitations: The systematic review highlighted the limited evidence on the topic. Additionally, there was too high a degree of statistical heterogeneity for the studies to be used in meta-analysis.

Generalizability to Medicare Population: Weak; the article does not focus on Medicare beneficiaries.

Methods: A systematic literature review was conducted.

Health Care Payment Learning & Action Network. Accelerating and Aligning Population-Based Payment Models: Performance Measurement. *HCPLAN*; 2016. Accessed December 1, 2023. <https://hcp-lan.org/pm-whitepaper/>, <https://hcp-lan.org/pm-whitepaper/>

Subtopic(s): Key Highlights; Best Practices for Measuring Performance

Type of Source: White paper

Objective: To document the principles and key characteristics of a measurement system that will sustain population-based payment (PBP) models that meet the Triple Aim of better care, better health, and lower costs. This white paper also provides guidance on performance measurement and recommendations for use nationally, as well as immediate next steps for stakeholders.

Main Findings: The Population-Based Payment Work Group developed four principles to illustrate its thinking in making recommendations for the future of population-based payment models. The principles are: 1) performance measurement is foundational to the PBP models

potential to advance the Triple Aim; 2) measurements must address the continuum of care; 3) performance measurements should be more outcome-based than current fee-for-service measures; and 4) PBP models must meaningfully incentivize improvement. The recommendations that follow are: 1) ensure that measures are based on results that matter to patients; 2) ensure that core measures are continuously refined to make sure they are comprehensive and outcome-oriented; 3) there should be a governance process in place to oversee and improve development of new measures; 4) improve national infrastructure; 5) develop meaningful incentives for high-quality care; 6) the measurement of outcomes should encourage system-wide improvements rather than one that mandates a system where there are winners or losers; and 7) measures must be developed from good science and implementation.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the performance measures that are developed as a result of this white paper would be utilized in programs intended for improving quality of care for the Medicare population.

Methods: N/A

Heidenreich P, Sandhu A. Pursuing Equity in Performance Measurement. *Circulation*. 2023;147(15):1134-1136. doi:10.1161/CIRCULATIONAHA.123.064123

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Editorial

Objective: To provide commentary and compare different approaches to performance measurements and how they address equity concerns.

Main Findings: The authors of this editorial posit that there are gaps in current performance measurements for quality of hospital care. These gaps lead to an inaccurate rating system that has implications for best practice guidelines and does not lead to equitable improvement across all hospitals. The authors review a study in the same journal which used separate risk models for race and ethnicity and aggregated the models with equal weights to estimate hospital outcome measures. The editorial authors conclude that race-specific models are useful for helping to detect inequitable care but can be tricky to use in drawing conclusions if the user is trying to discern the effect of differences of social determinants of health versus racist care, especially if non-White is used as a proxy for social disadvantage. The editorial authors conclude that the proposed outcome measurement tool developed in the study address both equity and optimal risk prediction for all groups, which is more useful than current methods to determine penalizing or rewarding hospitals based on the outcomes of the majority alone.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the study being reviewed is proposing an alternative method to CMS' current approach to measure quality of hospital care.

Methods: N/A; this is an editorial about another study in the same issue of this journal.

Hotstetter M, Klein S. Using Patient-Reported Outcomes to Improve Health Care Quality. *The Commonwealth Fund*. <https://www.commonwealthfund.org/publications/newsletter-article/using-patient-reported-outcomes-improve-health-care-quality>

Subtopic(s): Background

Type of Source: News article

Objective: To explain what patient-reported outcome measures (PROMs) are, provide examples, and show how they can be used to improve health care quality.

Main Findings: This article provides examples of how PROMs are used in clinical practice and suggest options for how they could be developed to help improve care coordination. The authors address obstacles to PROM use including the possibility that collection will disrupt the workflow of clinicians and is not billable, as well as PROMs being imperfectly linked to specific treatments. PROMIS, a program started in 2004, is expected to address some of these concerns by “creating more precise measures and reduce the number of questions needed to make them more feasible for use in clinical practice.” PROMs are expected to help with comparative effectiveness research and help improve health care quality. Sarah Scholle at the National Committee for Quality Assurance suggested that Accountable Care Organizations are particularly well-suited to roll up PROMs for use in this way due to their patient population being large enough for good sample sizes and the ability to have an effect on patient care.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article suggests that ACOs are ideal for using PROMs to improve health care quality.

Methods: N/A

Huang J, Yin S, Lin Y, Jiang Q, He Y, Du L. Impact of pay-for-performance on management of diabetes: A systematic review. *Journal of Evidence-Based Medicine*. 2013;6(3):173-184. doi:10.1111/jebm.12052

Subtopic(s): Performance Measures Used in Other Models

Type of Source: Journal article

Objective: To review studies that looked at the impact that pay-for-performance (P4P) has on management of diabetes, including the size of the effects and the quality of the body of evidence.

Main Findings: The authors found that a majority of the studies reported positive effects of P4P programs on management of diabetes. Larger effects were found in physician groups, particularly that one HbA1c indicator was high for physician groups but very low for patients. This review came to similar conclusions as other systematic reviews looking at P4P programs in finding that physician behavior measured by is directly affected by these programs. Additionally, patient outcomes (proportion of patients with managed levels of HbA1c, total cholesterol, and blood pressure) were found to improve after the introduction of P4P programs.

Strengths/Limitations: The lack of certain indicators in all studies, as well as the inherent heterogeneity of the different studies, made synthesizing data difficult. The authors also acknowledged that there may be other factors affecting implementation of P4P programs, including equity and cost-effectiveness. While comparing other reviews to their own, the authors noted that having updated research in their review created a more comprehensive picture of the effect of P4P implementation on management of diabetes.

Generalizability to Medicare Population: Moderate; of the 21 studies included in this review, eight of them were based in the U.S. health care system, though none specifically studied the Medicare population alone. However, diabetes is seen worldwide, and the same measures were compared across all studies, making this study useful for all providers who manage patients with diabetes.

Methods: Electronic search of online databases. Both meta and descriptive analyses were performed on included studies to synthesize results. The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system was used to determine the quality of the evidence.

Ibrahim H, Liu X, Zariffa N, Morris AD, Denniston AK. Health data poverty: An assailable barrier to equitable digital health care. *The Lancet Digital Health*. 2021;3(4):e260-e265. doi:10.1016/S2589-7500(20)30317-4

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To describe the problem of data health poverty and the inequities that will be created if access to health technologies is not expanded.

Main Findings: As health technologies become more advanced and more widely applicable, significant sectors of the population are being left out, which has significant implications for diagnostic care, as well as health systems improvement. The authors recommend four areas to improve the current state of health data: 1) increase the awareness data health poverty among digital health communities by including information about the composition of training datasets which can improve accountability of digital health tool developers; 2) transparently and effectively communicate to citizens how their data can improve the health care environment; 3) improve the equity of access to digital devices and services so that populations are not left out of data collection efforts; and 4) build inclusive and representative datasets to support equitable discovery and innovation in digital health care.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; this article is aimed toward generalized improvement of the quality of data collection efforts and building better data sets to use in health policy and health care provision.

Methods: Electronic search through PubMed, Google Scholar and Google search engine using specific search terms.

Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS — The Universal Foundation. *N Engl J Med*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures; Performance Measures Used in CMMI Models; Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To introduce and explain the Universal Foundation as a new standard to streamline quality measures that will apply to as many CMS quality-rating and value-based care programs as possible.

Main Findings: The goals of the Universal Foundation are to focus providers on measurements that are meaningful for the health of the broadest segments of the population; reduce provider burden; advance equity; aid the transition to seamless, automatic digital reporting; and allow comparisons among various programs. Quality measures selected are aimed to target the diseases and conditions with the highest morbidity and mortality rates in the United States. Measures also reflect the provision of high-quality preventive care, identification and treatment of depression and substance use disorders, and care coordination after hospitalization.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; this article was published by leaders of several CMS centers, and the measures are designed to improve quality improvements in CMS programs.

Methods: N/A

Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. *McKinsey & Company Healthcare Systems and Services Practice*. 2018. Accessed December 13, 2023. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>

Subtopic(s): Key Highlights; Performance Measures Used in the Medicare Advantage Star Ratings Program

Type of Source: Report

Objective: To assess whether the Medicare Advantage (MA) Star Ratings program achieves the program's goals of incentivizing health insurers to improve their MA plans and helping beneficiaries enroll in high-quality health plans.

Main Findings: Since the beginning of the quality bonus payment (QBP) demonstration, MA plans' Star Ratings have improved while controlling or reducing costs. Plans' scores and the percentage of MA beneficiaries enrolled in high-performing plans have increased over time. Average Star Ratings have improved despite increasingly higher standards used to assess plan performance. MA plans deliver Parts A and B benefits at 90 percent of the cost of Medicare fee-for-service (FFS) coverage.

Strengths/Limitations: The ratings calculated for a given year reflect the performance of MA plans two years earlier because of data reporting limitations.

Generalizability to Medicare Population: Strong; the report focuses on a Medicare program and its impact on MA beneficiaries.

Methods: An analysis of MA plan performance on quality (measured by Star Ratings) was conducted. In addition, the financial performance of MA plans was compared with Medicare FFS.

Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. *Br J Radiol*. 2019;92(1100):20190001. doi:10.1259/bjr.20190001

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To provide an overview of how machine learning can be applied to treatment planning and deliver of radiation oncology.

Main Findings: Machine learning has many practical applications for clinical workflow in radiation oncology, but there are still areas where human insight is the better choice. One advantage of using machine learning is to bring efficiency and consistency to solutions for problems which are well-defined, have a ground truth for which there is sufficient data, and if there is a quantitative measure that the algorithm can be optimized on through training. However, if the problems are not yet clearly defined or there are not enough data to support a ground truth, then the work is still better suited to humans for more efficiency.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; although the article does not focus specifically on Medicare beneficiaries receiving cancer treatment, the focus of this article applies to Medicare beneficiaries.

Methods: Literature review and analysis of the current state of machine learning technology in radiation oncology.

Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database Syst Rev.* 2021;2021(1):CD011865. doi:10.1002/14651858.CD011865.pub2

Subtopic(s): Key Highlights; Performance Measures Used in Other Models

Type of Source: Journal article

Objective: To compare payment methods used in outpatient health care settings in order to understand the different impacts they have on quantity and quality of health care services delivered.

Main Findings: The review looked at four comparison groups: 1) pay-for-performance (P4P) plus existing payment methods compared with existing payment methods; 2) fee-for-service (FFS) compared with existing payment methods; 3) FFS mixed with existing payment methods compared with existing payment methods; and 4) enhanced FFS compared with FFS. Comparison group 1 found that incentives were likely to improve the quantity of child immunization status; comparison group 2 had too much uncertainty of results to report anything substantive; comparison group 3 found that the mixed payment method may increase the quantity of health services provided; and comparison group 4 found that higher FFS payments probably increased child immunization rates.

Strengths/Limitations: The authors noted that none of the outcomes in the review were assessed as having high-certainty evidence, which may restrict the applicability of the findings for settings with limited resources. Additionally, there was limited information on any outcome other than quantity of health services provided.

Generalizability to Medicare Population: Moderate; the findings are not targeted to assess the Medicare population, but the payment mechanisms are applicable.

Methods: Literature review performed followed by a structured synthesis, meta-analysis, and mean/median reporting for categories that were not right for meta-analysis.

Joseph K, Udogwu UN, Manson TT, et al. Patient Satisfaction After Discharge Is Discordant With Reported Inpatient Experience. *Orthopedics.* 2021;44(3):e427-e433. doi:10.3928/01477447-20210415-01

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To determine if there is a significant difference between reporting patient satisfaction after discharge versus while still in inpatient care and to determine if method of survey improves response rate.

Main Findings: Patient satisfaction scores provided as an inpatient were often not in agreement with post-discharge surveys. The authors found that negative patient experiences specific to pain management, responsiveness, and hospital environment provided as an inpatient were more likely to be in agreement with post-discharge surveys, but positive experiences or neutral were not in agreement between the two time periods. Participants were also more likely to respond to a phone survey than a mailed survey. The authors suggest that the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey be reviewed for validity considering the discordance in inpatient and post-discharge responses.

Strengths/Limitations: The qualitative methods used in the study provided more context and allowed patients to provide a balanced view of the positive and negatives of their visit. However, the study used only one survey instrument at one academic medical center and therefore may have limited generalizability.

Generalizability to Medicare Population: Moderate; while the study was not targeted to the Medicare population, the HCAHPS is a common survey that is often used as a performance measurement in Medicare Alternative Payment Models.

Methods: Mixed methods study. Qualitative methods were used to assess patient experience, and quantitative methods were used to compare patient responses at two defined time points.

Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: A Narrative Review. *J Gen Intern Med.* 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8

Subtopic(s): Performance Measures Used in Other Programs; Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To review and classify the current (2017) Medicare payment models to understand the variability of their mechanisms and how they may inform future policymaking decisions.

Main Findings: Health care is rapidly moving into an era dominated by alternative payer models and value-based purchasing models, but program designs across the health care industry vary widely, and there are not enough data to provide evidence of what methods work best. In reviewing six different programs, the authors examined four areas and found that: (1) program scopes may benefit from specific, targeted programs for high-priority conditions and broad, frequently update programs to improve care more generally; (2) the difference between using absolute or relative performance targets may be determined by understanding the difference between the provider perspective and the payer perspective; (3) rewarding improvement or achievement can significantly impact which providers do well and poorly under a program; and (4) framing incentives as penalties or bonuses can affect how providers respond. Ultimately, the authors suggest more studies evaluating these programs in order to more directly compare approaches and inform future policy decisions.

Strengths/Limitations: A significant limitation in this review is the lack of comparative data on these programs.

Generalizability to Medicare Population: Strong; the review looks at Medicare payment models that are current as of writing in an effort to improve the future of the health care industry.

Methods: N/A

Kansagara D, Tuepker A, Joos S, Nicolaidis C, Skaperdas E, Hickam D. Getting Performance Metrics Right: A Qualitative Study of Staff Experiences Implementing and Measuring Practice Transformation. *J GEN INTERN MED.* 2014;29(S2):607-613. doi:10.1007/s11606-013-2764-y

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To understand how the Patient Aligned Care Team (PACT) in the Veterans Health Administration (VHA) respond to the use and implementation of performance metrics; whether they align with PACT principles or have unintended consequences on the provision of care.

Main Findings: The interviews with clinical staff illuminated significant issues in the implementation of performance metrics. Performance measures used were not always in line with PACT team-based care principles, leading to contradictory approaches to patient care; staff also felt that hitting certain measurement goals was perceived as more important than patient-centered care and required a substantial investment of time. Additionally, more transparency was needed for staff to understand why certain performance metrics were decided on over

others, especially when the metrics chosen did not adequately reflect a team-based model of care.

Strengths/Limitations: Since this was a qualitative analysis, the findings of this study may not be generalizable to the wider system, but do indicate a need to look further into the implementation of performance metrics in the VHA system.

Generalizability to Medicare Population: Moderate; is study is focused on the VHA system, but there may be crossover in the populations who utilize VA benefits and Medicare benefits.

Methods: Focus groups with qualitative analysis of interviews.

Kaufman BG, Bleser WK, Saunders R, et al. Prospective or retrospective ACO attribution matters for seriously ill patients. *The American Journal of Managed Care*. 2020;26(12):534-40. doi: 10.37765/ajmc.2020.88541

Subtopic(s): Key Highlights; Performance Measures Used in the Medicare Shared Savings Program

Type of Source: Journal article

Objective: To examine how Accountable Care Organizations' (ACOs') choice of retrospective or prospective attribution method influenced incentives for care among seriously ill patients.

Main Findings: Compared with beneficiaries surviving 270 days or longer, dying in the first 90 days of the performance year was associated with reduced odds of retrospective attribution. In addition, hospice use was associated with reduced odds of retrospective attribution. For ACOs that did not have shared savings, the average per capita Medicare expenditures was higher for prospective compared with retrospective ACO populations. Results suggest ACOs could use different care management strategies to improve performance depending on the attribution method.

Strengths/Limitations: Findings may be limited to the Medicare Shared Savings Program-attributed population that meets criteria for serious illness.

Generalizability to Medicare Population: Strong; the article focused on seriously ill Medicare beneficiaries.

Methods: A cross-sectional analysis was conducted on 100 percent Medicare Master Beneficiary Summary and Medicare Shared Savings Program beneficiary files (2014-2016). The average within-ACO difference between potential retrospective and prospective ACO populations was estimated.

Kemp E, Trivitt J, Davis C. Evidence-Based Performance Indicators of Positive Inpatient Experiences. *J Healthc Manag*. 2023;68(2):106-120. doi:10.1097/JHM-D-22-00147

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To understand what factors contribute to positive inpatient experiences.

Main Findings: Positive patient experience ratings were associated with hospital quietness, nurse communication, and care transition. Additionally, hospitals that had better clinical outcomes had higher patient experience ratings. Hospitals that serve higher-risk communities and those with more complications had lower patient experience scores.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; the study was not targeted for Medicare patient experience, but many of the measures collected are used to assess quality of care for Medicare patients.

Methods: Data were collected from four publicly available U.S. government datasets: HCAHPS (patient experience), CMS (hospital quality), Social Vulnerability Index (social determinants of health), and the Offices of Policy Development and research (zip code). A linear regression model was used to analyze the data.

Kessell E, Pegany V, Keolanui B, Fulton BD, Scheffler RM, Shortell SM. Review of Medicare, Medicaid, and commercial quality of care measures: Considerations for assessing accountable care organizations. *Journal of Health Politics, Policy and Law*. 2015;40(4):761-96. doi: 10.1215/03616878-3150050

Subtopic(s): Performance Measures Used in the Medicare Advantage Star Ratings Program

Type of Source: Journal article

Objective: To assess quality measurement across public and private sectors.

Main Findings: Two-thirds of quality measures were categorized as process measures, and approximately 20 percent of quality measures were categorized as outcome measures. Patient experience and structure measures made up approximately 8 percent and 7 percent of measures, respectively.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; although quality measures in Medicare were considered, the article considered quality measures across both public and private sectors.

Methods: Available quality measures were reviewed to estimate the number and domain of measures.

Khullar D, Bond AM, O'Donnell EM, Qian Y, Gans DN, Casalino LP. Time and Financial Costs for Physician Practices to Participate in the Medicare Merit-based Incentive Payment System: A Qualitative Study. *JAMA Health Forum*. 2021;2(5):e210527. doi:10.1001/jamahealthforum.2021.0527

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To understand the administrative, time, and cost burden experienced by physician practices who participate in the Medicare Merit-based Incentive Payment System (MIPS).

Main Findings: Providers reported significant time and costs associated with participating in MIPS. On average, in 2019, the study found that it cost practices \$12,811 per physician to participate in MIPS, and that physicians, clinical staff members, and administrators collectively spent 201.7 hours annually on the program. In terms of patients, the amount of time spent on MIPS could be used to provide care for an additional 212 patients per year. The authors suggest that the given the outsized burden and the studies showing its effects on improvement of quality and outcomes for patients, reducing reporting burden may be warranted.

Strengths/Limitations: This study collected information only on the 2019 performance year. Due to the qualitative study design, the study is subject to response bias; however, cost and time estimates were found to be consistent with prior studies done on MIPS.

Generalizability to Medicare Population: Strong; MIPS is a value-based purchasing program under the Medicare Access and CHIP Reauthorization Act.

Methods: Qualitative interviews of 30 physician practices based across the United States and in a variety of settings. The interview guide was developed based on a thorough literature review, and efforts were made to separate costs of MIPS from cost of other quality programs.

Kim KM, Max W, White JS, Chapman SA, Muench U. Do penalty-based pay-for-performance programs improve surgical care more effectively than other payment strategies? A systematic review. *Ann Med Surg (Lond)*. 2020;60:623-630. doi:10.1016/j.amsu.2020.11.060

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To determine the impact of pay-for-performance (P4P) penalty design on quality and cost outcomes on surgical care, as compared with a reward design or a combination reward and penalty design.

Main Findings: The authors found that P4P programs that utilize a penalty design could be more effective than programs that use a reward or combination reward and penalty design. There was a higher proportion of studies that showed positive effects due to a penalty design, whereas reward or combination designs showed null or non-significant effects. The authors attribute this to the behavioral economics theory of loss aversion, which shows that organizations tend to respond to losses more than gains.

Strengths/Limitations: There was a lack of studies that evaluated the same studies, which meant that a meta-analysis could not be performed. The studies reviewed for this article varied significantly in design features, and the findings are not generalizable to all surgical procedures as the P4P programs in this study primarily targeted coronary artery bypass graft, total hip arthroplasty, and total knee arthroplasty.

Generalizability to Medicare Population: Strong; many P4P programs were launched by Centers for Medicare & Medicaid Services to address quality and cost of care for their populations.

Methods: Systematic review of the literature focused on P4P programs that targeted surgical care with primary outcomes of interest being quality and cost of care.

Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open*. 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115

Subtopic(s): Key Highlights; Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To evaluate the association between the Hospital Acquired Conditions Present on Admission (HAC-POA) program and the impact on improved surgical care outcomes, specifically surgical care quality and costs.

Main Findings: The implementation of HAC-POA, a penalty design pay-for-performance program, was found to be associated with improved surgical care. Over the time period examined by the study, the incidence of surgical site infections (SSI) significantly decreased, as did the average length of stay (LOS) and hospital costs. However, there was not a statistically significant decrease in deep vein thrombosis (DVT) or in-hospital mortality.

Strengths/Limitations: There may be missing data on surgical complications as a result of undetected or underreporting and the reliance on administrative data which are often self-reported. Additionally there may be confounding factors from contemporaneous policies implemented during the study period.

Generalizability to Medicare Population: Strong; the HC-POA is a national CMS program targeted to address quality and cost of care for its populations.

Methods: Cross-sectional study using the National Inpatient Sample to examine the association between the HAC-POA program and the incidence of SSI and DVT, LOS, in-hospital mortality, and hospital costs.

Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med.* 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To understand the effect that key factors have on for pay-for-performance (P4P) program implementation.

Main Findings: The kinds of measures which tend to induce more positive change are process-of-care and clinical outcome measures that are evidence-based and viewed as clinically important. Incentivized measures should align with institutional priorities. While larger incentives are more likely to improve performance, they should not be too large as to not be cost-effective or encourage gaming. P4P programs need to be flexible and respond to evolving needs as the program continues; for example, by re-evaluating high-achieving measures and determining if thresholds should be raised or if high-achieving areas should be de-emphasized to encourage targeting areas of poor performance.

Strengths/Limitations: P4P programs have significantly heterogeneity, and therefore the findings from this study should not be assumed to be widely applicable. The authors' review and conclusions were also weighted heavily toward programs targeting ambulatory care, further reducing generalizability.

Generalizability to Medicare Population: Moderate; this study is focused on the VHA system and programs targeting patient populations similar to the VHA, but there may be crossover in the populations who utilize VA benefits and Medicare benefits.

Methods: Literature search and review supplemented with 14 Key Informant Interviews.

Krishnan M, Brunelli SM, Maddux FW, et al. Guiding principles and checklist for population-based quality metrics. *Clin J Am Soc Nephrol.* 2014;9(6):1124-1131. doi:10.2215/CJN.11061013

Subtopic(s): Key Highlights; Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To evaluate the National Quality Forum's (NQF's) criteria for quality measure development and to develop a parallel clinical performance measure checklist for use by the dialysis community in end-stage renal disease (ESRD) care.

Main Findings: The development of clinical performance measures should be reflective of the outcomes that providers, payers, and policy makers wish to see in a value-based care system. This means that measures should be derived from true facility-based management. For example, evidence-based measures should reflect studies based on patient population level rather than those that focus on individual patient treatment. The definition of reliable data should include statistical, year over year, and technical reliability with validation occurring at the facility level. Quality measures should be feasible for the infrastructure in use, and plausible conceptual models should exist that allow the facility or provider to impact the metric through quality improvement efforts.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; this paper was written based on care for ESRD patients for whom Medicare is a major payer.

Methods: Review NQF criteria and develop a checklist for quality measurement development.

Kristensen PK, Thillemann TM, Søballe K, Johnsen SP. Are process performance measures associated with clinical outcomes among patients with hip fractures? A population-based cohort study. *International Journal for Quality in Health Care*. 2016;28(6):698-708. doi:10.1093/intqhc/mzw093

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To understand the association between process performance measures and clinical outcomes among patients with hip fractures.

Main Findings: Early mobilization and higher-quality in-hospital care resulted in more positive clinical outcomes, including a lower 30-day mortality rate.

Strengths/Limitations: There were likely confounders and not all the process performance measures could provide information about whether the patients were treated appropriately based on the results of the assessments, but the study design took care to reduce the risk of selection and information bias.

Generalizability to Medicare Population: Moderate; the age range of the population studied was 65+ but the population was located in Denmark under the Danish health care system.

Methods: A population-based follow-up survey was conducted.

Le Corroller AG, Bonastre J. Patient-reported measures: How useful in health economics? *Eur J Health Econ*. 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z

Subtopic(s): Background; Best Practices for Measuring Performance

Type of Source: Editorial

Objective: To determine whether or not patient-reported outcomes measures are useful from health-economic point of view.

Main Findings: The use of patient-reported measures in health care can be categorized according to the end purpose of the measure and its recipient (patients, decision policy makers, and/or health care providers). PROMs have been used in national programs to compare health care provider performances in routine practice.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; patient-reported outcome measures are relevant to Medicare models and programs.

Methods: This was a commentary.

Liao JM, Navathe AS. Does the Merit-Based Incentive Payment System disproportionately affect safety-net practices? *JAMA Health Forum*. 2020;1(5):e200452. doi:10.1001/jamahealthforum.2020.0452

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To combine year 1 performance data as part of a portfolio monitoring and evaluating value-based payment programs.

Main Findings: Insight from granular MIPS year 1 performance data suggests that safety-net practices may perform more poorly than their non-safety-net counterparts. Policy makers should monitor for such dynamics and consider ways to adjust MIPS policy to ensure that these practices are not inappropriately penalized by the program.

Strengths/Limitations: There is little knowledge about practice-level MIPS performance. More research is needed for more in-depth reporting analysis.

Generalizability to Medicare Population: Strong; CMS continues using the program to incentivize clinicians to improve health care value for Medicare beneficiaries.

Methods: Not specified; used a year 1 track analysis of MIPS performance data.

Lilford RJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. *BMJ*. 2007;335(7621):648-650. doi:10.1136/bmj.39317.641296.AD

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To explore how quality should be measured.

Main Findings: Data on quality can be used either for internal quality improvement or for external reporting. In the first scenario, data are collected by an organization or individual for internal audit in the spirit of continuous improvement. The main disadvantage of measuring outcomes arises from the low signal to noise ratio: outcomes are likely to be affected by factors other than the quality of care. Process standards used in performance management should be valid in that they must either be self-evident measures of quality or be evidence-based. However, validity is not sufficient—the standards must also be genuinely important to health care. Process measures are the most suitable management tool for judging and rewarding quality. Clinical outcomes are likely to be affected by factors other than the quality of care. Outcome measures provide insufficient information about how to improve. Assessment of process encourages universal improvement rather than focusing on outliers. Selected measures must be valid and important.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; performance measurement is heavily utilized in Medicare.

Methods: Data analysis and meta-analysis were conducted.

Loehfelm TW. Artificial intelligence for quality improvement in radiology. *Radiol Clin North Am*. 2021;59(6):1053-1062. doi:10.1016/j.rcl.2021.07.005

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Report

Objective: To assess turnaround time (TAT) as an efficient and plausible measure as a quality metric.

Main Findings: Measuring the improvement of AI requires first that each radiology group identifies their customers, defines what “quality” and “value” means to radiology, and develop metrics that support the pursuit of improving quality and value. Using the Radiology Report Value Equation from Eberhardt and Heilbrun as a reference framework, informatics tools can help to provide automated and objective assessments of report accuracy, utility, clarity and conciseness (i.e., readability), and timeliness. Readability metrics from general linguistics fields can be applied to the authors’ reports as a first-pass effort to uphold accountability for producing comprehensible reports.

Strengths/Limitations: Simple turnaround times (TATs) are a “poor” surrogate for measuring quality in AI operational efficiencies.

Generalizability to Medicare Population: Moderate; understanding the use of AI for quality improvement in radiology could impact Medicare models and programs.

Methods: Metrics; quality initiatives; workflow assessment.

Machta R, Peterson G, Rotter J, Stewart K, Heitkamp S, Platt I, Whicher D, Calkins K, Kranker K, Barterian L, McCall N. *Evaluation of the Maryland Total Cost of Care Model: Implementation Report*. Mathematica; 2021. <https://www.mathematica.org/publications/evaluation-of-the-maryland-total-cost-of-care-model-implementation-report>

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To evaluate the Maryland Total Cost of Care (TCOC) Model.

Main Findings: In its first two years (2019 and 2020), the Maryland TCOC Model has engaged a wide range of providers and begun to transform care outside the hospital. Taken together, this engagement and care transformation can potentially improve targeted outcomes, capitalizing on the substantial room for improvement present at the start of the model. Although the state made progress in reducing avoidable hospital use and reducing hospital spending growth during the Maryland Alternative Payment Models, there remains meaningful room to further reduce avoidable acute care. There are substantial opportunities for improvement in areas newly targeted in the model, including reducing non-hospital spending, improving care coordination across providers, improving ambulatory care to reduce avoidable admissions, and reducing BMI and diabetes incidence. Future evaluation efforts will assess whether the model achieves these aims.

Strengths/Limitations: This is the first report.

Generalizability to Medicare Population: Strong; the report focuses primarily on Medicare Parts A and B.

Methods: Matched comparison group; interviews; Medicare Parts A and B Claims Data; CMS/HSCRC Implementation datasets; savings targets; fixed spending percentage; national spending growth; hospital spending growth.

Makhni EC. Meaningful clinical applications of patient-reported outcome measures in orthopaedics. *J Bone Joint Surg Am*. 2021;103(1):84-91. doi:10.2106/JBJS.20.00624

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Review

Objective: To understand the development and applications of PROMs in orthopedic medicine.

Main Findings: PROMs represent a valuable tool for patient-centered health state assessment in orthopedics that can serve as a foundation for clinical research. Implementing PROM collection in the ambulatory setting has many logistical, financial, and administrative challenges. Overcoming these challenges allows providers to incorporate these tools into routine clinical care, as well as into meaningful research, value, and quality initiatives. The most potentially important role of PROMs is as a facilitator of shared decision-making with patients along with post-interventional monitoring. Successful incorporation of PROMs into daily practice can improve not only outcomes but also quality and value.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; PROMs are used in Medicare models and programs.

Methods: Clinical examples of the application of PROMs.

Markovitz AA, Ayanian JZ, Sukul D, Ryan AM. The Medicare Advantage Quality Bonus Program has not improved plan quality. *Health Affairs*. 2021;40(12):1918-1925. doi:10.1377/hlthaff.2021.00606

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To evaluate the efficacy of the Medicare Advantage (MA) Quality Bonus Program's incentive structure and targeted performance measures from 2009 to 2018.

Main Findings: The authors analyzed insurance claims from 2009 through 2018 from the nation's largest MA claims database for 3,753,117 MA beneficiaries (treatment group) and 4,025,179 commercial enrollees (control group). The authors evaluated changes in performance on nine claims-based measures of quality in both groups before and after the start of the bonus program and with adjustment for differential pre-period trends. The authors observed no consistent differential improvement in quality for MA versus commercial enrollees under the quality bonus program. Participation in the quality bonus program was associated with significant quality improvements among MA beneficiaries on four measures, significant declines on four other measures, and no significant change in overall quality performance. The results suggest that the quality bonus program did not produce the intended improvement in overall quality performance of MA plans.

Strengths/Limitations: The authors' analysis was limited by the presence of non-parallel pre-period trends in quality between MA and commercial enrollees. Additionally, estimated changes in quality may be affected by differential changes in the composition of MA versus commercial enrollees. The authors did not capture changes in intermediate measures.

Generalizability to Medicare Population: Strong; the article focused on a Medicare program.

Methods: Analysis of MA Claims Database from 2009-2018; difference-in-differences framework.

Martino SC, Ahluwalia S, Harrison J, Kim A, Elliott MN. *Developing health equity measures*. The Office of the Assistant Secretary for Planning and Evaluation (ASPE) at the U.S. Department of Health and Human Services. 2021. Accessed December 13, 2023. <https://aspe.hhs.gov/reports/developing-health-equity-measures>

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To identify health equity measurement approaches and assess which approach(es) should be considered for inclusion in Medicare's value-based purchasing (VBP) programs, quality reporting efforts, and confidential reports.

Main Findings: Ten health equity measurement approaches were identified, which fit into three categories: approaches focused on measure identification, approaches focused on measure-by-measure comparisons, and approaches focused on creating a summary index. A health equity definition and a set of guidelines for measuring health equity were developed. Across all 10 approaches, the CMS Office of Minority Health (OMH) Health Equity Summary Score (HESS) received the highest rating from the technical expert panel (TEP). This measurement approach could potentially be incorporated into a Medicare VBP or quality reporting program.

Strengths/Limitations: The project team constructed a list of eight first-choice experts and eight alternates in case any of the first-choice experts could not join the TEP. One strength of this project is that all first-choice experts agreed to participate on the TEP.

Generalizability to Medicare Population: Strong; the report focused on understanding which health equity measurement approaches merit inclusion in Medicare programs and activities.

Methods: A literature review was conducted to identify health equity measurement approaches for assessing performance. Feedback from a TEP was collected and synthesized to understand

the use and potential impact of health equity measurement approaches in Medicare's VBP programs, quality reporting effects, and confidential reports.

Mathes T, Pieper D, Morche J, Polus S, Jaschinski T, Eikermann M. Pay for performance for hospitals. *Cochrane Database Syst Rev.* 2019;2019(7):CD011156. doi:10.1002/14651858.CD011156.pub2

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Review

Objective: To assess the impact of P4P for in-hospital delivered health care on the quality of care, resource use, and equity.

Main Findings: The authors included 27 studies (20 CBA, 7 ITS) on six different P4P programs. Authors analyzed between 10 and 4,267 centers. All P4P programs targeted acute or emergency physical conditions and compared a capitation-based payment scheme without P4P to the same capitation-based payment scheme combined with a P4P add-on. Two P4P program used rewards or penalties; one used first rewards and then penalties; two used penalties only; and one used rewards only. Four P4P programs were established and evaluated. It is uncertain whether P4P, compared with capitation-based payments without P4P for hospitals, has an impact on patient outcomes, quality of care, equity, or resource use as the certainty of the evidence was very low for all P4P programs.

Strengths/Limitations: The authors found no data on utilization, health care provider outcomes, and adverse effects.

Generalizability to Medicare Population: Moderate; the article was not specifically focused on CMS.

Methods: Literature review; clustered trials; narrative synthesis.

McGlynn EA. Improving the quality of U.S. health care — What will it take? *New England Journal of Medicine.* 2020;383(9):801-803. doi:10.1056/NEJMp2022644

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Commentary

Objective: To discuss the limited progress to date in reaching higher-quality care in the United States, which requires reconsidering approaches to measurement, financing, and organizational structures and a new emphasis on social needs.

Main Findings: The author commented that the U.S. health system should start by creating the financial and organizational conditions for changing care delivery from a reactive, fragmented enterprise to one that is coordinated and longitudinal, reflecting the need for systems that can effectively manage chronic disease. The U.S. current system of measuring discrete events reinforces fragmentation and may not lead to overall quality improvement. New approaches that use data from electronic health records rather than claims data and that allow for nuances that make clinical sense will require investments, testing, and deployment. Finally, as the COVID-19 epidemic has demonstrated, there is a need to link health care systems with appropriately resourced public health and community-based services. A variety of programs designed to make these linkages are under way throughout the country. These approaches must undergo systematic evaluations that will assess whether and under what conditions they work. There is opportunity to invest in quality in ways that lay a foundation for a healthier America.

Strengths/Limitations: The author did not mention Medicare or the Medicaid population.

Generalizability to Medicare Population: Weak; there was no mention of Medicare or Medicaid population.

Methods: Statistical review.

McWilliams JM. Pay for performance: When slogans overtake science in health policy. *JAMA*. 2022;328(21):2114-2116. doi:10.1001/jama.2022.20945

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Editorial/Opinion

Objective: To understand the impact of pay-for-performance and alternative approaches to improve quality.

Main Findings: To date, dozens of pay-for-performance arrangements have been implemented as stand-alone programs or components of broader risk contracts. No longer demonstrations, many are now fixtures in the payment system by statute. Conceptually, there are many reasons why pay-for-performance may not produce desired results. A core challenge is that health care quality is a complex, multidimensional construct that has evaded attempts to capture it accurately and comprehensively with a tractable measure set. The potential quality gains from Alternative Payment Models such as Accountable Care Organizations may lie not in the pay-for-performance component of those models but in the flexibility afforded by the population-based payment component to choose the mix of services that produces the best care.

Strengths/Limitations: Limited solutions for resolving issues in pay-for-performance model

Generalizability to Medicare Population: Strong; this article was directly related to the CMS measures program.

Methods: A meta-analysis and literature review were conducted.

McWilliams JM. Professionalism revealed: Rethinking quality improvement in the wake of a pandemic. *NEJM Catalyt*. 2020;1(5). doi:10.1056/CAT.20.0226

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To discuss the limitations of measure-focused approaches, describe a framework for conceiving a next generation of initiatives that aim to improve care by more productively leveraging professionalism, and provide directions for policy and practice.

Main Findings: First, population-based payments give organizations flexibility in selecting the inputs used to produce health by decoupling revenue from the services provided. Removing counterproductive fee-for-service (and pay-for-performance) incentives may be important for disinhibiting the motivation of organizational leaders and physicians. As physicians are increasingly employed by large organizations accountable principally to shareholders or boards, patients may nevertheless benefit from physicians having a say. Within organizations, administrators and managers can implement many strategies to fan the intrinsic motivation of physicians. Beyond replacing fee-for-service incentives with fixed or panel-based salaries, the role of physician-level financial incentives in these strategies is inherently limited. Because physicians hold themselves and others to high standards and value their reputation, peer comparisons and peer accountability have been advanced as particularly effective motivators. The provision of information on comparative performance alone may strengthen intrinsic incentives to improve and might explain any gains derived from public reporting. The application of these strategies has generally been limited to performance measures targeting specific actions, such as appropriate medication prescribing. Broader nudges could cultivate broader improvement.

Strengths/Limitations: The strategies presented in this article are speculative.

Generalizability to Medicare Population: Strong; the article discusses performance measures which are relevant to Medicare models and programs.

Methods: Performance-based measures review; meta-analysis.

Medicare Payment Advisory Commission. *Chapter 15: Moving beyond the Merit-Based Incentive Payment System (March 2018 Report)*.; 2018. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar18_medpac_ch15_sec.pdf

Subtopic(s): Key Highlights; Challenges Related to Developing and Implementing Performance Measures; Opportunities for APMs and PB-TCOC Models to Address Challenges Related to Developing and Implementing Performance Measures

Type of Source: Report

Objective: To outline for Congress how MIPS is not fulfilling its intended goals and put forth an alternative voluntary value program (VVP) that will fulfill the intended purpose of MIPS more efficiently. The Commission examined options for improving MIPS as it was implemented and provided constructive feedback as CMS established rules for the first two years of the program. However, it was determined after examination that the basic design of MIPS is fundamentally incompatible with the goals of a beneficiary-focused approach to quality measurement.

Main Findings: The basic design principle of MIPS is inequitable because clinicians will be evaluated and compared on dissimilar measures. MIPS scores are not comparable among clinicians because each clinician's composite MIPS score will reflect a mix of different, self-chosen, measures. MIPS imposes a significant reporting burden on clinicians (estimated by CMS as over \$1.3 billion in the first year). In contrast, the commission recommends the VVP, which is based on the premise that patient outcomes rely on the combined contributions of clinicians and emphasizes that quality improvement is a collective effort. A VVP would measure all clinicians based on the same set of measures: clinical quality, patient experience, and value.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the report is related directly to Medicare.

Methods: Evaluation of Medicare MACRA MIPS value-based program.

Medicare Payment Advisory Commission. *June 2014 Report to the Congress: Medicare and the Health Care Delivery System – MedPAC*. MedPAC; 2014. Accessed November 30, 2023.

https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-jun14_entirereport-pdf/

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To examine synchronizing policy across payment models with respect to spending benchmarks, quality measurement, and risk adjustment, and assess synchronizing regulatory oversight.

Main Findings: MPAC examined which approaches to quality measures would be appropriate to each payment model and consider using population-based outcome measures (e.g., potentially avoidable admissions for the FFS population in an area) to evaluate and compare quality within a local area across Medicare's three payment models. Provider-specific quality measures may still be needed for FFS payment adjustments. The authors considered an option to continue additional payments to primary care practitioners, but in the form of a per beneficiary payment.

The current FFS approach encourages volume. A per beneficiary approach could help encourage care coordination.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the report included a Medicare Advantage evaluation.

Methods: An analysis of claims-based data was conducted.

Medicare Payment Advisory Commission. *June 2018 Report to the Congress: Medicare and the Health Care Delivery System – MedPAC*. MedPAC; 2018. Accessed November 29, 2023.

<https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-jun18-medpacreporttocongress-rev-nov2019-note-sec-pdf/>

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To discuss refinements to Medicare payment systems and issues affecting the Medicare program, including broader changes in health care delivery and the market for health care services.

Main Findings: MPAC concluded that the Hospital Readmissions Reduction Program contributed to a significant decline in readmission rates without causing a material increase in emergency department visits or observation stays or an adverse effect on mortality rates. MPAC described a budget-neutral approach to rebalance the fee schedule that would increase payment rates for ambulatory evaluation and management services while reducing payment rates for other services. MPAC applies the principles to the design of a new hospital quality incentive program that combines measures of hospital outcomes, patient experience, and Medicare spending per beneficiary. MPAC found that the fee-for-service coverage process does not prevent the use of low-value services and that the use of such services is prevalent in Medicare. MPAC describes six tools that Medicare could consider for reducing the use of low-value care.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the report includes a recommendation for the Medicare program.

Methods: Claims-based data.

Medicare Payment Advisory Commission. *March 2022 Report to the Congress: Medicare Payment Policy – MedPAC*; 2022. Accessed November 29, 2023. <https://www.medpac.gov/document/march-2022-report-to-the-congress-medicare-payment-policy/>

Subtopic(s): Key Highlights; Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To discuss MPAC's assessment and recommendations of the Medicare program, including the near-term consequences of the coronavirus pandemic and the longer-term effects of program spending on the federal budget and the program's financial sustainability; Medicare Fee-for-Service (FFS) payment policy in 2023; Bipartisan Budget Act (BBA) of 2018 changes to the low-volume hospital payment adjustment; review the status of the Medicare Advantage (MA) program (Medicare Part C); performance of specialized MA plans; Medicare Part D drug coverage; prototype value-based payment program under a unified prospective payment

system (PPS) for post-acute care (PAC) services and analyze the impacts of the prototype's design.

Main Findings: MPAC recommended payment rate updates for nine FFS payment systems for 2023. These include acute care hospitals, physicians and other health professional services, ambulatory surgical centers, outpatient dialysis facilities, skilled nursing facilities, home health agencies, inpatient rehabilitation facilities, long-term care hospitals, and hospice providers. These recommendations are based on reviews of the latest available data. Given Medicare's financing challenges, restraining price growth may not be enough to ensure Medicare's financial sustainability and that the quantity and/or mix of health care services must also be changed. Medicare has piloted several Alternative Payment Models that give providers incentives to more closely manage and coordinate beneficiaries' care to keep them healthy and reduce unnecessary utilization. MPAC has made numerous recommendations that, if implemented, could address challenges with Medicare's payment systems and improve payment accuracy and equity. Medicare's fiscal challenges must be met in a manner that improves quality and reduces inequities in access to care across the Medicare population. Although quality of care appears stable, there is room for improvement.

Strengths/Limitations: Because of standard data lags, the most recent complete data for most payment adequacy indicators are from 2020.

Generalizability to Medicare Population: Strong; this report is an assessment of the Medicare payment system.

Methods: Claims-based data.

Mendelson A, Kondo K, Damberg C, et al. The effects of pay-for-performance programs on health, health care use, and processes of care. *Ann Intern Med.* 2017;166(5):341-353. doi:10.7326/M16-1881

Subtopic(s): Key Highlights; Performance Measures Used in Other Programs

Type of Source: Review

Objective: To update and expand a prior review examining the effects of P4P programs targeted at the physician, group, managerial, or institutional level on process-of-care and patient outcomes in ambulatory and inpatient settings.

Main Findings: The authors analyzed data synthesized from 69 trials and observational studies in ambulatory and inpatient settings reporting process-of-care, health, or utilization outcomes. The studies were collected from PubMed, MEDLINE, PsycINFO, CINAHL, Business Economics and Theory, Business Source Elite, Scopus, Faculty of 1000, and Gartner Research from June 2007 to February 2016. Low-strength evidence suggested that P4P programs in ambulatory settings may improve process-of-care outcomes over the short term (two to three years), whereas data on longer-term effects were limited. Many of the positive studies were conducted in the United Kingdom, where incentives were larger than in the United States. The largest improvements were seen in areas where baseline performance was poor. There was no consistent effect of P4P on intermediate health outcomes (low-strength evidence) and insufficient evidence to characterize any effect on patient health outcomes. In the hospital setting, there was low-strength evidence that P4P had little or no effect on patient health outcomes and a positive effect on reducing hospital readmissions. Pay-for-performance programs may be associated with improved processes of care in ambulatory settings, but consistently positive associations with improved health outcomes have not been demonstrated in any setting.

Strengths/Limitations: Few methodologically rigorous studies; heterogeneous population and program characteristics and incentive targets in the study.

Generalizability to Medicare Population: Moderate; although the study was not focused specifically on Medicare models or programs, findings are applicable to Medicare.

Methods: A literature review and meta-analysis were conducted.

Mentias A, Peterson ED, Keshvani N, et al. Achieving equity in hospital performance assessments using composite race-specific measures of risk-standardized readmission and mortality rates for heart failure. *Circulation*. 2023;147(15):1121-1133. doi:10.1161/CIRCULATIONAHA.122.061995

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To develop a novel approach to evaluate hospital performance among hospitalized patients with heart failure (HF).

Main Findings: First, there was a modest correlation between hospital-level 30-day risk standardized readmission rate (RSRR) and risk-standardized mortality rate (RSMR) for White patients compared with Black patients and patients of other races hospitalized with HF. Second, compared with race-agnostic performance metrics, composite race-specific RSMR and RSRR for all patients had a stronger correlation with the corresponding performance metrics for Black patients. Third, only ~50 percent of hospitals identified as high- or low-performing using race-agnostic RSRR and RSMR for all patients were concordantly high- or low-performing for patients of Black race. In contrast, the concordance in hospital performance for all patients compared with patients of Black race was significantly greater using composite race-specific performance metrics. Finally, compared with race-agnostic 30-day RSRR and RSMR, composite race-specific performance metrics led to a meaningful reclassification of 36 percent and 39 percent hospitals, respectively, with better 30-day and one-year outcomes in patients of all race and ethnicity groups among up-classified hospitals.

Strengths/Limitations: Future studies are needed to evaluate whether race-specific approaches to assessing hospital performance may be associated with more equitable gains in care quality and outcomes across all races.

Generalizability to Medicare Population: Strong; the sample included fee-for-service Medicare beneficiaries.

Methods: Claims data; hospitalization data; utilization data.

Naessens JM, Van Such MB, Nesse RE, et al. Looking Under the Streetlight? A Framework for Differentiating Performance Measures by Level of Care in a Value-Based Payment Environment. *Acad Med*. 2017;92(7):943-950. doi:10.1097/ACM.0000000000001654

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To describe a value measurement and payment framework for specialty care.

Main Findings: Measures should vary across different levels of care (primary/community care, secondary care, complex care), and risk adjustment should be used to reduce referral bias. Performance measures and reimbursement schemes should account for the increased intensity required in providing complex care. The authors propose a new model that would define the characteristics of patients requiring complex care and standardize metrics based on these definitions to improve performance measurement in specialty care.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Moderate; the authors discuss value measurement in several Medicare programs, and the approaches outlined in the paper are applicable within the context of Medicare.

Methods: N/A

National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Health Care Utilization and Adults with Disabilities. *Health-Care Utilization as a Proxy in Disability Determination*. Washington (DC): National Academies Press (US); 2018 Mar 1. 2, Factors That Affect Health-Care Utilization. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK500097/>

Subtopic(s): Background

Type of Source: Report

Objective: To identify health care utilization metrics that can serve as indicators of impairment severity for use by the disability program.

Main Findings: The committee did not identify an association between health care utilizations (e.g., inpatient hospitalizations, emergency department use, and hospital readmissions) and the Social Security Administration's (SSA's) definition of listing-level impairment severity. The SSA's definition of impairment severity goes beyond disease severity, requiring conditions be severe enough to prevent an individual from engaging in any gainful activity. Existing literature suggests that health care utilizations may be linked with disease severity, but there is no research linking utilizations with SSA's definition of impairment severity.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; the report focuses on metrics for the Social Security Disability Insurance (SSDI) program and the Supplemental Security Income (SSI) program.

Methods: A committee of 16 experts conducted a targeted literature search and convened five times to reach consensus about relevant metrics.

National Quality Forum. *Attribution - Principles and Approaches*. NQF; 2016. Accessed November 30, 2023. [https://www.qualityforum.org/Publications/2016/12/Attribution - Principles and Approaches.aspx](https://www.qualityforum.org/Publications/2016/12/Attribution_-_Principles_and_Approaches.aspx)

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Report

Objective: To determine the necessary elements of an attribution model that must be specified and identify recommendations for those developing and implementing attribution models.

Main Findings: At a minimum, accountable entities should consider the context and goal of the program, patient inclusion/exclusion criteria, eligible units, and attribution methodology during the selection and development of an attribution model. The committee also recommended testing and multi-stakeholder review of attribution models. In addition, accountable entities should be able to influence care and outcomes, and attribution models used by certain programs should have to meet minimum criteria.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the principles and approaches discussed are highly relevant to Medicare, and the report includes several specific examples of attribution in Medicare payment models.

Methods: A multi-stakeholder expert committee convened and conducted an environmental scan of current attribution models.

National Quality Forum. *Measure Evaluation Criteria*. NQF: n.d. Accessed November 17, 2023. https://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Webpage

Objective: To describe the measure evaluation criteria from the National Quality Forum's 2012 projects on patient-reported outcomes, composites, and eMeasure feasibility.

Main Findings: NQF considered and evaluated measures for suitability as voluntary consensus standards. Measures were evaluated based on their 1) importance to measure and report; 2) scientific acceptability of measure properties; 3) feasibility; 4) usability and use; and 5) related and competing measures. Evaluation of criteria occurred in sequential order. Assessment for each criterion was based on a continuous scale, rather than a binary "meets" or "does not meet" determination. However, if a measure did not meet the minimum threshold for a criterion, the measure was not evaluated against the remaining criteria.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; the webpage does not discuss Medicare-specific measures.

Methods: N/A

National Quality Forum. *Social Risk Trial Final Report*. NQF; 2017. Accessed November 29, 2023. https://www.qualityforum.org/Publications/2017/07/Social_Risk_Trial_Final_Report.aspx

Subtopic(s): Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To evaluate the NQF trial period for risk adjustment for social risk factors.

Main Findings: Over 300 measures were submitted for review by the NQF committee, but only 65 outcome measures had a conceptual basis for adjustment for social risk factors. Generally, model performance was not improved by including social risk-adjusted measures in risk-adjustment models. Developers noted challenges obtaining individual-level data on social risk factors. Nonetheless, NQF committee members and measure developers emphasized the importance of addressing both clinical and social factors through performance measures.

Strengths/Limitations: Data on social risk factors were limited; however, including external methodologists in the evaluation was a strength.

Generalizability to Medicare Population: Strong; several Medicare measures were evaluated, and findings on use of risk-adjusted measures are relevant to Medicare.

Methods: During a two-year trial period, NQF allowed inclusion of social risk factors in its risk-adjustment models. The NQF committee evaluated and tracked measures submitted during the trial period. In addition, measure developers and NQF committee members provided feedback on use of risk-adjusted outcome measures via open-ended text response surveys.

NEJM Catalyst. What Is Pay for Performance in Healthcare? *Catalyst Carryover*. 2018;4(2). doi:10.1056/CAT.18.0245

Subtopic(s): Key Highlights; Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To describe pay-for-performance approaches, pay-for-performance models, CMS programs, and pros and cons.

Main Findings: CMS has developed several pay-for-performance models and programs, including the Hospital Value-Based Purchasing Program, the Hospital Readmissions Reduction Program, the Hospital-Acquired Condition Reduction Program, the End-Stage Renal Disease Quality Initiative Program, the Skilled Nursing Facility Value-Based Program, the Home Health Value-Based Program, and the Value Modifier or Value-Based Modifier Program. Benefits of pay-for-performance include emphasis on quality over quantity of care, use of transparent metrics that improve accountability, use of existing FFS payment systems, and proven cost savings for some programs. Criticisms of pay-for-performance include reductions in access for disadvantaged populations, reductions in job satisfaction, requirements for expensive administrative systems, and challenges related to accurate provider attribution.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article discusses Medicare pay-for-performance programs in-depth.

Methods: N/A

NORC at the University of Chicago. *Fifth Evaluation Report: Next Generation Accountable Care Organization Model Evaluation*. NORC at the University of Chicago; 2022:44.

<https://www.cms.gov/priorities/innovation/data-and-reports/2022/nextgenaco-fifthevalrpt>

Subtopic(s): Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To report on program year 5 of the NGACO advanced Alternative Payment Model, changes made in response to COVID-19, as well as impacts on gross and net overall spending and its impacts on specific categories of Medicare spending and utilization.

Main Findings: NGACO was associated with a 1.5 percent reduction in spending and increased cumulative net Medicare spending. Participating organizations found that partnerships and resources developed during their time in the model allowed them to better respond to needs during the COVID-19 pandemic. These organizations reduced hospital spending and utilization, SNF stays and days and spending in institutional PAC settings.

Strengths/Limitations: COVID-19 severely impacted the model and resulted in significant changes to accommodate the needs during the pandemic. However, participating organizations found that the work done in previous program years allowed them to improve their ability to respond to the needs of their patients.

Generalizability to Medicare Population: Strong; this is an evaluation of a model that was developed through the Center for Medicare and Medicaid Innovation fund.

Methods: Difference-in-differences framework was used to estimate differential changes in spending and utilization between the baseline year and each program year.

NORC at the University of Chicago. *Third Evaluation Report: Next Generation Accountable Care Organization Model Evaluation*. NORC at the University of Chicago; 2020:137.

<https://innovation.cms.gov/data-and-reports/2020/nextgenaco-thirdevalrpt-fullreport>

Subtopic(s): Key Highlights; Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To evaluate the NGACO Model through performance year three (end of 2018).

Main Findings: Across the first three performance years, gross Medicare expenditures decreased; however, net Medicare spending did not decrease. Cumulative net and gross spending patterns differed across cohort years with the 2016 cohort demonstrating the highest net spending increase and the 2017 cohort demonstrated the greatest reduction in gross spending. In its first year, the 2018 cohort had statistically significant reductions in gross spending. With respect to spending in the third performance year specifically, NGACOs decreased gross spending but did not reduce net spending. Additionally, the effect size of the model-wide reduction in gross spending in PY3 was larger than the gross spending reduction in PY2. Regarding utilization, there were no observed model-wide reductions in acute care hospital spending, though there was a 12 percent increase in annual wellness visits across NGACOs. There were no significant changes in quality of care measures detected in PY3 or cumulatively.

Strengths/Limitations: The evaluation draws on both quantitative and qualitative methods and effectively synthesizes findings from these different methods. Additionally, the model employs a difference-in-differences design, which is an effective model for assessing causal relationships between the model and observed outcomes. The evaluation notes that in future reports, researchers plan to further categorize NGACOs according to their care management/coordination/delivery and risk stratification approaches so as to better isolate organizational and structural characteristics associated with improved outcomes. Additionally, the evaluation captures only the first three performance years; some outcomes may take more time to see changes.

Generalizability to Medicare Population: Strong; the model evaluated directly served Medicare beneficiaries and providers.

Methods: The evaluation used both quantitative and qualitative methods, including regression modeling such as difference-in-differences modeling to assess causal effects of the model, qualitative comparative analysis to examine NGACOs' contextual and structural pathways to reduce Medicare spending, interviews with ACO leaders, and surveys with NGACO leadership and affiliated physicians.

Nouri SS, Adler-Milstein J, Thao C, et al. Patient characteristics associated with objective measures of digital health tool use in the United States: A literature review. *Journal of the American Medical Informatics Association*. 2020;27(5):834-841. doi:10.1093/jamia/ocaa024

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Journal Article

Objective: To identify patient characteristics linked to the utilization of patient-facing digital health tools in the United States.

Main Findings: The analysis included 29 studies, with 13 deemed to have robust methodology. Most studies focused on smartphone apps and text messaging programs for chronic disease management, primarily assessing 1-3 patient characteristics such as age and gender. Overall, there was no consistent association between patient characteristics and tool usage. However, in studies with robust methodology, higher use was associated with white race and poor health status.

Strengths/Limitations: Limited studies assess objective measures of digital health tool use based on patient characteristics.

Generalizability to Medicare Population: Weak; the article does not discuss Medicare payment.

Methods: A literature review was conducted.

Ogundeji YK, Bland JM, Sheldon TA. The effectiveness of payment for performance in health care: A meta-analysis and exploration of variation in outcomes. *Health Policy*. 2016;120(10):1141-1150. doi:10.1016/j.healthpol.2016.09.002

Subtopic(s): Performance Measures Used in Other Models

Type of Source: Journal Article

Objective: To evaluate the effectiveness of pay-for-performance schemes on health outcomes.

Main Findings: A total of 37 studies were included in the meta-analysis and meta-regression. When evaluating pay-for-performance schemes based on processes rather than outcomes, larger effects were observed. Larger incentives tended to have greater impacts compared with smaller incentives. Pay-for-performance schemes that underwent more thorough evaluations demonstrated smaller estimated effects.

Strengths/Limitations: This is the first meta-analysis that examined how design features and evaluation methods influence the effectiveness of pay-for-performance schemes. However, the authors noted inadequate reporting of effect estimates, sample sizes, and standard errors in 60 studies, which prevented their inclusion in the meta-analysis.

Generalizability to Medicare Population: Moderate; although the paper does not reference Medicare specifically, the findings on pay-for-performance schemes are applicable within the context of Medicare.

Methods: A systematic review and meta-analysis were conducted.

Petersen LA, Simpson K, Pietz K, et al. Effects of Individual Physician-Level and Practice-Level Financial Incentives on Hypertension Care: A Randomized Trial. *JAMA*. 2013;310(10):1042-1050. doi:10.1001/jama.2013.276303

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To assess the impact of explicit financial incentives on provision of guideline-recommended hypertension care.

Main Findings: Physicians in the individual incentive group were more likely than controls to achieve blood pressure control and make adjustments to anti-hypertensive medications in response to uncontrolled blood pressure. The use of guideline-recommended medications did not differ between intervention and control groups. Participants in the intervention groups were more likely to view their performance reports. However, the effect of financial incentives was not sustained after the incentive was withdrawn. There was not a higher incidence of hypotension in the panels of physicians randomized to the incentive groups.

Strengths/Limitations: The sample included Veterans Affairs enrollees who tend to be male. The larger proportion of male enrollees could impact the generalizability of findings.

Generalizability to Medicare Population: Moderate; the paper does not discuss Medicare payment in-depth; however, the findings on financial incentives are applicable within the context of Medicare.

Methods: A cluster randomized trial was conducted.

Pirritano M, Miller Parrish K, Kim Y, Solomon H, Keene J. It takes quality improvement to cross the chasm. *BMJ Open Qual*. 2023;12(3):e001906. doi:10.1136/bmjopen-2022-001906

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To evaluate the effect of a comprehensive quality improvement program on health outcomes in a large county-based Medicaid health plan.

Main Findings: There were significant improvements in various HEDIS metrics following the implementation of the quality improvement program, including prenatal and postpartum measures. The majority of improvement occurred on the lower end of metrics, close to the thresholds. Opportunities for improvement included using an online reporting portal and streamlining data source formats.

Strengths/Limitations: This study did not evaluate return on investment, and longer-term research is needed to confirm the results of this evaluation.

Generalizability to Medicare Population: Moderate; the study discusses the effects of a quality improvement program on Medicaid beneficiaries. However, findings are applicable to CMS programs, including Medicare, more broadly.

Methods: Regression discontinuity design.

Qaseem A, Mount CA, Campos K, et al. Quality Indicators for Osteoporosis in Adults: A Review of Performance Measures by the American College of Physicians. *Ann Intern Med.* 2023;176(10):1386-1391. doi:10.7326/M23-1291

Subtopic(s): Consumer Assessment of Healthcare Providers and Systems Survey

Type of Source: Journal Article

Objective: To review existing performance measures related to osteoporosis.

Main Findings: Six osteoporosis performance measures relevant to internal medicine physicians were identified. However, only one of these measures was deemed valid across all levels of attribution (individual physician, group practice, health plan). Despite the limitations of existing osteoporosis performance measures, MIPS uses these measures to incentivize physicians who provide care to Medicare patients.

Strengths/Limitations: Measures were reviewed American College of Physicians (ACP) Performance Measurement Committee (PMC), which includes physicians from across the United States, working in a variety of settings, with expertise in performance measurement.

Generalizability to Medicare Population: Strong; this study reviewed measures from CMS Measure Inventory Tool and CMS Quality Payment Program. The study also discussed Medicare programs such as MIPS.

Methods: The ACP PMC reviewed osteoporosis performance measures using a modified RAND/UCLA Appropriateness Method. Measures were evaluated on five domains: importance, appropriate care, clinical evidence base, measure specifications, and feasibility and applicability. Following a moderated committee discussion, 11 committee members voted on each measure.

RAND Health Care. *Developing Health Equity Measures.* The Office of the Assistant Secretary for Planning and Evaluation (ASPE) at the U.S. Department of Health and Human Services. 2021. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//200651/developing-health-equity-measures.pdf

Subtopic(s): Background

Type of Source: Report

Objective: To identify and evaluate existing approaches to measuring health equity and determine which, if any, merited consideration for inclusion in Medicare's Value-Based Purchasing (VBP) programs.

Main Findings: Three programs were found to meet the 10 approaches developed by the Technical Expert Panel to evaluate health equity measures. The National Quality Forum (NQF) Disparities-Sensitive Measure Assessment was determined to have the most favorable approach for measure identification; the Minnesota Healthcare Disparities Report had the most favorable approach for measure-by-measure comparisons; and the Centers for Medicare & Medicaid Services Office of Minority Health's Health Equity Summary Score (CMS OMH HESS) approach was most favorable for summary indices. Overall, the HESS was deemed to be the best approach, coming the closest to meeting the full scope of goals outlined by ASPE to incorporate a measure of health equity into a Medicare VBP or quality reporting program.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the developed measures are intended for use by Medicare's VBP programs to improve health equity.

Methods: A formal definition of a health equity measure was developed to guide a thorough literature review which resulted in 11 articles and reports that were selected as fitting the eligibility criteria which was to exclude any reports that (1) did not describe a specific health equity measurement approach developed or (2) were focused on risk adjustment.

Ratitch B, Trigg A, Majumder M, Vlajnic V, Rethemeier N, Nkulikiyinka R. Clinical Validation of Novel Digital Measures: Statistical Methods for Reliability Evaluation. *Digital Biomarkers*. 2023;7(1):74-91. doi:10.1159/000531054

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Journal article

Objective: To describe statistical methodologies and analytical tools for reliability assessment.

Main Findings: Most literature that describes methodologies for reliability analyses applies to the validation of patient-reported outcomes (PROs), wet biomarkers, and other clinical measures, rather than digital clinical measures. The authors discuss a general modeling framework and statistical metrics commonly used for reliability assessments in clinical validation. The authors also provide an example using actigraphy-based digital measures of physical activity in heart failure patients. The example emphasized several considerations for reliability analyses, including availability of data, critical pathways of interest, and repeated measurements.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; the article does not discuss Medicare payment.

Methods: N/A

Rawal P, Jacobs D, Fowler E, Seshamani M. Building On CMS's Accountable Care Vision To Improve Care For Medicare Beneficiaries. *Health Affairs Forefront*. doi:10.1377/forefront.20230727.802728

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To describe advancements made in CMS accountable care strategy and identify opportunities for future improvement in CMS ACOs.

Main Findings: CMS supports care delivery redesign for physician practices, aiming to improve quality, outcomes, and care experiences for beneficiaries. The 2022 CMS accountable care strategy created new model tests to increase access to high-quality, integrated, and coordinated care for beneficiaries. For example, CMS updated the Shared Savings Program policies and

successfully executed ACO REACH. The ultimate objective is to ensure that all traditional Medicare beneficiaries are in accountable care relationships by 2030.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article is about improving care for Medicare beneficiaries.

Methods: N/A

Reeves S, Dombkowski K, Madden B, et al. Considerations when Aggregating Data to Measure Performance across Levels of the Healthcare System. *Acad Pediatr.* 2022;22(3 Suppl):S119-S124. doi:10.1016/j.acap.2021.11.013

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To compare two different attribution models and the effect of the attribution models by population characteristics.

Main Findings: In the first case study, the proportion of children with sickle cell anemia who were not attributed to a Medicaid health plan ranged from 12.2 percent to 89.0 percent across years. In the second case study, 22 percent of patients were not attributed to a health plan for the 30 days post-discharge. The authors highlighted that when selecting attribution models, it is crucial to consider the potential to induce health disparities. Differential attribution can have unintended consequences on health disparities.

Strengths/Limitations: The study used simulated ACO data to study rare pediatric events, for which real data are limited.

Generalizability to Medicare Population: Moderate; although the study analyzes Medicaid data, considerations for aggregating data are relevant to Medicare as well.

Methods: This study conducted two case studies to illustrate the use of two different methods of attribution using real datasets. The authors compared differences in the proportion of patients that could not be attributed between the two methods.

Report of the Council on Medical Service, Improving Risk Adjustment in Alternative Payment Models, CMS Report 3-1-19. 2019. <https://www.ama-assn.org/system/files/2019-12/i19-cms-report3-apms.pdf>

Subtopic(s): Risk Adjustment Methods Used for PB-TCOC Models

Type of Source: Report

Objective: To describe background on risk adjustment and identify policy recommendations for the American Medical Association (AMA) to improve risk adjustment.

Main Findings: The report identified 8 recommendations for AMA, including that AMA should work with CMS to identify better data sources and advocate for less burdensome documentation requirements. The report also recommended the use of risk stratification systems that adjust payments based on patient characteristics, offer outlier payments for high-cost individual patients, use risk corridors, and consider external price changes beyond the physician's control. In addition, the report recommended that physicians not be held accountable for patients' outcomes that they do not influence, and risk adjustment mechanisms should be flexible to allow for changes in science and practice.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the report focuses on risk adjustment for Medicare Alternative Payment Models.

Methods: N/A

Revere L, Langland-Orban B, Large JT, Yang Y. Evaluating the robustness of the CMS Hospital Value-Based Purchasing measurement system. *Health Services Research*. 2021;56(3):464-473. doi:10.1111/1475-6773.13608

Subtopic(s): Appendix E. Summary of Current Performance Measures for Hospital VBP, MA Star Ratings Program, MIPS, and MSSP

Type of Source: Journal article

Objective: To explore variation in overall and individual-hospital total performance score (TPS) and embedded domains for hospitals during 2014-2018.

Main Findings: The annual TPS score shows a skewed distribution, with a substantial gap between top-performing and other hospitals. Yearly TPS results demonstrate significant fluctuations, with hospitals moving in and out of the top and bottom performance categories. Hospitals categorized in the bottom 5 percent in 2014 did not attain a top 5 percent rank throughout the study period.

Strengths/Limitations: The study sample size was small.

Generalizability to Medicare Population: Strong; the study evaluates the measurement system of a Medicare program, the Hospital Value-Based Purchasing Program.

Methods: Descriptive statistics and transitional analysis to evaluate annual movement in the TPS ranking for outlier hospitals in the top and bottom 5 percent.

Rhee C, Strich JR, Chiotos K, et al. Improving Sepsis Outcomes in the Era of Pay-for-Performance and Electronic Quality Measures: A Joint IDSA/ACEP/PIDS/SHEA/SHM/SIDP Position Paper. *Clin Infect Dis*. 2023:ciad447. doi:10.1093/cid/ciad447

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Journal article

Objective: To describe concerns with changing the Severe Sepsis/Septic Shock Management Bundle (SEP-1) from a pay-for-reporting measure to a pay-for-performance measure in the Hospital Value-Based Purchasing Program.

Main Findings: Existing research does not suggest an association between SEP-1 implementation and decreased mortality. The authors recommend retiring SEP-1 and shifting to new sepsis metrics that focus on patient outcomes, such as the community-onset sepsis 30-day mortality electronic clinical quality measure (eCQM). However, the eCQM should be combined with the CDC Adult Sepsis Event surveillance metric to reduce hospital reporting burden.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the paper discusses concerns related to a CMS measure used in the Hospital Value-Based Purchasing Program, a Medicare program.

Methods: This is a position paper drafted by members of a task force, including representatives from several medical professional groups.

Richard C, Urick BY, Pathak S, Jackson J, Livet M. Performance-based pharmacy payment models: Key components and critical implementation considerations for successful uptake and integration. *J Manag Care Spec Pharm*. 2021;27(11). doi:10.18553/jmcp.2021.27.11.1568

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To recommend design improvements that increase transparency, alignment of incentive structures, and improve patient care performance measure within performance-based pharmacy payment models (PBPPMs).

Main Findings: Attribution, performance and quality measures, incentive structures, and patient care services were identified as the four major components of PVPPMs. To improve implementation of PBPPMs, the study recommends that innovative business models be embraced, a culture of quality is fostered throughout health care, transparency is increased within the incentive structure, and roadmaps for successful uptake and implementation must be carefully planned and followed. Building critical infrastructure such as data sharing tools is critical in facilitating the implementation of said recommendations.

Strengths/Limitations: Outcomes and effectiveness of PBPPMs were not studied. The sample size was limited, and interviewees were pharmacists on the independent and regional level, meaning retail pharmacists were not included.

Generalizability to Medicare Population: Moderate; Medicare beneficiaries were considered in the analysis; however, this was not the central focus of the study.

Methods: A qualitative study consisting of an environmental scan and stakeholder interviews was conducted. The first author conducted a content analysis of the interview transcripts and created a coding structure based on categories of interest.

Richardson A, Robbins CB, Wisely CE, Heno R, Grewal DS, Fekrat S. Artificial intelligence in dementia. *Curr Opin Ophthalmol.* 2022;33(5):425-431. doi:10.1097/ICU.0000000000000881

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To review current trends in the use of artificial intelligence tools in clinical environments to anticipate possible incorporation in dementia diagnostic paradigms.

Main Findings: Alzheimer diagnostic efforts may be simplified if machine learning models can be incorporated into clinical care. These models may be able to distinguish those with symptomatic Alzheimer's dementia from patients with mild cognitive impairment and normal cognition. The models may also be able to predict progressive disease.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Low; article not directly related to the Medicare population.

Methods: Literature review conducted on emerging trends in artificial intelligence tools use in clinical settings.

Rivera SC, Liu X, Hughes SE, et al. Embedding patient-reported outcomes at the heart of artificial intelligence health-care technologies. *The Lancet Digital Health.* 2023;5(3):e168-e173. doi:10.1016/S2589-7500(22)00252-7

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To analyze the current and potential use of artificial intelligence health care technologies in patient-reported outcomes measures.

Main Findings: Artificial intelligence health care technologies have the potential to promote survival over wellbeing. Careful consideration must be used to integrate patient-reported outcomes in AI health care to avoid. Integration PROs and AI could help ensure that patient voices are being heard as health care becomes more automated with AI.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Weak; Medicare is not specially referenced or considered in this article.

Methods: Authors analyzed relevant literature to determine the current use of AI in health care and PROs, and theorize what embedding the two could look like.

RTI. *ACO Realizing Equity, Access, and Community Health (REACH) Model - PY2024 Quality Measurement Methodology*. Centers for Medicare & Medicaid Services; n.d. Accessed December 1, 2023.

<https://www.cms.gov/files/document/aco-reach-quality-msr-meth-py24.pdf>

Subtopic(s): Key Highlights; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To provide an overview of quality measurement and performance evaluation methodology for ACOs participating in the ACO REACH Model.

Main Findings: The ACO REACH Model makes three important changes to the GPDC Model: advancing health equity, promoting provider leadership and governance, and protecting beneficiaries. CMS uses Quality Performance Benchmarks (QPBs) for high-needs populations ACOS and standard and new entrant ACOs. The goal of the ACO REACH Model is to lower the cost of care for Medicare beneficiaries while maintaining or improving the quality of care provided.

Strengths/Limitations: This document is updated as new relevant policies become applicable.

Generalizability to Medicare Population: Strong; Medicare ACO initiatives are central to this report.

Methods: Scoring systems and benchmarks are used to assess quality performance of ACOs participating in the model.

Ryan AM, Markovitz AA. Estimated savings from the Medicare Shared Savings Program. *JAMA Health Forum*; 2023;4(12): e234449-e234449. doi:10.1001/jamahealthforum.2023.4449

Subtopic(s): Performance Measures Used in the Medicare Shared Saving Program

Type of Source: Journal article

Objective: To examine the budgetary impact of the Medicare Shared Savings Program (MSSP) to the Centers for Medicare & Medicaid Services (CMS) between 2013 and 2021.

Main Findings: The MSSP was associated with overall net savings for CMS, including Medicare Advantage. However, the MSSP was associated with net losses to traditional Medicare. Despite potential savings overall, the total estimated budgetary impact of the MSSP to CMS was small.

Strengths/Limitations: One limitation was the amount of missing data on hospital affiliation for approximately 12 percent of the beneficiaries. In addition, administrative or opportunity costs of the MSSP could not be accounted for, potentially resulting in an underestimate of the total costs to CMS.

Generalizability to Medicare Population: Strong; this article evaluated the budgetary impact of the MSSP to CMS.

Methods: Publicly available data on the MSSP, as well as data extracted from two studies were analyzed.

Salinas J, Sprinkhuizen SM, Ackerson T, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. *Stroke*. 2016;47(1):180-186. doi:10.1161/STROKEAHA.115.010898

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To define an international standard set of patient-centered stroke outcome measures.

Main Findings: Patients presenting with ischemic stroke or intracerebral hemorrhage were selected as the target population for recommendations. Survival and disease control, acute complications, and patient-reported health status were the three most prevalent outcomes and measures selected for the standard set. The implementation of a standard set has the potential to help inform the delivery of effective, equitable, patient-centered stroke care.

Strengths/Limitations: The efficacy of the standard set in practice has not been evaluated.

Generalizability to Medicare Population: Low; Medicare populations are not referenced in this article.

Methods: A modified Delphi process was utilized, convening an international expert panel representing patients, advocates, and clinical specialists in stroke outcomes.

Sankaran R, Kumar A, Parasuram H. Role of Artificial Intelligence and Machine Learning in the prediction of the pain: A scoping systematic review. *Proc Inst Mech Eng H*. 2022;236(10):1478-1491. doi:10.1177/09544119221122012

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To conduct a review of the role of artificial intelligence and machine learning in diagnostics and treatment management.

Main Findings: Artificial intelligence and machine learning have the potential to play a role in pain prediction for patients. AI/ML can differentiate between health controls and people experiencing pain. This ability could be used to compare subjective and objective measures of pain.

Strengths/Limitations: There are not many other similar reviews of this nature to compare to.

Generalizability to Medicare Population: Low; Medicare populations are not referenced in this article.

Methods: Literature review was conducted from the inception of artificial intelligence and machine learning techniques until 2021.

Shirley ED, Sanders JO. Measuring Quality of Care with Patient Satisfaction Scores. *J Bone Joint Surg Am*. 2016;98(19):e83. doi:10.2106/JBJS.15.01216

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To investigate preferred ways to measure quality of care.

Main Findings: Satisfaction surveys may not be an appropriate measure of overall quality of care. To accurately measure quality of care from a patient's perspective, a multidimensional approach is needed. The use of satisfaction scores that include specific measures for each domain is recommended.

Strengths/Limitations: Patient satisfaction is a complex construct with varying factors that are hard to assess.

Generalizability to Medicare Population: Low; Medicare populations are not referenced in this article, though this content may affect them.

Methods: Investigation of quality-of-care methods using existing literature.

Simon L, Robert C, Meyer P. Artificial intelligence for quality assurance in radiotherapy. *Cancer Radiother*. 2021;25(6-7):623-626. doi:10.1016/j.canrad.2021.06.012

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To analyze what contributions artificial intelligence could bring to patient-specific quality assurance and machine quality assurance.

Main Findings: Using artificial intelligence in patient-specific quality assurance and machine quality assurance is in very early stages. Using AI in radiotherapy quality assurance takes advanced skills and algorithms, and there are currently no commercial options.

Strengths/Limitations: This article is specific to France and guidelines from the European Federation of Organizations for Medical Physics; the article does not look at implementation or guidelines in America. Additionally, the article is originally written in French; therefore there could be translation barriers.

Generalizability to Medicare Population: Low; the Medicare population was not considered in this journal article, and the article relates to the population of France.

Methods: A literature review was conducted to determine AI use in patient-specific and machine quality assurance.

Skopec L, Berenson RA. The Medicare Advantage Quality Bonus Program: High Cost for Uncertain Gain. *Urban Institute*; 2023. Accessed November 30, 2023.

<https://www.urban.org/research/publication/medicare-advantage-quality-bonus-program>

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Report

Objective: To describe the Medicare Advantage (MA) quality bonus program (QBP), Star Ratings measures, and issues with the QBP with potential solutions.

Main Findings: There are several issues with the Star Rating system and the QBP, such as score inflation and limitations with underlying measure sets. A large proportion of a contract's Star Rating is determined by beneficiary experiences with the health plan and providers and administrative effectiveness rather than clinical quality. The MA QBP should be modified so that the program encourages MA organizations to improve quality of care and helps beneficiaries make informed decisions when selecting health plans.

Strengths/Limitations: Some analyses could have been impacted by challenges with data collection due to the COVID-19 pandemic.

Generalizability to Medicare Population: Strong; the article identifies and describes the strengths and limitations of a Medicare program.

Methods: A literature review of published literature was conducted to identify issues with the QBP, as well as potential solutions to the issues. In addition, an analysis of 2023 MA Star Ratings data was conducted, specifically focusing on measure technical specifications and weights, performance and Star Ratings for each measure, and overall Star Ratings by MA contract. MA plan enrollment data were analyzed to understand the geographic reach of MA contracts, as well as the proportion of MA enrollment in contracts with varying levels of Star Ratings by measure.

Skopec L, Bowen G, Zuckerman S. Accounting for Social Risk in Value-Based Payment and Quality Measurement. *Urban Institute*; 2022. Accessed December 1, 2023.

<https://www.urban.org/research/publication/accounting-social-risk-value-based-payment-and-quality-measurement>

Subtopic(s): Risk Adjustment Methods Used for PB-TCOC Models; Best Practices for Measuring Performance; Opportunities to Improve Performance Measures

Type of Source: Report

Objective: To recommend approaches to improve health equity throughout Medicare, Medicaid, and VBP models.

Main Findings: To preserve access to health care for all beneficiaries, adjusting payments and quality measures for social risks may be necessary. This, however, is not sufficient to improve equity. Reinforcing preexisting inequities should be avoided in creating new approaches. Incentivizing providers to improve quality of care and reduce disparities differs from investing directly in health equity resources, though both may be necessary.

Strengths/Limitations: More research is needed on strategies to effectively promote value and equity in payment and delivery systems.

Generalizability to Medicare Population: Strong; research questions directly relate to improving health equity for Medicare populations.

Methods: A panel of national experts was convened to identify policy and technical issues to help CMS consider changes to Medicare, Medicaid, and VBP models.

Swankoski K, O'Malley A, Tu, H et al. *Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Fourth Annual Report*. Mathematica; 2022:200.

<https://www.cms.gov/priorities/innovation/data-and-reports/2022/cpc-plus-fourth-annual-eval-report>

Subtopic(s): Performance Measures Used in CMMI Models

Type of Source: Report

Objective: To examine how CPC+ has been implemented and its impact on Medicare FFS beneficiaries in regions that joined in 2017.

Main Findings: Over the first four years, CPC+ reduced key measures of acute care utilization and expenditures on acute inpatient care but did not reduce total Medicare expenditures across Track 1 and Track 2 overall. CPC+ also led to small improvements in some quality-of-care measures. The impact findings are largely consistent with findings from other studies of primary care interventions, which found mixed results from practice transformation initiatives. The reductions in acute care utilization are consistent with the time path of the theory of change for CPC+ and findings from the evaluation of the implementation of CPC+ at the practice level. While CPC+ has not shown reductions in total expenditures in the first four years across Track 1 and Track 2 overall, the reductions in hospitalizations and accompanying reduction in expenditures on acute inpatient care in PY 3 and PY 4 are promising.

Strengths/Limitations: The authors did not report on the regions that joined CPC+ in 2018 that accounted for 5 percent of the total number of practices participating in CPC+, and the first-year implementation experiences of practices and payers in the regions that joined CPC+ in 2018 were very similar to the first-year experiences of those that joined CPC+ in 2017.

Generalizability to Medicare Population: Strong; the report evaluated a CMS model.

Methods: Claims-based data; tracking of four-year model impacts; participation rosters; surveys; vendor data.

Taylor KI, Staunton H, Lipsmeier F, Nobbs D, Lindemann M. Outcome measures based on digital health technology sensor data: Data- and patient-centric approaches. *npj Digit Med*. 2020;3(1):1-8.

doi:10.1038/s41746-020-0305-8

Subtopic(s): Assessment of the Different Types of Data Sources Used for Construction of Performance Measures

Type of Source: Journal article

Objective: To propose a path to developing outcome measures based on digital health technology tools (DHTT) data.

Main Findings: Two roadmaps to developing DHTT outcome measures are proposed, both focusing on measures of disease progression for use in longitudinal clinical research studies. The first approach is data-centric and maximizes sensitivity to detect disease progression. The second approach is patient-centric and measures progression of symptoms patients consider most meaningful.

Strengths/Limitations: Some DHTT data are missing as they come from disease-relevant factors, which could bias the outcome measure.

Generalizability to Medicare Population: Low; the Medicare population is not specifically examined or considered.

Methods: DHTT data summarized into meaningful outcome measures.

Tew M, Dalziel K, Clarke P, Smith A, Choong PF, Dowsey M. Patient-reported outcome measures (PROMs): Can they be used to guide patient-centered care and optimize outcomes in total knee replacement? *Qual Life Res.* 2020;29:3273-3283. doi:10.1007/s11136-020-02577-4

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal

Objective: To examine health-related quality-of-life (QoL) trajectories and associated patient characteristics in a five-year span post total knee replacement (TKR).

Main Findings: Six unique QoL trajectories were identified: patients with the lowest gains from TKR; patients with moderate sustained gains; patients exhibiting slow progressive improvement; patients with large, sustained improvement; patients whose improvement peaked early; and patients consistently reporting high QoL. Younger patients with co-morbidities reporting greater pain at pre-surgery were most likely to be in the most positive QoL trajectory. 18.4 percent of patients were identified in said most positive trajectory.

Strengths/Limitations: Patients were from a single center, which could be limiting.

Generalizability to Medicare Population: Low; Medicare populations are not specially acknowledged or considered.

Methods: A single-institution registry was used to pull data on patients who underwent TKRs from 2006 to 2011. QoL trajectories were modeled using latent class growth analysis. Multinomial logistic regression analyses were conducted to examine association between baseline patient characteristics and trajectory groups.

Valuck TB, Sampsel S, Sloan DM, Van Meter J. Improving Quality Measure Maintenance: Navigating the Complexities of Evolving Evidence. *The American Journal of Managed Care.* 2019;25(6):e188-e191. Accessed November 28, 2023. <https://www.ajmc.com/view/improving-quality-measure-maintenance-navigating-the-complexities-of-evolving-evidence>

Subtopic(s): Opportunities to Improve Performance Measures

Type of Source: Journal article

Objective: To encourage collaboration between quality measurement stakeholders and advance shared responsibility for timely measurement updates.

Main Findings: If quality measures are not maintained and reported in a timely fashion, this can directly affect a patient's care. CMS, health plans, clinicians, and other stakeholders need to collaborate to invest in quality measurement and measurement sharing.

Strengths/Limitations: The case study focuses on only two measures; it is unknown how often other measures may have similar issues.

Generalizability to Medicare Population: Strong; the measures focused on in the case study are two heart failure measures included in Medicare programs.

Methods: A case study examining two measures was conducted.

Verma S. 2019 Medicare Shared Savings Program ACO Performance: Lower Costs And Promising Results Under 'Pathways To Success'. *Health Affairs Forefront*. 2020. doi:10.1377/forefront.20200914.598838

Subtopic(s): Key Highlights; Performance Measures Used in the Medicare Shared Savings Program

Type of Source: Journal Article

Objective: To examine Accountable Care Organization (ACO) financial and quality performance in the Medicare Shared Savings Program (MSSP) in performance year 2019.

Main Findings: Over 11.2 million fee-for-service Medicare beneficiaries are served by providers in ACOs. Five hundred forty-one ACOs in the MSSP generated \$1.19 billion in total net savings to Medicare in 2019. 2019 was the third year in a row that the program achieved net program savings. There is strong interest from new ACOs to elect the Pathways to Success policies.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the article focused on financial and quality performance in ACOs delivering care to Medicare beneficiaries.

Methods: N/A

Tooker J. The importance of measuring quality and performance in healthcare. *MedGenMed*. 2005;7(2):49. PMID: 16369427

Subtopic(s): Background

Type of Source: Journal article

Objective: To stress the need for measuring quality and performance.

Main Findings: Quality improvement strategies can increase both patient and physician satisfaction. The main goal in creating quality and performance measures in health care is to build a system that promotes top clinical standards and high-quality patient care. Transparency, credibility, and accountability are three tools for creating quality and performance measures.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Low; Medicare populations are not specifically referenced, though the concepts could affect those populations as well.

Methods: N/A

United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures

Type of Source: Report

Objective: To gain insight on strengths and challenges of the incentive payment system in Medicare from the provider perspective.

Main Findings: From 2017-2019, over 90 percent of providers received a less than 2 percent increase to their Medicare payments. The largest adjustment in any year was 1.88 percent. 72 to 84 percent of providers achieved a performance bonus depending on the year.

Strengths/Limitations: There was varying consensus among providers; for example, some said they were not sure the incentives helped improve quality of care at all, while others said it helped smaller practices.

Generalizability to Medicare Population: Strong; the Medicare population is directly studied.

Methods: Qualitative interviews were conducted among providers.

Van Buchem MM, Neve OM, Kant IMJ, Steyerberg EW, Boosman H, Hensen EF. Analyzing patient experiences using natural language processing: Development and validation of the artificial intelligence patient reported experience measure (AI-PREM). *BMC Med Inform Decis Mak.* 2022;22(1):183. doi:10.1186/s12911-022-01923-5

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To develop an Artificial Intelligence Patient-Reported Experience Measures (AI-PREM) survey tool of open-ended questions to evaluate patient experiences.

Main Findings: Five open-ended questions were developed regarding personal approach, collaboration, general information, organization of care, and general experiences. The AI-PREM shows a trend toward positive responses. The ability to give free text responses opposed to a set response selection option leads to more in-depth feedback.

Strengths/Limitations: Limitations include the topic models not being able to be compared with other literature. Strengths include the use of both qualitative and quantitative data.

Generalizability to Medicare Population: Low; the Medicare population was not specifically discussed or considered.

Methods: An open-ended questionnaire was designed along with a Natural Language Processing (NLP) pipeline to analyze responses.

Wilensky G. The Need to Simplify Measuring Quality in Health Care. *JAMA.* 2018;319(23):2369. doi:10.1001/jama.2018.6858

Subtopic(s): Challenges Related to Developing and Implementing Performance Measures

Type of Source: Journal article

Objective: To explore the concerns of clinicians regarding the burden of varying required quality and performance metrics.

Main Findings: Approximately 1,700 health care measures are used by CMS alone. The multitude of measures have led to a need to simplify the way quality is measured in order to relieve the burden on practicing clinicians. The author recommends CMS convene stakeholders to develop simplified metrics important to patients.

Strengths/Limitations: This article is based on the opinion of the author.

Generalizability to Medicare Population: Low; the concept of the article is relevant to Medicare population, but no Medicare-specific recommendations are given.

Methods: N/A

Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev.* 2017;(3):CD011153. doi:10.1002/14651858.CD011153.pub2

Subtopic(s): Performance Measures Used in Other Programs

Type of Source: Report

Objective: To determine the impact of differing payment models regarding the performance of outpatient care facilities and analyze how payment methods differ in varying settings.

Main Findings: 21 international studies were included in the report. It was found that a slight improvement in health professionals' use of health services or outcomes can be achieved if incentives to pay health facilities providing outpatient services are offered. However, this may not affect patients' outcomes.

Strengths/Limitations: Studies included in the review did not account for considerations encouraging patient outcomes.

Generalizability to Medicare Population: Strong; Medicare payments are central to the payment research.

Methods: Literature review was conducted for relevant studies. The reference lists of studies were searched as well and an additional citation search was conducted.

Zhang H, Cowling DW, Graham JM, Taylor E. Impact of a Commercial Accountable Care Organization on Prescription Drugs. *Health Services Research.* 2021;56(4):592-603. doi:10.1111/1475-6773.13626

Subtopic(s): Best Practices for Measuring Performance

Type of Source: Journal article

Objective: To deduce the long-term impact of a commercial ACO on prescription drug spending, utilization, and related quality of care.

Main Findings: Inconsequential differential changes were found in association with commercial ACOs and prescription drug spending, utilization, and related quality of care. There were no significant differential changes throughout the five years examined for the following measures: average total spend, average total prescriptions filled, average generic shares of total prescription drug spending, and annual rates of outpatient process quality of care metrics.

Strengths/Limitations: Actual spending and cost savings for prescription drugs may differ from the allowed payment comparisons as the study allowed payments from PBM pharmacy claims to be included, which did not include retroactive lump sum rebates from PBM financial reconciliation.

Generalizability to Medicare Population: Moderate; although participants were non-Medicare members, findings apply to the Medicare population.

Methods: A longitudinal retrospective cohort study with propensity weighted difference-in-differences regression models was conducted. The sample consisted of members continuously enrolled in a commercial HMO from 2008 through 2014 in the Sacramento area.

Zhu M, Japinga M, Saunders R, McClellan M. *The Future of Risk Adjustment: Supporting Equitable, Comprehensive Health Care.* Duke Margolis Center for Health Policy; 2022.

https://healthpolicy.duke.edu/sites/default/files/2022-06/Margolis%20Future%20Risk%20Adjustment%20Paper%20v3_0.pdf

Subtopic(s): Risk Adjustment Methods Used for PB-TCOC Models

Type of Source: White paper

Objective: To address the fundamental challenges that will arise for risk adjustment as providers shift into accountable care models and health plans become accountable for risk-adjusted spending and quality of care.

Main Findings: Current risk adjustment methodologies fall short on multiple accounts, including accuracy in their predictions and the incentives those predictions create. These shortcomings may be reinforcing structural barriers to health care access and may not be addressing the health of the whole person. The authors suggest that in the short term, incremental steps should be taken by CMS to build better adjustors into both Medicare Advantage and across Medicare Alternative Payment Models. In the long-term, some of their recommendations include linking risk adjustment to population health, aligning incentives and models, and creating and implementing an evidence-driven strategy for risk adjustment to improve equity.

Strengths/Limitations: N/A

Generalizability to Medicare Population: Strong; the report is focused on risk adjustment of payments for Medicare Advantage and Medicare Alternative Payment Models.

Methods: N/A

Appendix K. References

- ¹ Centers for Medicare & Medicaid Services. New to Measures. October 2023. <https://mmshub.cms.gov/about-quality/new-to-measures/what-is-a-measure>
- ² Agency for Healthcare Research and Quality. Types of Healthcare Quality Measures. July 2015. <https://www.ahrq.gov/talkingquality/measures/types.html>
- ³ National Quality Forum. Improving Healthcare Quality. N.D. https://www.qualityforum.org/Setting_Priorities/Improving_Healthcare_Quality.aspx
- ⁴ University of North Carolina. A Model for Measuring Quality Care. N.D. <https://www.med.unc.edu/ihqi/wp-content/uploads/sites/463/2021/01/A-Model-for-Measuring-Quality-Care-NHS-Improvement-brief.pdf>
- ⁵ American Medical Association. Performance Improvement Continuing Medical Education. N.D. <https://www.ama-assn.org/education/ama-pra-credit-system/performance-improvement-continuing-medical-education-pi-cme>
- ⁶ Partnership for Quality Management. Endorsement & Maintenance. N.D. <https://p4qm.org/EM>
- ⁷ Agency for Healthcare Research and Quality. Types of Healthcare Quality Measures. July 2015. <https://www.ahrq.gov/talkingquality/measures/types.html>
- ⁸ National Quality Forum. Improving Healthcare Quality. N.D. https://www.qualityforum.org/Setting_Priorities/Improving_Healthcare_Quality.aspx
- ⁹ American Academy of Family Physicians. Quality Measures. N.D. <https://www.aafp.org/family-physician/practice-and-career/managing-your-practice/quality-measures.html>
- ¹⁰ Partnership for Quality Management. Endorsement & Maintenance. N.D. <https://p4qm.org/EM>
- ¹¹ Centers for Medicare & Medicaid Services. *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹² Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258
- ¹³ Khullar D, Bond AM, O'Donnell EM, Qian Y, Gans DN, Casalino LP. Time and Financial Costs for Physician Practices to Participate in the Medicare Merit-based Incentive Payment System: A Qualitative Study. *JAMA Health Forum*. 2021;2(5):e210527. doi:10.1001/jamahealthforum.2021.0527
- ¹⁴ Medicare Basics Series: MIPS Data Problems. American Medical Association. Published August 18, 2023. Accessed November 28, 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹⁵ Bond AM, Schpero WL, Casalino LP, Zhang M, Khullar D. Association Between Individual Primary Care Physician Merit-based Incentive Payment System Score and Measures of Process and Patient Outcomes. *JAMA*. 2022;328(21):2136-2146. doi:10.1001/jama.2022.20619
- ¹⁶ Liao JM, Navathe AS. Does the Merit-Based Incentive Payment System Disproportionately Affect Safety-Net Practices? *JAMA Health Forum*. 2020;1(5):e200452. doi:10.1001/jamahealthforum.2020.0452
- ¹⁷ Honorable Elizabeth Richter. Published online April 13, 2021. <https://searchlf.ama-assn.org/letter/documentDownload?uri=%2Funstructured%2Fbinary%2Fletter%2FLETTERS%2F2021-4-13-MIPS-Benchmarks-Sign-on-Letter-to-CMS.pdf>
- ¹⁸ Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS — The Universal Foundation. *N Engl J Med*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539
- ¹⁹ Aligning Quality Measures Across CMS — The Universal Foundation. Centers for Medicare & Medicaid Services. September 2023. <https://www.cms.gov/aligning-quality-measures-across-cms-universal-foundation>
- ²⁰ Centers for Medicare & Medicaid Services. *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>

-
- ²¹ Damberg CL, Sorbero ME, Lovejoy SL, Martsof GR, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. *Rand Health Q.* 2014;4(3):9.
- ²² Medicare Payment Advisory Commission. *Chapter 15: Moving beyond the Merit-Based Incentive Payment System (March 2018 Report).*; 2018. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar18_medpac_ch15_sec.pdf
- ²³ Fisher H. Watchdog Highlights Challenges of Medicare Payment Program. *The Regulatory Review.* January 2022. <https://www.theregreview.org/2022/01/27/fisher-watchdog-highlights-challenges-medicare-payment-program/>
- ²⁴ Centers for Medicare & Medicaid Services. Measure Inventory Tool. <https://cmit.cms.gov/cmit/#/MeasureInventory>
- ²⁵ Abt Associates. Evaluation of the Oncology Care Model. Abt Associates; 2021. <https://www.cms.gov/priorities/innovation/data-and-reports/2021/ocm-ar4-eval-part-persp-report>
- ²⁶ Arbor Research. HHVBP Evaluation Final Report. Arbor Research; 2023. <https://www.cms.gov/priorities/innovation/data-and-reports/2023/hhvbp-seventh-ann-rpt>
- ²⁷ Mathematica. Evaluation of the Primary Care First Model. Mathematica; 2022:201. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/pcf-first-eval-rpt>
- ²⁸ Mathematica. Independent Evaluation of the Comprehensive Primary Care Plus (CPC+): Fourth Annual Report. Mathematica; 2022:200. <https://www.mathematica.org/publications/independent-evaluation-of-the-comprehensive-primary-care-plus-cpc-fourth-annual-report>
- ²⁹ American Hospital Association. Comments on HHS Proposal to Sunset the Regulatory Review Schedule. 2020. <https://www.aha.org/lettercomment/2020-12-07-comments-hhs-proposal-sunset-regulatory-review-schedule>
- ³⁰ Centers for Medicare & Medicaid Services. Synthesis of Evaluation Results across 21 Medicare Models, 2012-2020. Centers for Medicare & Medicaid Services; 2022:49. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/wp-eval-synthesis-21models>
- ³¹ NORC at the University of Chicago. Third Evaluation Report: Next Generation Accountable Care Organization Model Evaluation. NORC at the University of Chicago; 2020:137. <https://innovation.cms.gov/data-and-reports/2020/nextgenaco-thirdevalrpt-fullreport>
- ³² Kaufman BG, Bleser WK, Saunders R, et al. Prospective or retrospective ACO attribution matters for seriously ill patients. *The American Journal of Managed Care.* 2020;26(12):534-40. doi: 10.37765/ajmc.2020.88541
- ³³ Verma S. 2019 Medicare Shared Savings Program ACO Performance: Lower Costs And Promising Results Under 'Pathways To Success'. *Health Affairs Forefront.* 2020. Doi: 10.1377/forefront.20200914.598838
- ³⁴ CMS. Medicare Shared Savings Program Saves Medicare More Than \$1.8 Billion in 2022 and Continues to Deliver High-quality Care | CMS. cms.gov. Published 2023. Accessed December 13, 2023. <https://www.cms.gov/newsroom/press-releases/medicare-shared-savings-program-saves-medicare-more-18-billion-2022-and-continues-deliver-high>
- ³⁵ Centers for Medicare & Medicaid Services. 2024 Medicare Advantage and Part D Star Ratings. Centers for Medicare & Medicaid Services. Published October 2023. Accessed January 24, 2024. <https://www.cms.gov/newsroom/fact-sheets/2024-medicare-advantage-and-part-d-star-ratings>
- ³⁶ Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. mckinsey.com. Accessed December 13, 2023. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>
- ³⁷ Urban Institute. The Medicare Advantage Quality Bonus Program: High Cost for Uncertain Gain. June 2023. <https://www.urban.org/sites/default/files/2023-06/The%20Medicare%20Advantage%20Quality%20Bonus%20Program.pdf>
- ³⁸ Five-Star Quality Rating System. Centers for Medicare & Medicaid Services. February 2024. <https://www.cms.gov/medicare/health-safety-standards/certification-compliance/five-star-quality-rating-system>
- ³⁹ Home Health Quality of Patient Care Star Rating Provider Preview Report. Centers for Medicare & Medicaid Services. <https://www.cms.gov/files/document/hhqualityofpatientcarestarratingjanuary2023sampleproviderpreview.pdf>
- ⁴⁰ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med.* 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0

-
- ⁴¹ Hamilton FL, Greaves F, Majeed A, Millett C. Effectiveness of providing financial incentives to healthcare professionals for smoking cessation activities: systematic review. *Tob Control*. 2013;22(1):3-8. doi:10.1136/tobaccocontrol-2011-050048
- ⁴² Pirritano M, Miller Parrish K, Kim Y, Solomon H, Keene J. It takes quality improvement to cross the chasm. *BMJ Open Qual*. 2023;12(3):e001906. doi:10.1136/bmj-2022-001906
- ⁴³ Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open*. 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ⁴⁴ Pirritano M, Miller Parrish K, Kim Y, Solomon H, Keene J. It takes quality improvement to cross the chasm. *BMJ Open Qual*. 2023;12(3):e001906. doi:10.1136/bmj-2022-001906
- ⁴⁵ Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev*. 2017;2017(3):CD011153. doi:10.1002/14651858.CD011153.pub2
- ⁴⁶ Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev*. 2017;2017(3):CD011153. doi:10.1002/14651858.CD011153.pub2
- ⁴⁷ Huang J, Yin S, Lin Y, Jiang Q, He Y, Du L. Impact of pay-for-performance on management of diabetes: a systematic review. *Journal of Evidence-Based Medicine*. 2013;6(3):173-184. doi:10.1111/jebm.12052
- ⁴⁸ Rhee C, Strich JR, Chiotos K, et al. Improving Sepsis Outcomes in the Era of Pay-for-Performance and Electronic Quality Measures: A Joint IDSA/ACEP/PIDS/SHEA/SHM/SIDP Position Paper. *Clin Infect Dis*. Published online October 13, 2023:ciad447. doi:10.1093/cid/ciad447
- ⁴⁹ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). doi:10.1056/CAT.18.0245
- ⁵⁰ Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open*. 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ⁵¹ Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open*. 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ⁵² Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open*. 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ⁵³ Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine*. 2022;37(2):359. doi:10.1007/s11606-021-06617-8
- ⁵⁴ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). doi:10.1056/CAT.18.0245
- ⁵⁵ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). doi:10.1056/CAT.18.0245
- ⁵⁶ Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open*. 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ⁵⁷ Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database Syst Rev*. 2021;2021(1):CD011865. doi:10.1002/14651858.CD011865.pub2
- ⁵⁸ Mendelson A, Kondo K, Damberg C, et al. The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care. *Ann Intern Med*. 2017;166(5):341-353. doi:10.7326/M16-1881
- ⁵⁹ Mendelson A, Kondo K, Damberg C, et al. The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care. *Ann Intern Med*. 2017;166(5):341-353. doi:10.7326/M16-1881
- ⁶⁰ Ogundejì YK, Bland JM, Sheldon TA. The effectiveness of payment for performance in health care: A meta-analysis and exploration of variation in outcomes. *Health Policy*. 2016;120(10):1141-1150. doi:10.1016/j.healthpol.2016.09.002

-
- ⁶¹ Health Care Payment Learning & Action Network. Alternative Payment Model Framework. 2017. <https://hcp-lan.org/workproducts/apm-refresh-whitepaper-final.pdf>
- ⁶² Kim KM, Max W, White JS, Chapman SA, Muench U. Do penalty-based pay-for-performance programs improve surgical care more effectively than other payment strategies? A systematic review. *Ann Med Surg (Lond)*. 2020;60:623-630. doi:10.1016/j.amsu.2020.11.060
- ⁶³ Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med*. 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8
- ⁶⁴ Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med*. 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8
- ⁶⁵ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). doi:10.1056/CAT.18.0245
- ⁶⁶ Conway A, Satin D. The role of pay-for-performance in reducing healthcare disparities: A narrative literature review. *Prev Med*. 2022;164:107274. doi:10.1016/j.ypmed.2022.107274
- ⁶⁷ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). Doi:10.1056/CAT.18.0245
- ⁶⁸ NORC at the University of Chicago. Environmental Scan on Improving Management of Care Transitions in Population-Based Models. 2023. <https://aspe.hhs.gov/sites/default/files/documents/61e603e1beb3f5eb4d528b1e91fadf12/PTAC-Jun-12-Escan.pdf>
- ⁶⁹ Grants NRC (US) P on PM and D for PHPP, Perrin EB, Durch JS, Skillman SM. Data and Information Systems: Issues for Performance Measurement. In: *Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network*. National Academies Press (US); 1999. Accessed November 30, 2023. <https://www.ncbi.nlm.nih.gov/books/NBK231008/>
- ⁷⁰ IMPACT Act of 2014 Data Standardization & Cross Setting Measures. Centers for Medicare & Medicaid Services. September 2023. <https://www.cms.gov/medicare/quality/initiatives/pac-quality-initiatives/impact-act-2014-data-standardization-cross-setting-measures#:~:text=The%20IMPACT%20Act%20requires%20the%20Secretary%20to%20implement%20specified%20clinical,from%20five%20quality%20m>
- ⁷¹ Berenson RA, Shartzter A, Pham HH. Beyond demonstrations: implementing a primary care hybrid payment model in Medicare. *Health Affairs Scholar*. 2023;1(2):qxad024. doi:10.1093/haschl/qxad024
- ⁷² Skopec L, Bowen G, Zuckerman S. *Accounting for Social Risk in Value-Based Payment and Quality Measurement* / Urban Institute. Urban Institute; 2022. Accessed December 1, 2023. <https://www.urban.org/research/publication/accounting-social-risk-value-based-payment-and-quality-measurement>
- ⁷³ Centers for Medicare & Medicaid Services. Consumer Assessment of Healthcare Providers & Systems. September 2023. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems>
- ⁷⁴ Agency for Healthcare Research and Quality. Understanding CAHPS Surveys: A Primer for New Users. January 2019. <https://www.ahrq.gov/sites/default/files/wysiwyg/cahps/news-and-events/events/webinars/understanding-cahps-101-infographic.pdf>
- ⁷⁵ Bland C, Zuckerbraun S, Lines LM, et al. Challenges Facing CAHPS Surveys and Opportunities for Modernization [Internet]. Research Triangle Park (NC): RTI Press; 2022 Nov. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK592584/> doi: 10.3768/rtipress.2022.op.0080.2211
- ⁷⁶ Health Care Payment Learning & Action Network. *Accelerating and Aligning Population-Based Payment Models: Performance Measurement*. HCPLAN; 2016. <https://hcp-lan.org/pm-whitepaper/>, <https://hcp-lan.org/pm-whitepaper/>
- ⁷⁷ Krishnan M, Brunelli SM, Maddux FW, et al. Guiding principles and checklist for population-based quality metrics. *Clin J Am Soc Nephrol*. 2014;9(6):1124-1131. doi:10.2215/CJN.11061013

-
- ⁷⁸ MedPAC. March 2022 Report to the Congress: Medicare Payment Policy. <https://www.medpac.gov/document/march-2022-report-to-the-congress-medicare-payment-policy/>
- ⁷⁹ Anderson AC, O'Rourke E, Chin MH, Ponce NA, Bernheim SM, Burstin H. Promoting Health Equity And Eliminating Disparities Through Performance Measurement And Payment. *Health Aff (Millwood)*. 2018;37(3):371-377. doi:10.1377/hlthaff.2017.1301
- ⁸⁰ Centers for Medicare & Medicaid Services. Digital Quality Strategic Roadmap. CMS; 2022. https://ecqi.healthit.gov/sites/default/files/CMSdQMStrategicRoadmap_032822.pdf.
- ⁸¹ RTI. ACO Realizing Equity, Access, and Community Health (REACH) Model - PY2024 Quality Measurement Methodology. Centers for Medicare & Medicaid Services; n.d. <https://www.cms.gov/files/document/aco-reach-quality-msr-meth-py24.pdf>
- ⁸² MedPAC. March 2022 Report to the Congress: Medicare Payment Policy – MedPAC; 2022. <https://www.medpac.gov/document/march-2022-report-to-the-congress-medicare-payment-policy/>
- ⁸³ 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/reports/second-report-congress-social-risk-medicare-value-based-purchasing-programs>
- ⁸⁴ Basch E, Spertus J, Adams Dudley R, et al. Methods for Developing Patient-Reported Outcome-Based Performance Measures (PRO-PMs). *Value in Health*. 2015;18(4):493-504. doi:10.1016/j.jval.2015.02.018
- ⁸⁵ Greenhalgh J, Dalkin S, Gooding K, et al. Functionality and Feedback: A Realist Synthesis of the Collation, Interpretation and Utilisation of Patient-Reported Outcome Measures Data to Improve Patient Care. *NIHR Journals Library*; 2017. <http://www.ncbi.nlm.nih.gov/books/NBK409450/>
- ⁸⁶ Agency for Healthcare Research and Quality. Types of Healthcare Quality Measures. July 2015 <https://www.ahrq.gov/talkingquality/measures/types.html>
- ⁸⁷ National Quality Forum. Improving Healthcare Quality. N.D. https://www.qualityforum.org/Setting_Priorities/Improving_Healthcare_Quality.aspx
- ⁸⁸ American Academy of Family Physicians. Quality Measures. N.D. <https://www.aafp.org/family-physician/practice-and-career/managing-your-practice/quality-measures.html>
- ⁸⁹ Partnership for Quality Management. Endorsement & Maintenance. N.D. <https://p4qm.org/EM>
- ⁹⁰ The Agency for Healthcare Research and Quality. Major Physician Measurement Sets. January 2023. <https://www.ahrq.gov/talkingquality/measures/setting/physician/measurement-sets.html>
- ⁹¹ Physician Advocacy Institute. What are specialty measure sets? <https://www.physiciansadvocacyinstitute.org/Physician-Practice-Resources/Medicare-QPP/FAQs/mips-quality-category-4-what-are-specialty-measure-sets>
- ⁹² Agency for Healthcare Research and Quality. Measuring the Quality of Physician Care. September 2019. <https://www.ahrq.gov/talkingquality/measures/setting/physician/index.html>
- ⁹³ Agency for Healthcare Research and Quality. Measures of Quality for Different Health Care Settings. January 2023. <https://www.ahrq.gov/talkingquality/measures/setting/index.html>
- ⁹⁴ Centers for Medicare & Medicaid Services. IMPACT Act of 2014 Data Standardization & Cross Setting Measures. September 6, 2023. <https://www.cms.gov/medicare/quality/initiatives/pac-quality-initiatives/impact-act-2014-data-standardization-cross-setting-measures>
- ⁹⁵ IMPACT Act of 2014 Data Standardization & Cross Setting Measures. Centers for Medicare & Medicaid Services. September 2023. <https://www.cms.gov/medicare/quality/initiatives/pac-quality-initiatives/impact-act-2014-data-standardization-cross-setting-measures#:~:text=The%20IMPACT%20Act%20requires%20the%20Secretary%20to%20implement%20specified%20clinical,from%20five%20quality%20measure%20domains.>
- ⁹⁶ Inpatient Rehabilitation Facility (IRF) Quality Reporting Program (QRP) Measures Information. Centers for Medicare & Medicaid Services. March 2024. <https://www.cms.gov/medicare/quality/inpatient-rehabilitation-facility/irf-quality-reporting-measures-information>
- ⁹⁷ Skilled Nursing Facility (SNF) Quality Reporting Program (QRP) Measures and Technical Information. Centers for Medicare & Medicaid Services. March 2024. <https://www.cms.gov/medicare/quality/snf-quality-reporting-program/measures-and-technical-information>

-
- ⁹⁸ Home Health Quality Measures. Centers for Medicare & Medicaid Services. February 2024. <https://www.cms.gov/medicare/quality/home-health/home-health-quality-measures>
- ⁹⁹ Long-Term Care Hospital (LTCH) Quality Reporting Program (QRP) Measures Information. Centers for Medicare & Medicaid Services. March 2024. <https://www.cms.gov/medicare/quality/long-term-care-hospital/lrch-quality-reporting-measures-information>
- ¹⁰⁰ Physician-Focused Payment Model Technical Advisory Committee. *Report to the Secretary of Health and Human Services: Optimizing Population-Based Total Cost of Care (PB-TCOC) Models in the Context of Alternative Payment Models (APMs) and Physician-Focused Payment Models (PFPs)*. March 2023. <https://aspe.hhs.gov/sites/default/files/documents/4b65476c58e363735aa9065a82a35df4/PTAC-TCOC-RTS.pdf>
- ¹⁰¹ Centers for Medicare & Medicaid Services. New to Measures. August 2023. <https://mmshub.cms.gov/about-quality/new-to-measures/roles>
- ¹⁰² American Academy of Family Physicians. 2023. Quality Measures. <https://www.aafp.org/family-physician/practice-and-career/managing-your-practice/quality-measures.html>
- ¹⁰³ Centers for Medicare & Medicaid Services. New to Measures. August 2023. <https://mmshub.cms.gov/about-quality/new-to-measures/roles>
- ¹⁰⁴ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹⁰⁵ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹⁰⁶ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹⁰⁷ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹⁰⁸ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹⁰⁹ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹¹⁰ Centers for Medicare & Medicaid Services. Measure Production & Monitoring. October 2023. <https://mmshub.cms.gov/measure-lifecycle/measure-use/production-and-monitoring>
- ¹¹¹ Centers for Medicare & Medicaid Services. Blueprint Measure Lifecycle Overview. August 2023.
- ¹¹² American Academy of Family Physicians. Performance Measure Criteria. 2023. <https://www.aafp.org/about/policies/all/performance-measures.html>
- ¹¹³ American Academy of Orthopedic Surgeons. Quality Performance Measures. 2018. https://www.aaos.org/globalassets/quality-and-practice-resources/quality-performance-measures-resouces/pm_methods_v2.pdf
- ¹¹⁴ Tooker J; Ambulatory Care Quality Alliance. The importance of measuring quality and performance in healthcare. *MedGenMed*. 2005 Jun 3;7(2):49. PMID: 16369427; PMCID: PMC1681600.
- ¹¹⁵ Agency for Healthcare Research and Quality. Section 4: Ways To Approach the Quality Improvement Process. January 2020.
- ¹¹⁶ Committee on Quality Measures for the Healthy People Leading Health Indicators; Board on Population Health and Health Practice; Institute of Medicine. *Toward Quality Measures for Population Health and the Leading Health Indicators*. Washington (DC): National Academies Press (US); 2013 Apr 9. 2, Criteria for Selecting Measures. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK202189/>
- ¹¹⁷ DeGroff A, Schooley M, Chapel T, Poister TH. Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning*. 2010;33(4):365-372. doi:10.1016/j.evalprogplan.2010.02.003
- ¹¹⁸ Kansagara D, Tuepker A, Joos S, Nicolaidis C, Skaperdas E, Hickam D. Getting Performance Metrics Right: A Qualitative Study of Staff Experiences Implementing and Measuring Practice Transformation. *J GEN INTERN MED*. 2014;29(S2):607-613. doi:10.1007/s11606-013-2764-y
- ¹¹⁹ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹²⁰ DeGroff A, Schooley M, Chapel T, Poister TH. Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning*. 2010;33(4):365-372. doi:10.1016/j.evalprogplan.2010.02.003

-
- ¹²¹ Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258
- ¹²² Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258
- ¹²³ Traditional MIPS Overview. CMS Quality Payment Program. Accessed December 13, 2023. <https://qpp.cms.gov/mips/traditional-mips>
- ¹²⁴ Khullar D, Bond AM, O'Donnell EM, Qian Y, Gans DN, Casalino LP. Time and Financial Costs for Physician Practices to Participate in the Medicare Merit-based Incentive Payment System: A Qualitative Study. *JAMA Health Forum*. 2021;2(5):e210527. doi:10.1001/jamahealthforum.2021.0527
- ¹²⁵ Khullar D, Bond AM, O'Donnell EM, Qian Y, Gans DN, Casalino LP. Time and Financial Costs for Physician Practices to Participate in the Medicare Merit-based Incentive Payment System: A Qualitative Study. *JAMA Health Forum*. 2021;2(5):e210527. doi:10.1001/jamahealthforum.2021.0527
- ¹²⁶ United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>
- ¹²⁷ United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>
- ¹²⁸ United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>
- ¹²⁹ Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258
- ¹³⁰ Berdahl CT, Easterlin MC, Ryan G, Needleman J, Nuckols TK. Primary Care Physicians in the Merit-Based Incentive Payment System (MIPS): A Qualitative Investigation of Participants' Experiences, Self-Reported Practice Changes, and Suggestions for Program Administrators. *J GEN INTERN MED*. 2019;34(10):2275-2281. doi:10.1007/s11606-019-05207-z
- ¹³¹ Fisher H. Watchdog Highlights Challenges of Medicare Payment Program. The Regulatory Review. Published January 27, 2022. Accessed November 28, 2023. <https://www.theregreview.org/2022/01/27/fisher-watchdog-highlights-challenges-medicare-payment-program/>
- ¹³² Fisher H. Watchdog Highlights Challenges of Medicare Payment Program. The Regulatory Review. Published January 27, 2022. Accessed November 28, 2023. <https://www.theregreview.org/2022/01/27/fisher-watchdog-highlights-challenges-medicare-payment-program/>
- ¹³³ Khullar D, Bond AM, O'Donnell EM, Qian Y, Gans DN, Casalino LP. Time and Financial Costs for Physician Practices to Participate in the Medicare Merit-based Incentive Payment System: A Qualitative Study. *JAMA Health Forum*. 2021;2(5):e210527. doi:10.1001/jamahealthforum.2021.0527
- ¹³⁴ Berdahl CT, Easterlin MC, Ryan G, Needleman J, Nuckols TK. Primary Care Physicians in the Merit-Based Incentive Payment System (MIPS): A Qualitative Investigation of Participants' Experiences, Self-Reported Practice Changes, and Suggestions for Program Administrators. *J GEN INTERN MED*. 2019;34(10):2275-2281. doi:10.1007/s11606-019-05207-z
- ¹³⁵ Medicare Basics Series: MIPS Data Problems. American Medical Association. Published August 18, 2023. Accessed November 28, 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹³⁶ United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>
- ¹³⁷ CCW White Paper: Medicare Claims Maturity. Chronic Condition Data Warehouse. 2017. <https://www2.ccwdata.org/documents/10280/19002256/medicare-claims-maturity.pdf>

-
- ¹³⁸ Medicare Basics Series: MIPS Data Problems. American Medical Association. Published August 18, 2023. Accessed November 28, 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹³⁹ United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>
- ¹⁴⁰ Medicare Payment Advisory Commission. *Chapter 15: Moving beyond the Merit-Based Incentive Payment System (March 2018 Report)*; 2018. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar18_medpac_ch15_sec.pdf
- ¹⁴¹ Bond AM, Schpero WL, Casalino LP, Zhang M, Khullar D. Association Between Individual Primary Care Physician Merit-based Incentive Payment System Score and Measures of Process and Patient Outcomes. *JAMA*. 2022;328(21):2136-2146. doi:10.1001/jama.2022.20619
- ¹⁴² Bond AM, Schpero WL, Casalino LP, Zhang M, Khullar D. Association Between Individual Primary Care Physician Merit-based Incentive Payment System Score and Measures of Process and Patient Outcomes. *JAMA*. 2022;328(21):2136-2146. doi:10.1001/jama.2022.20619
- ¹⁴³ Bond AM, Schpero WL, Casalino LP, Zhang M, Khullar D. Association Between Individual Primary Care Physician Merit-based Incentive Payment System Score and Measures of Process and Patient Outcomes. *JAMA*. 2022;328(21):2136-2146. doi:10.1001/jama.2022.20619
- ¹⁴⁴ Medicare Basics Series: MIPS Data Problems. American Medical Association. August 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹⁴⁵ Liao JM, Navathe AS. Does the Merit-Based Incentive Payment System Disproportionately Affect Safety-Net Practices? *JAMA Health Forum*. 2020;1(5):e200452. doi:10.1001/jamahealthforum.2020.0452
- ¹⁴⁶ Medicare Basics Series: MIPS Data Problems. American Medical Association. August 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹⁴⁷ Medicare Basics Series: MIPS Data Problems. American Medical Association. August 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹⁴⁸ Honorable Elizabeth Richter. April 2021. <https://searchlf.ama-assn.org/letter/documentDownload?uri=%2Funstructured%2Fbinary%2Fletter%2FLETTERS%2F2021-4-13-MIPS-Benchmarks-Sign-on-Letter-to-CMS.pdf>
- ¹⁴⁹ United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. <https://www.gao.gov/products/gao-22-104667>
- ¹⁵⁰ Medicare Basics Series: MIPS Data Problems. American Medical Association. August 2023. <https://www.ama-assn.org/practice-management/payment-delivery-models/medicare-basics-series-mips-data-problems>
- ¹⁵¹ National Quality Measures Clearinghouse. Agency for Healthcare Research and Quality. Published 2018. <https://qualitymeasures.ahrq.gov/browse/setting>
- ¹⁵² Centers for Medicare & Medicaid Services Measures Inventory Tool. <https://cmit.cms.gov/cmit/#/>
- ¹⁵³ DeGroff A, Schooley M, Chapel T, Poister TH. Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning*. 2010;33(4):365-372. doi:10.1016/j.evalproplan.2010.02.003
- ¹⁵⁴ Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258
- ¹⁵⁵ Chou P. HEDIS Diabetes Measures – Understanding Comprehensive Care. Stability Health. <https://stabilityhealth.com/hedis-diabetes-measures/>
- ¹⁵⁶ American Diabetes Association. Standards of Medical Care in Diabetes—2014. *Diabetes Care*. 2014;37(Supplement_1):S14-S80. doi:10.2337/dc14-S014
- ¹⁵⁷ Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS – The Universal Foundation. *N Engl J Med*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539
- ¹⁵⁸ Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):401-406. doi:10.1377/hlthaff.2015.1258

-
- ¹⁵⁹ Wilensky G. The Need to Simplify Measuring Quality in Health Care. *JAMA*. 2018;319(23):2369. doi:10.1001/jama.2018.6858
- ¹⁶⁰ Rowley R. Claims Data – The Good, The Bad and The Ugly. HealthIT Answers. Published May 28, 2014. Accessed January 10, 2024. <https://www.healthitanswers.net/claims-data-the-good-the-bad-and-the-ugly/>
- ¹⁶¹ Aligning Quality Measures Across CMS – The Universal Foundation. Centers for Medicare & Medicaid Services. Published September 6, 2023. <https://www.cms.gov/aligning-quality-measures-across-cms-universal-foundation>
- ¹⁶² Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS – The Universal Foundation. *N Engl J Med*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539
- ¹⁶³ Aligning Quality Measures Across CMS – The Universal Foundation. Centers for Medicare & Medicaid Services. Published September 6, 2023. <https://www.cms.gov/aligning-quality-measures-across-cms-universal-foundation>
- ¹⁶⁴ Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS – The Universal Foundation. *N Engl J Med*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539
- ¹⁶⁵ Meaningful Measures 2.0: Moving from Measure Reduction to Modernization. CMS. Published September 6, 2023. Accessed December 13, 2023. <https://www.cms.gov/medicare/quality/meaningful-measures-initiative/meaningful-measures-20>
- ¹⁶⁶ Centers for Medicare & Medicaid Services. 2024 National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report. Baltimore, MD: U.S. Department of Health and Human Services; 2024. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-AssessmentInstruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports>
- ¹⁶⁷ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁶⁸ Harmonization. The Measures Management System. Published July 2023. Accessed January 11, 2024. <https://mmshub.cms.gov/measure-lifecycle/measure-specification/harmonization>
- ¹⁶⁹ Damberg CL, Sorbero ME, Lovejoy SL, Martsof GR, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. *Rand Health Q*. 2014;4(3):9.
- ¹⁷⁰ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁷¹ National Quality Forum. NQF's Aligned Innovation Initiative to Advance Next Generation Measures of Outcomes for Behavioral Health and Maternity Care. April 13, 2023. https://www.qualityforum.org/News_And_Resources/Press_Releases/2023/NQF_s_Aligned_Innovation_Initiative_to_Advance_Next_Generation_Measures_of_Outcomes_for_Behavioral_Health_and_Maternity_Care.aspx
- ¹⁷² United States Government Accountability Office. *Medicare: Provider Performance and Experiences Under the Merit-Based Incentive Payment System*. United States Government Accountability Office; 2021. Accessed November 28, 2023. <https://www.gao.gov/products/gao-22-104667>
- ¹⁷³ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁷⁴ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>

-
- ¹⁷⁵ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁷⁶ CMS Consensus-Based Entity (CBE) Endorsement and Maintenance. Published December 2022. <https://mmshub.cms.gov/sites/default/files/Blueprint-CMS-CBE-Endorsement-Maintenance.pdf>
- ¹⁷⁷ Partnership for Quality Measurement. Endorsement and Maintenance (E&M) Guidebook. Published online October 2023. Accessed December 13, 2023. https://p4qm.org/sites/default/files/2023-10/Del-3-6-Endorsement-and-Maintenance-Guidebook-Final_0.pdf
- ¹⁷⁸ Partnership for Quality Measurement. Endorsement and Maintenance (E&M) Guidebook. Published online October 2023. Accessed December 13, 2023. https://p4qm.org/sites/default/files/2023-10/Del-3-6-Endorsement-and-Maintenance-Guidebook-Final_0.pdf
- ¹⁷⁹ Partnership for Quality Measurement. Endorsement and Maintenance (E&M) Guidebook. Published online October 2023. Accessed December 13, 2023. https://p4qm.org/sites/default/files/2023-10/Del-3-6-Endorsement-and-Maintenance-Guidebook-Final_0.pdf
- ¹⁸⁰ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁸¹ Measure Implementation. CMS Measures Management System (MMS). March 2023. <https://mmshub.cms.gov/measure-lifecycle/measure-implementation/overview>
- ¹⁸² Measure Implementation. CMS Measures Management System (MMS). March 2023. <https://mmshub.cms.gov/measure-lifecycle/measure-implementation/overview>
- ¹⁸³ Measure Implementation. CMS Measures Management System (MMS). March 2023. <https://mmshub.cms.gov/measure-lifecycle/measure-implementation/overview>
- ¹⁸⁴ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁸⁵ Medicare Payment Advisory Commission. *Chapter 15: Moving beyond the Merit-Based Incentive Payment System (March 2018 Report)*.; 2018. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar18_medpac_ch15_sec.pdf
- ¹⁸⁶ *CMS Quality Measure Development Plan: Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs)*. Centers for Medicare & Medicaid Services; 2016. <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/value-based-programs/macra-mips-and-apms/final-mdp.pdf>
- ¹⁸⁷ Damberg CL, Sorbero ME, Lovejoy SL, Martsolf GR, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. *Rand Health Q.* 2014;4(3):9.
- ¹⁸⁸ Damberg CL, Sorbero ME, Lovejoy SL, Martsolf GR, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions. *Rand Health Q.* 2014;4(3):9.
- ¹⁸⁹ Medicare Payment Advisory Commission. *Chapter 15: Moving beyond the Merit-Based Incentive Payment System (March 2018 Report)*.; 2018. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar18_medpac_ch15_sec.pdf
- ¹⁹⁰ Medicare Payment Advisory Commission. *Chapter 15: Moving beyond the Merit-Based Incentive Payment System (March 2018 Report)*.; 2018. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar18_medpac_ch15_sec.pdf

-
- ¹⁹¹ Fisher H. Watchdog Highlights Challenges of Medicare Payment Program. The Regulatory Review. January 2022. Accessed November 28, 2023. <https://www.theregreview.org/2022/01/27/fisher-watchdog-highlights-challenges-medicare-payment-program/>
- ¹⁹² LaPointe J. The Most Successful Alternative Payment Models from CMMI, To Date. Revcycle Intelligence. Published December 13, 2022. <https://revcycleintelligence.com/features/the-most-successful-alternative-payment-models-from-cmmi-to-date>
- ¹⁹³ Giovannetti, E. R., Dy, S. M., Leff, B., Weston, C., Adams, K., Valuck, T., Pittman, A. T., Blaum, C. S., McCann, B. A., & Boyd, C. M. (2013). Performance measurement for people with multiple chronic conditions: conceptual model. *PubMed*. <https://pubmed.ncbi.nlm.nih.gov/24304183>
- ¹⁹⁴ Duseja R, Andress J, Sandhu AT, et al. Development of Episode-Based Cost Measures for the US Medicare Merit-based Incentive Payment System. *JAMA Health Forum*. 2021;2(5):e210451. doi:10.1001/jamahealthforum.2021.0451
- ¹⁹⁵ Centers for Medicare & Medicaid Services. Expanded Home Health Value-Based Purchasing (HHVBP) Model An Overview: Risk Adjustment Process and Use in the Expanded HHVBP Model. May 2022. <https://www.cms.gov/priorities/innovation/media/document/hhvbp-exp-risk-adj-resource>
- ¹⁹⁶ Jacobs DB, Schreiber M, Seshamani M, et al. Aligning Quality Measures Across CMS—The Universal Foundation. *New England Journal of Medicine*. 2023;388(9):776-9. doi: 10.1056/NEJMp2215539
- ¹⁹⁷ Arbor Research. HHVBP Evaluation Final Report. Arbor Research; 2023. <https://www.cms.gov/priorities/innovation/data-and-reports/2023/hhvbp-seventh-ann-rpt>
- ¹⁹⁸ Abt Associates. Evaluation of the Oncology Care Model. Abt Associates; 2021. <https://www.cms.gov/priorities/innovation/data-and-reports/2021/ocm-ar4-eval-part-persp-report>
- ¹⁹⁹ Mathematica. *Evaluation of the Primary Care First Model*. Mathematica; 2022:201. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/pcf-first-eval-rpt>
- ²⁰⁰ Mathematica. *Evaluation of the Primary Care First Model*. Mathematica; 2022:201. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/pcf-first-eval-rpt>
- ²⁰¹ Mathematica. *Independent Evaluation of the Comprehensive Primary Care Plus (CPC+): Fourth Annual Report*. Mathematica; 2022:200. <https://www.mathematica.org/publications/independent-evaluation-of-the-comprehensive-primary-care-plus-cpc-fourth-annual-report>
- ²⁰² American Hospital Association. Comments on HHS Proposal to Sunset the Regulatory Review Schedule. 2020. <https://www.aha.org/lettercomment/2020-12-07-comments-hhs-proposal-sunset-regulatory-review-schedule>
- ²⁰³ Arbor Research. *Evaluation of the Home Health Value-Based Purchasing (HHVBP) Model*. Arbor Research; 2023:252. <https://www.cms.gov/priorities/innovation/data-and-reports/2023/hhvbp-sixth-ann-rpt-app>
- ²⁰⁴ NORC at the University of Chicago. Third Evaluation Report: Next Generation Accountable Care Organization Model Evaluation. NORC at the University of Chicago; 2020:137. <https://innovation.cms.gov/data-and-reports/2020/nextgenaco-thirdevalrpt-fullreport>
- ²⁰⁵ NORC at the University of Chicago. Fifth Evaluation Report: Next Generation Accountable Care Organization Model Evaluation. NORC at the University of Chicago; 2022:44. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/nextgenaco-fifthevalrpt>
- ²⁰⁶ Abt Associates. *Evaluation of the Oncology Care Model: Executive Summary*. Abt Associates; 2023:10 <https://www.cms.gov/priorities/innovation/data-and-reports/2023/ocm-evaluation-pp1-9-exec-sum>
- ²⁰⁷ Mathematica. *Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Fourth Annual Report*. Mathematica; 2022:200 <https://www.cms.gov/priorities/innovation/data-and-reports/2022/cpc-plus-fourth-annual-eval-report>
- ²⁰⁸ Arbor Research. HHVBP Evaluation Final Report. Arbor Research; 2023. <https://www.cms.gov/priorities/innovation/data-and-reports/2023/hhvbp-seventh-ann-rpt>
- ²⁰⁹ Kaufman BG, Bleser WK, Saunders R, et al. Prospective or retrospective ACO attribution matters for seriously ill patients. *The American Journal of Managed Care*. 2020;26(12):534-40. doi: 10.37765/ajmc.2020.88541
- ²¹⁰ CMS. CMS Finalizes Physician Payment Rule that Advances Health Equity. cms.gov. Published 2023. Accessed December 13, 2023. <https://www.cms.gov/newsroom/press-releases/cms-finalizes-physician-payment-rule-advances-health-equity>

-
- ²¹¹ Ryan AM, Markovitz AA. Estimated Savings From the Medicare Shared Savings Program. *JAMA Health Forum*; 2023;4(12): e234449-e234449. doi:10.1001/jamahealthforum.2023.4449
- ²¹² Verma S. 2019 Medicare Shared Savings Program ACO Performance: Lower Costs And Promising Results Under ‘Pathways To Success’. *Health Affairs Forefront*. 2020. Doi: 10.1377/forefront.20200914.598838
- ²¹³ Centers for Medicare & Medicaid Services. Medicare Shared Savings Program Saves Medicare More Than \$1.8 Billion in 2022 and Continues to Deliver High-quality Care. 2023. <https://www.cms.gov/newsroom/press-releases/medicare-shared-savings-program-saves-medicare-more-18-billion-2022-and-continues-deliver-high>
- ²¹⁴ Centers for Medicare & Medicaid Services. Medicare Shared Savings Program Saves Medicare More Than \$1.8 Billion in 2022 and Continues to Deliver High-quality Care. 2023. <https://www.cms.gov/newsroom/press-releases/medicare-shared-savings-program-saves-medicare-more-18-billion-2022-and-continues-deliver-high>
- ²¹⁵ Urban Institute. The Medicare Advantage Quality Bonus Program: High Cost for Uncertain Gain. June 2023. <https://www.urban.org/sites/default/files/2023-06/The%20Medicare%20Advantage%20Quality%20Bonus%20Program.pdf>
- ²¹⁶ Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. *mckinsey.com*. July 2018. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>
- ²¹⁷ Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. *mckinsey.com*. July 2018. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>
- ²¹⁸ Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. *mckinsey.com*. July 2018. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>
- ²¹⁹ Urban Institute. The Medicare Advantage Quality Bonus Program: High Cost for Uncertain Gain. June 2023. <https://www.urban.org/sites/default/files/2023-06/The%20Medicare%20Advantage%20Quality%20Bonus%20Program.pdf>
- ²²⁰ Gifford K, Eileen E, Edwards BC, et al. States focus on quality and outcomes amid waiver changes. Henry J. Kaiser Family Foundation. 2019. <https://www.kff.org/medicaid/report/states-focus-on-quality-and-outcomes-amid-waiver-changes-results-from-a-50-state-medicaid-budget-survey-for-state-fiscal-years-2018-and-2019/>
- ²²¹ Kessell E, Pegany V, Keolanui B, et al. Review of Medicare, Medicaid, and commercial quality of care measures: considerations for assessing accountable care organizations. *Journal of Health Politics, Policy and Law*. 2015;40(4):761-96. Doi: 10.1215/03616878-3150050
- ²²² Chen AH, Lee PV. California’s Marketplace Innovations: Driving Health Plan Accountability For Quality And Equity. *Health Affairs Forefront*. 2022. Doi: 10.1377/forefront.20220928.429170
- ²²³ Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine*. 2022;37(2):359. doi:10.1007/s11606-021-06617-8
- ²²⁴ Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev*. 2017;2017(3):CD011153. doi:10.1002/14651858.CD011153.pub2
- ²²⁵ Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev*. 2017;2017(3):CD011153. doi:10.1002/14651858.CD011153.pub2
- ²²⁶ Pirritano M, Miller Parrish K, Kim Y, Solomon H, Keene J. It takes quality improvement to cross the chasm. *BMJ Open Qual*. 2023;12(3):e001906. doi:10.1136/bmjopen-2022-001906
- ²²⁷ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med*. 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0
- ²²⁸ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med*. 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0
- ²²⁹ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med*. 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0
- ²³⁰ Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database Syst Rev*. 2021(1):CD011865. doi:10.1002/14651858.CD011865.pub2

-
- ²³¹ Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database Syst Rev.* 2021(1):CD011865. doi:10.1002/14651858.CD011865.pub2
- ²³² Pirritano M, Miller Parrish K, Kim Y, Solomon H, Keene J. It takes quality improvement to cross the chasm. *BMJ Open Qual.* 2023;12(3):e001906. doi:10.1136/bmjoc-2022-001906
- ²³³ Pirritano M, Miller Parrish K, Kim Y, Solomon H, Keene J. It takes quality improvement to cross the chasm. *BMJ Open Qual.* 2023;12(3):e001906. doi:10.1136/bmjoc-2022-001906
- ²³⁴ Hamilton FL, Greaves F, Majeed A, Millett C. Effectiveness of providing financial incentives to healthcare professionals for smoking cessation activities: systematic review. *Tob Control.* 2013;22(1):3-8. doi:10.1136/tobaccocontrol-2011-050048
- ²³⁵ Benabbas R, Shan G, Akindutire O, Mehta N, Sinert R. The Effect of Pay-for-Performance Compensation Model Implementation on Vaccination Rate: A Systematic Review. *Quality Management in Healthcare.* 2019;28(3):155. doi:10.1097/QMH.0000000000000219
- ²³⁶ Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine.* 2022;37(2):359. doi:10.1007/s11606-021-06617-8
- ²³⁷ Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database Syst Rev.* 2021;2021(1):CD011865. doi:10.1002/14651858.CD011865.pub2
- ²³⁸ Huang J, Yin S, Lin Y, Jiang Q, He Y, Du L. Impact of pay-for-performance on management of diabetes: a systematic review. *Journal of Evidence-Based Medicine.* 2013;6(3):173-184. doi:10.1111/jebm.12052
- ²³⁹ Rhee C, Strich JR, Chiotos K, et al. Improving Sepsis Outcomes in the Era of Pay-for-Performance and Electronic Quality Measures: A Joint IDSA/ACEP/PIDS/SHEA/SHM/SIDP Position Paper. *Clin Infect Dis.* Published online October 13, 2023:ciad447. doi:10.1093/cid/ciad447
- ²⁴⁰ Rhee C, Strich JR, Chiotos K, et al. Improving Sepsis Outcomes in the Era of Pay-for-Performance and Electronic Quality Measures: A Joint IDSA/ACEP/PIDS/SHEA/SHM/SIDP Position Paper. *Clin Infect Dis.* Published online October 13, 2023:ciad447. doi:10.1093/cid/ciad447
- ²⁴¹ Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine.* 2022;37(2):359. doi:10.1007/s11606-021-06617-8
- ²⁴² Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev.* 2017;2017(3):CD011153. doi:10.1002/14651858.CD011153.pub2
- ²⁴³ Yuan B, He L, Meng Q, Jia L. Payment methods for outpatient care facilities. *Cochrane Database Syst Rev.* 2017;2017(3):CD011153. doi:10.1002/14651858.CD011153.pub2
- ²⁴⁴ Petersen LA, Simpson K, Pietz K, et al. Effects of Individual Physician-Level and Practice-Level Financial Incentives on Hypertension Care: A Randomized Trial. *JAMA.* 2013;310(10):1042-1050. doi:10.1001/jama.2013.276303
- ²⁴⁵ Huang J, Yin S, Lin Y, Jiang Q, He Y, Du L. Impact of pay-for-performance on management of diabetes: a systematic review. *Journal of Evidence-Based Medicine.* 2013;6(3):173-184. doi:10.1111/jebm.12052
- ²⁴⁶ Bardach NS, Wang JJ, De Leon SF, Shih SC, Boscardin WJ, Goldman LE, Dudley RA. Effect of pay-for-performance incentives on quality of care in small practices with electronic health records: A randomized trial. *JAMA.* 2013;310(10):1051-9. doi:10.1001/jama.2013.277353
- ²⁴⁷ Huang J, Yin S, Lin Y, Jiang Q, He Y, Du L. Impact of pay-for-performance on management of diabetes: a systematic review. *Journal of Evidence-Based Medicine.* 2013;6(3):173-184. doi:10.1111/jebm.12052
- ²⁴⁸ Rhee C, Strich JR, Chiotos K, et al. Improving Sepsis Outcomes in the Era of Pay-for-Performance and Electronic Quality Measures: A Joint IDSA/ACEP/PIDS/SHEA/SHM/SIDP Position Paper. *Clin Infect Dis.* Published online October 13, 2023:ciad447. doi:10.1093/cid/ciad447
- ²⁴⁹ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). doi:10.1056/CAT.18.0245
- ²⁵⁰ Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open.* 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115

-
- ²⁵¹ Mathes T, Pieper D, Morche J, Polus S, Jaschinski T, Eikermann M. Pay for performance for hospitals. *Cochrane Database Syst Rev.* 2019;2019(7):CD011156. doi:10.1002/14651858.CD011156.pub2
- ²⁵² Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open.* 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ²⁵³ Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine.* 2022;37(2):359. doi:10.1007/s11606-021-06617-8
- ²⁵⁴ Burstein DS, Liss DT, Linder JA. Association of Primary Care Physician Compensation Incentives and Quality of Care in the United States, 2012-2016. *Journal of General Internal Medicine.* 2022;37(2):359. doi:10.1007/s11606-021-06617-8
- ²⁵⁵ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med.* 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0
- ²⁵⁶ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med.* 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0
- ²⁵⁷ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). doi:10.1056/CAT.18.0245
- ²⁵⁸ Kim KM, White JS, Max W, Chapman SA, Muench U. Evaluation of Clinical and Economic Outcomes Following Implementation of a Medicare Pay-for-Performance Program for Surgical Procedures. *JAMA Netw Open.* 2021;4(8):e2121115. doi:10.1001/jamanetworkopen.2021.21115
- ²⁵⁹ Beauvis B, Whitaker Z, Kim F, Anderson B. Is the Hospital Value-Based Purchasing Program Associated with Reduced Hospital Readmissions? *J Multidiscip Healthc.* 2022;15: 1089-1099 doi: 10.2147/JMDH.S358733
- ²⁶⁰ Mathes T, Pieper D, Morche J, Polus S, Jaschinski T, Eikermann M. Pay for performance for hospitals. *Cochrane Database Syst Rev.* 2019;2019(7):CD011156. doi:10.1002/14651858.CD011156.pub2
- ²⁶¹ Jia L, Meng Q, Scott A, Yuan B, Zhang L. Payment methods for healthcare providers working in outpatient healthcare settings. *Cochrane Database Syst Rev.* 2021;2021(1):CD011865. doi:10.1002/14651858.CD011865.pub2
- ²⁶² Mendelson A, Kondo K, Damberg C, et al. The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care. *Ann Intern Med.* 2017;166(5):341-353. doi:10.7326/M16-1881
- ²⁶³ Ogundeji YK, Bland JM, Sheldon TA. The effectiveness of payment for performance in health care: A meta-analysis and exploration of variation in outcomes. *Health Policy.* 2016;120(10):1141-1150. doi:10.1016/j.healthpol.2016.09.002
- ²⁶⁴ Health Care Payment Learning & Action Network. Alternative Payment Model Framework. 2017. <https://hcp-lan.org/workproducts/apm-refresh-whitepaper-final.pdf>
- ²⁶⁵ Kazungu JS, Barasa EW, Obadha M, Chuma J. What characteristics of provider payment mechanisms influence health care providers' behaviour? A literature review. *Int J Health Plann Manage.* 2018 Oct 1; 33(4): e892–e905. Published online 2018 Jul 8. doi: [10.1002/hpm.2565](https://doi.org/10.1002/hpm.2565)
- ²⁶⁶ Kim KM, Max W, White JS, Chapman SA, Muench U. Do penalty-based pay-for-performance programs improve surgical care more effectively than other payment strategies? A systematic review. *Ann Med Surg (Lond).* 2020;60:623-630. doi:10.1016/j.amsu.2020.11.060
- ²⁶⁷ Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med.* 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8
- ²⁶⁸ Kondo KK, Damberg CL, Mendelson A, et al. Implementation Processes and Pay for Performance in Healthcare: A Systematic Review. *J Gen Intern Med.* 2016;31(Suppl 1):61-69. doi:10.1007/s11606-015-3567-0
- ²⁶⁹ Ogundeji YK, Bland JM, Sheldon TA. The effectiveness of payment for performance in health care: A meta-analysis and exploration of variation in outcomes. *Health Policy.* 2016;120(10):1141-1150. doi:10.1016/j.healthpol.2016.09.002
- ²⁷⁰ Scott A, Liu M, Yong J. Financial Incentives to Encourage Value-Based Health Care. *Medical Care Research and Review.* 2018;75(1):3-32. doi:10.1177/1077558716676594

-
- ²⁷¹ Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med*. 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8
- ²⁷² Scott A, Liu M, Yong J. Financial Incentives to Encourage Value-Based Health Care. *Medical Care Research and Review*. 2018;75(1):3-32. doi:10.1177/1077558716676594
- ²⁷³ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). Doi:10.1056/CAT.18.0245
- ²⁷⁴ Conway A, Satin D. The role of pay-for-performance in reducing healthcare disparities: A narrative literature review. *Prev Med*. 2022;164:107274. Doi:10.1016/j.yjmed.2022.107274
- ²⁷⁵ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). Doi:10.1056/CAT.18.0245
- ²⁷⁶ Bailey V. Stratifying Hospital Readmissions Reduction Program Improved Equity. *Value-Based Care News*. RevCycle Intelligence. <https://revcycleintelligence.com/news/stratifying-hospital-readmissions-reduction-program-improved-equity>
- ²⁷⁷ Conway A, Satin D. The role of pay-for-performance in reducing healthcare disparities: A narrative literature review. *Prev Med*. 2022;164:107274. Doi:10.1016/j.yjmed.2022.107274
- ²⁷⁸ Centers for Medicare & Medicaid Services. 2024 National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report. Baltimore, MD: U.S. Department of Health and Human Services; 2024. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-AssessmentInstruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports>
- ²⁷⁹ Conway A, Satin D. The role of pay-for-performance in reducing healthcare disparities: A narrative literature review. *Prev Med*. 2022;164:107274. Doi:10.1016/j.yjmed.2022.107274
- ²⁸⁰ NEJM Catalyst. What Is Pay for Performance in Healthcare? Catalyst Carryover. 2018;4(2). Doi:10.1056/CAT.18.0245
- ²⁸¹ Naessens JM, Van Such MB, Nesse RE, et al. Looking Under the Streetlight? A Framework for Differentiating Performance Measures by Level of Care in a Value-Based Payment Environment. *Acad Med*. 2017;92(7):943-950. Doi:10.1097/ACM.0000000000001654
- ²⁸² Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med*. 2017;32(11):1249-1254. Doi:10.1007/s11606-017-4125-8
- ²⁸³ Centers for Medicare & Medicaid Services. Measure Specification. October 2023. <https://mmshub.cms.gov/measure-lifecycle/measure-specification/develop-specification/denominator>
- ²⁸⁴ Grants NRC (US) P on PM and D for PHPP, Perrin EB, Durch JS, Skillman SM. Data and Information Systems: Issues for Performance Measurement. In: *Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network*. National Academies Press (US); 1999. Accessed November 30, 2023. <https://www.ncbi.nlm.nih.gov/books/NBK231008/>
- ²⁸⁵ Agency for Healthcare Research and Quality. Data Sources for Health Care Quality Measures. July 2018. <https://www.ahrq.gov/talkingquality/measures/understand/index.html>.
- ²⁸⁶ Centers for Medicare & Medicaid Services. Data Sources for Quality Measurement. Supplement Material to the CMS MMS Blueprint. September 2020. <https://www.cms.gov/files/document/blueprint-data-sources.pdf>
- ²⁸⁷ National Academies of Sciences, Engineering, and Medicine. Vital Signs: Core Metrics for Health and Health Care Progress. Washington, DC: The National Academies Press. 2015. <https://doi.org/10.17226/19402>.
- ²⁸⁸ Chen MM, Rosenkrantz AB, Nicola GN, et al. The Qualified Clinical Data Registry: A Pathway to Success within MACRA. *AJNR Am J Neuroradiol*. 2017. 38(7):1292-1296. doi:10.3174/ajnr.A5220
- ²⁸⁹ Chen MM, Rosenkrantz AB, Nicola GN, et al. The Qualified Clinical Data Registry: A Pathway to Success within MACRA. *AJNR Am J Neuroradiol*. 2017. 38(7):1292-1296. doi:10.3174/ajnr.A5220
- ²⁹⁰ Centers for Medicare & Medicaid Services. Measure Specification. October 2023. <https://mmshub.cms.gov/measure-lifecycle/measure-specification/develop-specification/denominatorGrants>

-
- ²⁹¹ Perrin EB, Durch JS, Skillman SM. Data and Information Systems: Issues for Performance Measurement. In: Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network. National Academies Press (US); 1999. <https://www.ncbi.nlm.nih.gov/books/NBK231008/>
- ²⁹² Agency for Healthcare Research and Quality. Data Source for Healthcare Quality Measures. July 2018. <https://www.ahrq.gov/talkingquality/measures/understand/index.html>
- ²⁹³ Redman TC. To Improve Data Quality, Start at the Source. Harvard Business Review. Published online February 10, 2020. <https://hbr.org/2020/02/to-improve-data-quality-start-at-the-source>
- ²⁹⁴ Centers for Medicare & Medicaid Services. Data Highlight. April 2022. <https://www.cms.gov/files/document/data-highlight-april-2022.pdf>
- ²⁹⁵ Centers for Medicare & Medicaid Services. 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>
- ²⁹⁶ American Academy of Family Physicians. New “Z codes” and a tool to help physicians use them. October 2, 2023. <https://www.aafp.org/pubs/fpm/blogs/gettingpaid/entry/sdoh-codes-info.html>
- ²⁹⁷ Dusetzina SB, Tyree S, Meyer AM, et al. Linking Data for Health Services Research: A Framework and Instructional Guide. Rockville (MD): Agency for Healthcare Research and Quality. September 2014. <https://www.ncbi.nlm.nih.gov/books/NBK253315/>
- ²⁹⁸ Mylod DE, Lee TH. Fixing Data Overload in Health Care. Harvard Business Review. March 16, 2022. <https://hbr.org/2022/03/fixing-data-overload-in-health-care>
- ²⁹⁹ Taylor KI, Staunton H, Lipsmeier F, Nobbs D, Lindemann M. Outcome measures based on digital health technology sensor data: data- and patient-centric approaches. npj Digit Med. 2020. 3(1):1-8. doi:10.1038/s41746-020-0305-8
- ³⁰⁰ Nouri SS, Adler-Milstein J, Thao C, et al. Patient characteristics associated with objective measures of digital health tool use in the United States: A literature review. Journal of the American Medical Informatics Association. 2020;27(5):834-841. doi:10.1093/jamia/ocaa024
- ³⁰¹ Ratitch B, Trigg A, Majumder M, Vlajnic V, Rethemeier N, Nkulikiyinka R. Clinical Validation of Novel Digital Measures: Statistical Methods for Reliability Evaluation. Digital Biomarkers. 2023. 7(1):74-91. doi:10.1159/000531054
- ³⁰² Duarte M, Pereira-Rodrigues P, Ferreira-Santos D. The Role of Novel Digital Clinical Tools in the Screening or Diagnosis of Obstructive Sleep Apnea: Systematic Review. Journal of Medical Internet Research. 2023. 25(1):e47735. doi:10.2196/47735
- ³⁰³ Moschella, Susan. MIPS eCOMs vs Registry Measures: What’s the difference? Medisolv. 2019. <https://blog.medisolv.com/articles/mips-ecqms-vs-registry-measures-whats-the-difference-2019>
- ³⁰⁴ Office of the National Coordinator for Health Information Technology. What is FHIR? ONC Fact Sheet. <https://www.healthit.gov/sites/default/files/2019-08/ONCFHIRFSWhatIsFHIR.pdf>.
- ³⁰⁵ Berenson RA, Shartz A, Pham HH. Beyond demonstrations: implementing a primary care hybrid payment model in Medicare. Health Affairs Scholar. 2023. 1(2):qxad024. doi:10.1093/haschl/qxad024
- ³⁰⁶ Skopec L, Bowen G, Zuckerman S. Accounting for Social Risk in Value-Based Payment and Quality Measurement. Urban Institute; 2022. <https://www.urban.org/research/publication/accounting-social-risk-value-based-payment-and-quality-measurement>
- ³⁰⁷ Zhu M, Japinga M, Saunders R, McClellan M. The Future of Risk Adjustment: Supporting Equitable, Comprehensive Health Care. The Duke Margolis Center for Health Policy. 2022. https://healthpolicy.duke.edu/sites/default/files/2022-06/Margolis%20Future%20Risk%20Adjustment%20Paper%20v3_0.pdf
- ³⁰⁸ Zhu M, Japinga M, Saunders R, McClellan M. The Future of Risk Adjustment: Supporting Equitable, Comprehensive Health Care. The Duke Margolis Center for Health Policy. 2022. https://healthpolicy.duke.edu/sites/default/files/2022-06/Margolis%20Future%20Risk%20Adjustment%20Paper%20v3_0.pdf
- ³⁰⁹ Antonucci, J. Proposal- An innovative model for primary care office payment. PTAC Proposal. April 2018. <https://aspe.hhs.gov/sites/default/files/private/pdf/255906/ProposalAntonucci.pdf>

-
- ³¹⁰ Centers for Medicare & Medicaid Services. Report of the Council on Medical Service, Improving Risk Adjustment in Alternative Payment Models, CMS Report. March 2019, <https://www.ama-assn.org/system/files/2019-12/i19-cms-report3-apms.pdf> .
- ³¹¹ Centers for Medicare & Medicaid Services. Consumer Assessment of Healthcare Providers & Systems (CAHPS). CMS.gov. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems>
- ³¹² Agency for Healthcare Research and Quality. About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/about-cahps/index.html>
- ³¹³ Agency for Healthcare Research and Quality. About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/about-cahps/index.html>
- ³¹⁴ Agency for Healthcare Research and Quality. CAHPS Patient Experience Surveys and Guidance. Agency for Healthcare Research and Quality <https://www.ahrq.gov/cahps/surveys-guidance/index.html>
- ³¹⁵ Agency for Healthcare Research and Quality. Frequently Asked Questions About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/faq/index.html#:~:text=CMS%20requires%20or%20administers%20several,outpatient%20and%20ambulatory%20surgery%20centers.>
- ³¹⁶ Agency for Healthcare Research and Quality. Frequently Asked Questions About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/faq/index.html#:~:text=CMS%20requires%20or%20administers%20several,outpatient%20and%20ambulatory%20surgery%20centers.>
- ³¹⁷ Agency for Healthcare Research and Quality. Understanding CAHPS Surveys: A Primer for New Users. AHRQ. January 2019. <https://www.ahrq.gov/sites/default/files/wysiwyg/cahps/news-and-events/events/webinars/understanding-cahps-101-infographic.pdf>
- ³¹⁸ Agency for Healthcare Research and Quality. CAHPS Patient Experience Surveys and Guidance. AHRQ.gov <https://www.ahrq.gov/cahps/surveys-guidance/index.html>
- ³¹⁹ Centers for Medicare & Medicaid Services. Consumer Assessment of Healthcare Providers & Systems (CAHPS). CMS.gov. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems>
- ³²⁰ Agency for Healthcare Research and Quality. CAHPS Patient Experience Surveys and Guidance. Agency for Healthcare Research and Quality <https://www.ahrq.gov/cahps/surveys-guidance/index.html>
- ³²¹ Agency for Healthcare Research and Quality. Frequently Asked Questions About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/faq/index.html#:~:text=CMS%20requires%20or%20administers%20several,outpatient%20and%20ambulatory%20surgery%20centers.>
- ³²² Agency for Healthcare Research and Quality. Frequently Asked Questions About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/faq/index.html#:~:text=CMS%20requires%20or%20administers%20several,outpatient%20and%20ambulatory%20surgery%20centers.>
- ³²³ Agency for Healthcare Research and Quality. Frequently Asked Questions About CAHPS. AHRQ.gov. <https://www.ahrq.gov/cahps/faq/index.html#:~:text=CMS%20requires%20or%20administers%20several,outpatient%20and%20ambulatory%20surgery%20centers.>
- ³²⁴ Agency for Healthcare Research and Quality. CAHPS Patient Experience Surveys and Guidance. Agency for Healthcare Research and Quality <https://www.ahrq.gov/cahps/surveys-guidance/index.html>
- ³²⁵ Agency for Healthcare Research and Quality. CAHPS Surgical Care Survey. AHRQ.gov. Content last reviewed February 2018. <https://www.ahrq.gov/cahps/surveys-guidance/surgical/index.html>
- ³²⁶ CAHPS American Indian Survey. Content last reviewed May 2022. Agency for Healthcare Research and Quality, Rockville, MD. <https://www.ahrq.gov/cahps/surveys-guidance/american-indian/index.html>
- ³²⁷ Agency for Healthcare Research and Quality. CAHPS Cancer Care Survey. AHRQ.gov. Content last reviewed July 2020. <https://www.ahrq.gov/cahps/surveys-guidance/cancer/index.html>
- ³²⁸ Agency for Healthcare Research and Quality. CAHPS Mental Health Care Surveys. AHRQ.gov. Content last reviewed May 2022. <https://www.ahrq.gov/cahps/surveys-guidance/echo/index.html>
- ³²⁹ Agency for Healthcare Research and Quality. CAHPS Health Plan Survey. AHRQ.gov. Content last reviewed January 2023. <https://www.ahrq.gov/cahps/surveys-guidance/hp/index.html>
- ³³⁰ Agency for Healthcare Research and Quality. CAHPS Dental Plan Survey. AHRQ.gov. Content last reviewed March 2018. <https://www.ahrq.gov/cahps/surveys-guidance/dental/index.html>

-
- ³³¹ Agency for Healthcare Research and Quality. CAHPS Emergency Department Survey. AHRQ.gov. Content last reviewed December 2023. <https://www.ahrq.gov/cahps/surveys-guidance/ed-cahps/index.html>
- ³³² Agency for Healthcare Research and Quality. CAHPS Child Hospital Survey. AHRQ.gov. Content last reviewed December 2023. https://www.ahrq.gov/cahps/surveys-guidance/hospital/about/child_hp_survey.html
- ³³³ Agency for Healthcare Research and Quality. CAHPS Nursing Home Surveys. AHRQ.gov. Content last reviewed August 2023. <https://www.ahrq.gov/cahps/surveys-guidance/nh/index.html>
- ³³⁴ Agency for Healthcare Research and Quality. CAHPS Patient Experience Surveys and Guidance. AHRQ.gov. Content last reviewed August 2023.
- ³³⁵ Agency for Healthcare Research and Quality. CAHPS Home Health Care Survey. AHRQ.gov. Content last reviewed November 2023. <https://www.ahrq.gov/cahps/surveys-guidance/home/index.html>
- ³³⁶ Agency for Healthcare Research and Quality. CAHPS Home and Community-Based Services Survey. AHRQ.gov. Content last reviewed November 2023. <https://www.ahrq.gov/cahps/surveys-guidance/hcbs/index.html>
- ³³⁷ Centers for Medicare & Medicaid Services. Fee-for-Service (FFS) CAHPS. Cms.gov. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems/fee-service-cahps#:~:text=Overview%3A%20The%20Medicare%20Fee%2Dfor,Medicare%20and%20their%20FFS%20provider.>
- ³³⁸ Centers for Medicare & Medicaid Services. Medicare Advantage and Prescription Drug Plan CAHPS (MA and PDP CAHPS). Cms.gov. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems/medicare-advantage-and-prescription-drug-plan-cahps>
- ³³⁹ Agency for Healthcare Research and Quality. CAHPS In-Center Hemodialysis Survey. AHRQ.gov. Content last reviewed November 2023., Rockville, MD. <https://www.ahrq.gov/cahps/surveys-guidance/ich/index.html>
- ³⁴⁰ Centers for Medicare & Medicaid Services. Nationwide Adult Medicaid Consumer Assessment of Healthcare Providers & Systems. CMS.gov. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems/nationwide-adult-medicaid-cahps>
- ³⁴¹ Centers for Medicare & Medicaid Services. CAHPS Hospice Survey. www.hospicecahpsurvey.org
- ³⁴² Agency for Healthcare Research and Quality. CAHPS Outpatient and Ambulatory Surgery Survey. AHRQ.gov. November 2023. <https://www.ahrq.gov/cahps/surveys-guidance/oas/index.html>
- ³⁴³ Centers for Medicare & Medicaid Services. CAHPS for MIPS Survey. CMS. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems/cahps-mips>
- ³⁴⁴ Centers for Medicare & Medicaid Services. Emergency Department CAHPS. CMS. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems/emergency-department-cahps>
- ³⁴⁵ Bland C, Zuckerbraun S, Lines LM, et al. Challenges Facing CAHPS Surveys and Opportunities for Modernization. Research Triangle Park (NC): RTI Press. November 2022. <https://www.ncbi.nlm.nih.gov/books/NBK592584/doi:10.3768/rtipress.2022.op.0080.2211>
- ³⁴⁶ Bland C, Zuckerbraun S, Lines LM, et al. Challenges Facing CAHPS Surveys and Opportunities for Modernization. Research Triangle Park (NC): RTI Press. November 2022. <https://www.ncbi.nlm.nih.gov/books/NBK592584/doi:10.3768/rtipress.2022.op.0080.2211>
- ³⁴⁷ Krzyewski K. Payment Model Features Contributing to Successful Population-Based Total Cost of Care Models. PTAC Public Meeting Listening Session Presentation. September 2019. <https://aspe.hhs.gov/sites/default/files/documents/1c798bde8b881d82ae5b994368b3a/PTAC-Sep-19-SME-LS-Slides.pdf>
- ³⁴⁸ Krzyewski K. Payment Model Features Contributing to Successful Population-Based Total Cost of Care Models. PTAC Public Meeting Listening Session Presentation. September 2019. <https://aspe.hhs.gov/sites/default/files/documents/1c798bde8b881d82ae5b994368b3a/PTAC-Sep-19-SME-LS-Slides.pdf>
- ³⁴⁹ Agency for Healthcare Research and Quality. Consumer Assessment of Healthcare Providers and Systems (CAHPS). AHRQ. <https://www.ahrq.gov/cahps/index.html>
- ³⁵⁰ Center for Medicare and Medicaid Innovation. Person-Centered Innovation – An Update on the Implementation of the CMS Innovation Center’s Strategy – Supplemental Document. Centers for Medicare & Medicaid Services. 2022. <https://innovation.cms.gov/data-and-reports/2022/cmimi-strategy-refresh-imp-tech-report>
- ³⁵¹ <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems>

-
- ³⁵² Health Care Payment Learning & Action Network. *Accelerating and Aligning Population-Based Payment Models: Performance Measurement*. HCPLAN; 2016. <https://hcp-lan.org/pm-whitepaper/>, <https://hcp-lan.org/pm-whitepaper/>
- ³⁵³ National Quality Forum. Measure Evaluation Criteria. NQF: n.d https://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx
- ³⁵⁴ Krishnan M, Brunelli SM, Maddux FW, et al. Guiding principles and checklist for population-based quality metrics. *Clin J Am Soc Nephrol*. 2014;9(6):1124-1131. doi:[10.2215/CJN.11061013](https://doi.org/10.2215/CJN.11061013)
- ³⁵⁵ American Medical Association. A Guide to Physician-Focused Alternative Payment Models. 2015. <https://www.ama-assn.org/system/files/corp/media-browser/specialty%20group/washington/alternative-payment-models-physician-guide.pdf>
- ³⁵⁶ Machta R, Peterson G, Rotter J, et al. Evaluation of the Maryland Total Cost of Care Model: Implementation Report. Mathematica; 2021. <https://www.mathematica.org/publications/evaluation-of-the-maryland-total-cost-of-care-model-implementation-report>
- ³⁵⁷ Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med*. 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8
- ³⁵⁸ Joynt Maddox KE, Sen AP, Samson LW, Zuckerman RB, DeLew N, Epstein AM. Elements of Program Design in Medicare's Value-based and Alternative Payment Models: a Narrative Review. *J Gen Intern Med*. 2017;32(11):1249-1254. doi:10.1007/s11606-017-4125-8
- ³⁵⁹ Blumenthal D, Malphrus E, McGinnis JM. Vital Signs: Core Metrics for Health and Health Care Progress. Committee on Core Metrics for Better Health at Lower Cost, Institute of Medicine. National Academies Press (US); 2015. <http://www.ncbi.nlm.nih.gov/books/NBK316120/>
- ³⁶⁰ Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS — The Universal Foundation. *New England Journal of Medicine*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539
- ³⁶¹ Richard C, Urick BY, Pathak S, Jackson J, Livet M. Performance-based pharmacy payment models: key components and critical implementation considerations for successful uptake and integration. *J Manag Care Spec Pharm*. 2021;27(11):10.18553/jmcp.2021.27.11.1568. doi:10.18553/jmcp.2021.27.11.1568
- ³⁶² Richard C, Urick BY, Pathak S, Jackson J, Livet M. Performance-based pharmacy payment models: key components and critical implementation considerations for successful uptake and integration. *J Manag Care Spec Pharm*. 2021;27(11):10.18553/jmcp.2021.27.11.1568. doi:10.18553/jmcp.2021.27.11.1568
- ³⁶³ Afendulis CC, Fendrick AM, Song Z, et al. The Impact of Global Budgets on Pharmaceutical Spending and Utilization. *Inquiry*. 2014;51:0046958014558716. doi:10.1177/0046958014558716
- ³⁶⁴ Zhang, Hui, David W. Cowling, Joanne M. Graham, and Erik Taylor. Impact of a Commercial Accountable Care Organization on Prescription Drugs. *Health Services Research* 56, no. 4. August 2021. : 592–603. <https://doi.org/10.1111/1475-6773.13626>.
- ³⁶⁵ Cuellar A, Helmchen LA, Gimm G, et al. The CareFirst Patient-Centered Medical Home Program: Cost and Utilization Effects in Its First Three Years. *J Gen Intern Med*. 2016;31(11):1382-1388. doi:10.1007/s11606-016-3814-z
- ³⁶⁶ Mentias A, Peterson ED, Keshvani N, et al. Achieving Equity in Hospital Performance Assessments Using Composite Race-Specific Measures of Risk-Standardized Readmission and Mortality Rates for Heart Failure. *Circulation*. 2023;147(15):1121-1133. doi:10.1161/CIRCULATIONAHA.122.061995
- ³⁶⁷ Heidenreich P, Sandhu A. Pursuing Equity in Performance Measurement. *Circulation*. 2023;147(15):1134-1136. doi:10.1161/CIRCULATIONAHA.123.064123
- ³⁶⁸ Anderson AC, O'Rourke E, Chin MH, Ponce NA, Bernheim SM, Burstin H. Promoting Health Equity And Eliminating Disparities Through Performance Measurement And Payment. *Health Aff (Millwood)*. 2018;37(3):371-377. doi:10.1377/hlthaff.2017.1301
- ³⁶⁹ Skopec L, Bowen G, Zuckerman S. Accounting for Social Risk in Value-Based Payment and Quality Measurement. Urban Institute; 2022. <https://www.urban.org/research/publication/accounting-social-risk-value-based-payment-and-quality-measurement>

-
- ³⁷⁰ MedPAC. March 2022 Report to the Congress: Medicare Payment Policy. <https://www.medpac.gov/document/march-2022-report-to-the-congress-medicare-payment-policy/>
- ³⁷¹ MedPAC. June 2018 Report to the Congress: Medicare and the Health Care Delivery System. <https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-jun18-medpacreporttocongress-rev-nov2019-note-sec-pdf/>
- ³⁷² Centers for Medicare & Medicaid Services. Hospital Readmissions Reduction Program (HRRP) Overview. <https://qualitynet.cms.gov/inpatient/hrrp>
- ³⁷³ National Quality Forum. Social Risk Trial Final Report. NQF; 2017. https://www.qualityforum.org/Publications/2017/07/Social_Risk_Trial_Final_Report.aspx
- ³⁷⁴ National Quality Forum. Social Risk Trial Final Report. NQF; 2017. https://www.qualityforum.org/Publications/2017/07/Social_Risk_Trial_Final_Report.aspx
- ³⁷⁵ Skopec L, Bowen G, Zuckerman S. Accounting for Social Risk in Value-Based Payment and Quality Measurement. Urban Institute; 2022. <https://www.urban.org/research/publication/accounting-social-risk-value-based-payment-and-quality-measurement>
- ³⁷⁶ Assistant Secretary for Planning and Evaluation. Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. June 2020. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>
- ³⁷⁷ Assistant Secretary for Planning and Evaluation. Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. June 2020. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>
- ³⁷⁸ Assistant Secretary for Planning and Evaluation. Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. June 2020. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>
- ³⁷⁹ Rawal P, Jacobs D, Fowler E, Seshamani M. Building On CMS's Accountable Care Vision To Improve Care For Medicare Beneficiaries. Health Affairs Forefront. doi:10.1377/forefront.20230727.802728
- ³⁸⁰ National Quality Forum. Attribution - Principles and Approaches. 2016. https://www.qualityforum.org/Publications/2016/12/Attribution_-_Principles_and_Approaches.aspx
- ³⁸¹ National Quality Forum. Attribution - Principles and Approaches. 2016. https://www.qualityforum.org/Publications/2016/12/Attribution_-_Principles_and_Approaches.aspx
- ³⁸² Reeves S, Dombkowski K, Madden B, et al. Considerations when Aggregating Data to Measure Performance across Levels of the Healthcare System. Acad Pediatr. 2022;22(3 Suppl):S119-S124. doi:10.1016/j.acap.2021.11.013
- ³⁸³ Qaseem A, Mount CA, Campos K, et al. Quality Indicators for Osteoporosis in Adults: A Review of Performance Measures by the American College of Physicians. Ann Intern Med. 2023;176(10):1386-1391. doi:10.7326/M23-1291
- ³⁸⁴ Conlin PR, Zhang L, Li D, Nelson RE, Prentice JC, Mohr DC. Association of hemoglobin A1c stability with mortality and diabetes complications in older adults with diabetes. BMJ Open Diabetes Res Care. 2023;11(2):e003211. doi:10.1136/bmjdr-2022-003211
- ³⁸⁵ Kristensen PK, Thillemann TM, Søballe K, Johnsen SP. Are process performance measures associated with clinical outcomes among patients with hip fractures? A population-based cohort study. International Journal for Quality in Health Care. 2016;28(6):698-708. doi:10.1093/intqhc/mzw093
- ³⁸⁶ Skopec L, Berenson R. The Medicare Advantage Quality Bonus Program. Urban Institute; 2023. <https://www.urban.org/research/publication/medicare-advantage-quality-bonus-program>.
- ³⁸⁷ McWilliams JM. Professionalism Revealed: Rethinking Quality Improvement in the Wake of a Pandemic. NEJM Catalyst. 2020;1(5). doi:10.1056/CAT.20.0226
- ³⁸⁸ Centers for Medicare & Medicaid Services. 2024 National Impact Assessment of the Centers for Medicare & Medicaid Services (CMS) Quality Measures Report. Baltimore, MD: U.S. Department of Health and Human Services; 2024. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-AssessmentInstruments/QualityMeasures/National-Impact-Assessment-of-the-Centers-for-Medicare-and-Medicaid-Services-CMS-Quality-Measures-Reports>

-
- ³⁸⁹ Burke LG, Orav EJ, Zheng J, Jha AK. Healthy Days at home: A novel population-based outcome measure. *Healthcare*. 2020;8(1):100378. doi:10.1016/j.hjdsi.2019.100378
- ³⁹⁰ Le Corroller AG, Bonastre J. Patient-reported measures: how useful in health economics? *Eur J Health Econ*. 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z
- ³⁹¹ Basch E, Spertus J, Adams Dudley R, et al. Methods for Developing Patient-Reported Outcome-Based Performance Measures (PRO-PMs). *Value in Health*. 2015;18(4):493-504. doi:10.1016/j.jval.2015.02.018
- ³⁹² Basch E, Spertus J, Adams Dudley R, et al. Methods for Developing Patient-Reported Outcome-Based Performance Measures (PRO-PMs). *Value in Health*. 2015;18(4):493-504. doi:10.1016/j.jval.2015.02.018
- ³⁹³ Salinas J, Sprinkhuizen SM, Ackerson T, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. *Stroke*. 2016;47(1):180-186. doi:10.1161/STROKEAHA.115.010898
- ³⁹⁴ Tew M, Dalziel K, Clarke P, Smith A, Choong PF, Dowsey M. Patient-reported outcome measures (PROMs): can they be used to guide patient-centered care and optimize outcomes in total knee replacement? | SpringerLink. *Qual Life Res*. 2020;29:3273-3283. doi:https://doi.org/10.1007/s11136-020-02577-4
- ³⁹⁵ Makhni EC. Meaningful Clinical Applications of Patient-Reported Outcome Measures in Orthopaedics. *J Bone Joint Surg Am*. 2021;103(1):84-91. doi:10.2106/JBJS.20.00624
- ³⁹⁶ Dudgeon D. The Impact of Measuring Patient-Reported Outcome Measures on Quality of and Access to Palliative Care. *J Palliat Med*. 2017;20(Suppl 1):S-76-S-80. doi:10.1089/jpm.2017.0447
- ³⁹⁷ Austin E, LeRouge C, Hartzler AL, Segal C, Lavalley DC. Capturing the patient voice: implementing patient-reported outcomes across the health system. *Qual Life Res*. 2020;29(2):347-355. doi:10.1007/s11136-019-02320-8
- ³⁹⁸ Greenhalgh J, Dalkin S, Gooding K, et al. Functionality and Feedback: A Realist Synthesis of the Collation, Interpretation and Utilisation of Patient-Reported Outcome Measures Data to Improve Patient Care. *NIHR Journals Library*; 2017. <http://www.ncbi.nlm.nih.gov/books/NBK409450/>
- ³⁹⁹ Le Corroller AG, Bonastre J. Patient-reported measures: how useful in health economics? *Eur J Health Econ*. 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z
- ⁴⁰⁰ Centers for Medicare & Medicaid Services. Consumer Assessment of Healthcare Providers and Systems (CAHPS). September 2023. <https://www.cms.gov/data-research/research/consumer-assessment-healthcare-providers-systems>
- ⁴⁰¹ Shirley ED, Sanders JO. Measuring Quality of Care with Patient Satisfaction Scores. *J Bone Joint Surg Am*. 2016;98(19):e83. doi:10.2106/JBJS.15.01216
- ⁴⁰² Joseph K, Udogwu UN, Manson TT, et al. Patient Satisfaction After Discharge Is Discordant With Reported Inpatient Experience. *Orthopedics*. 2021;44(3):e427-e433. doi:10.3928/01477447-20210415-01
- ⁴⁰³ Kemp E, Trivitt J, Davis C. Evidence-Based Performance Indicators of Positive Inpatient Experiences. *J Healthc Manag*. 2023;68(2):106-120. doi:10.1097/JHM-D-22-00147
- ⁴⁰⁴ Simon L, Robert C, Meyer P. Artificial intelligence for quality assurance in radiotherapy. *Cancer Radiother*. 2021;25(6-7):623-626. doi:10.1016/j.canrad.2021.06.012
- ⁴⁰⁵ Loehfelm TW. Artificial Intelligence for Quality Improvement in Radiology. *Radiol Clin North Am*. 2021;59(6):1053-1062. doi:10.1016/j.rcl.2021.07.005
- ⁴⁰⁶ Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. *Br J Radiol*. 2019;92(1100):20190001. doi:10.1259/bjr.20190001
- ⁴⁰⁷ Sankaran R, Kumar A, Parasuram H. Role of Artificial Intelligence and Machine Learning in the prediction of the pain: A scoping systematic review. *Proc Inst Mech Eng H*. 2022;236(10):1478-1491. doi:10.1177/09544119221122012
- ⁴⁰⁸ Richardson A, Robbins CB, Wisely CE, Henao R, Grewal DS, Fekrat S. Artificial intelligence in dementia. *Curr Opin Ophthalmol*. 2022;33(5):425-431. doi:10.1097/ICU.0000000000000881
- ⁴⁰⁹ Battista P, Salvatore C, Berlingeri M, Cerasa A, Castiglioni I. Artificial intelligence and neuropsychological measures: The case of Alzheimer's disease. *Neurosci Biobehav Rev*. 2020;114:211-228. doi:10.1016/j.neubiorev.2020.04.026
- ⁴¹⁰ Dullabh P, Dhopeswarkar R, Leapart D, et al. Trustworthy Artificial Intelligence (TAI) for Patient-Centered Outcomes Research (PCOR). Assistant Secretary for Planning and Evaluation; 2023. <https://aspe.hhs.gov/reports/trustworthy-ai-pcor>

-
- ⁴¹¹ Rivera SC, Liu X, Hughes SE, et al. Embedding patient-reported outcomes at the heart of artificial intelligence health-care technologies. *The Lancet Digital Health*. 2023;5(3):e168-e173. doi:10.1016/S2589-7500(22)00252-7
- ⁴¹² van Buchem MM, Neve OM, Kant IMJ, Steyerberg EW, Boosman H, Hensen EF. Analyzing patient experiences using natural language processing: development and validation of the artificial intelligence patient reported experience measure (AI-PREM). *BMC Med Inform Decis Mak*. 2022;22(1):183. doi:10.1186/s12911-022-01923-5
- ⁴¹³ Ibrahim H, Liu X, Zariffa N, Morris AD, Denniston AK. Health data poverty: an assailable barrier to equitable digital health care. *The Lancet Digital Health*. 2021;3(4):e260-e265. doi:10.1016/S2589-7500(20)30317-4
- ⁴¹⁴ Centers for Medicare & Medicaid Services. Digital Quality Strategic Roadmap. CMS; 2022. https://ecqi.healthit.gov/sites/default/files/CMSdQMStrategicRoadmap_032822.pdf.
- ⁴¹⁵ Centers for Medicare & Medicaid Services. Lessons Learned from the CMS Artificial Intelligence Health Outcomes Challenge. May 2021. <https://www.cms.gov/blog/lessons-learned-cms-artificial-intelligence-health-outcomes-challenge>
- ⁴¹⁶ Berenson RA, Shartz A, Pham HH. Beyond demonstrations: implementing a primary care hybrid payment model in Medicare. *Health Affairs Scholar*. 2023;1(2):qxad024. doi:10.1093/haschl/qxad024
- ⁴¹⁷ McGlynn EA. Improving the Quality of U.S. Health Care — What Will It Take? *New England Journal of Medicine*. 2020;383(9):801-803. doi:10.1056/NEJMp2022644
- ⁴¹⁸ McWilliams JM. Pay for Performance: When Slogans Overtake Science in Health Policy. *JAMA*. 2022;328(21):2114-2116. doi:10.1001/jama.2022.20945
- ⁴¹⁹ McWilliams JM. Professionalism Revealed: Rethinking Quality Improvement in the Wake of a Pandemic. *NEJM Catalyst*. 2020;1(5). doi:10.1056/CAT.20.0226
- ⁴²⁰ Markovitz AA, Ayanian JZ, Sukul D, Ryan AM. The Medicare Advantage Quality Bonus Program Has Not Improved Plan Quality. *Health Affairs*. 2021;40(12):1918-1925. doi:10.1377/hlthaff.2021.00606
- ⁴²¹ Thomas B, Valuck MD, Sarah Sampsel MPH, David M. Sloan P, Jennifer Van Meter P. Improving Quality Measure Maintenance: Navigating the Complexities of Evolving Evidence. 2019;25. <https://www.ajmc.com/view/improving-quality-measure-maintenance-navigating-the-complexities-of-evolving-evidence>
- ⁴²² Center for Healthcare Quality & Payment Reform. 5 Fatal Flaws in Total Cost of Care & Population-Based Payment Models. Center for Healthcare Quality & Payment Reform: n.d. Accessed November 28, 2023. https://chqpr.org/downloads/Flaws_in_TCOC_and_Population-Based_Payments.pdf
- ⁴²³ Berenson RA, Upadhyay D, Delbanco SF, Murray R. Payment Methods: How They Work | Urban Institute. Urban Institute; 2016. Accessed December 1, 2023. <https://www.urban.org/research/publication/payment-methods-how-they-work>
- ⁴²⁴ Ahluwalia SC, Damberg CL, Haas A, Shekelle PG. How are medical groups identified as high-performing? The effect of different approaches to classification of performance. *BMC Health Serv Res*. 2019;19(1):500. doi:10.1186/s12913-019-4293-9
- ⁴²⁵ Chernew ME, McWilliams JM, Shah SA. The Case for Administrative Benchmarks (and Some Challenges). *NEJM Catalyst*. 2023;4(10). doi:10.1056/CAT.23.0194
- ⁴²⁶ Jacobs DB, Schreiber M, Seshamani M, Tsai D, Fowler E, Fleisher LA. Aligning Quality Measures across CMS — The Universal Foundation. *New England Journal of Medicine*. 2023;388(9):776-779. doi:10.1056/NEJMp2215539
- ⁴²⁷ Assistant Secretary for Planning and Evaluation. Second Report to Congress on Social Risk and Medicare's Value-Based Purchasing Programs. ASPE; 2020. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>
- ⁴²⁸ National Quality Forum. Social Risk Trial Final Report. NQF; 2017. https://www.qualityforum.org/Publications/2017/07/Social_Risk_Trial_Final_Report.aspx
- ⁴²⁹ Skopec L, Bowen G, Zuckerman S. Accounting for Social Risk in Value-Based Payment and Quality Measurement. Urban Institute; 2022. <https://www.urban.org/research/publication/accounting-social-risk-value-based-payment-and-quality-measurement>
- ⁴³⁰ RAND Health Care. *Developing Health Equity Measures*. ASPE; 2021. <https://aspe.hhs.gov/reports/developing-health-equity-measures>

-
- ⁴³¹ RTI. ACO Realizing Equity, Access, and Community Health (REACH) Model - PY2024 Quality Measurement Methodology. Centers for Medicare & Medicaid Services; n.d. <https://www.cms.gov/files/document/aco-reach-quality-msr-meth-py24.pdf>
- ⁴³² Mentias A, Peterson ED, Keshvani N, et al. Achieving Equity in Hospital Performance Assessments Using Composite Race-Specific Measures of Risk-Standardized Readmission and Mortality Rates for Heart Failure. *Circulation*. 2023;147(15):1121-1133. doi:10.1161/CIRCULATIONAHA.122.061995
- ⁴³³ Heidenreich P, Sandhu A. Pursuing Equity in Performance Measurement. *Circulation*. 2023;147(15):1134-1136. doi:10.1161/CIRCULATIONAHA.123.064123
- ⁴³⁴ Burke LG, Orav EJ, Zheng J, Jha AK. Healthy Days at home: A novel population-based outcome measure. *Healthcare*. 2020;8(1):100378. doi:10.1016/j.hjdsi.2019.100378
- ⁴³⁵ Lilford RJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. *BMJ*. 2007;335(7621):648-650. doi:10.1136/bmj.39317.641296.AD
- ⁴³⁶ MedPAC. June 2014 Report to the Congress: Medicare and the Health Care Delivery System – MedPAC. MedPAC; 2014. <https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-jun14-entirereport-pdf/>
- ⁴³⁷ Duseja R, Andress J, Sandhu AT, et al. Development of Episode-Based Cost Measures for the US Medicare Merit-based Incentive Payment System. *JAMA Health Forum*. 2021;2(5):e210451. doi:10.1001/jamahealthforum.2021.0451
- ⁴³⁸ Duseja R, Andress J, Sandhu AT, et al. Development of Episode-Based Cost Measures for the US Medicare Merit-based Incentive Payment System. *JAMA Health Forum*. 2021;2(5):e210451. doi:10.1001/jamahealthforum.2021.0451
- ⁴³⁹ Assistant Secretary for Planning and Evaluation. *Second Report to Congress on Social Risk and Medicare’s Value-Based Purchasing Programs*. June 2020. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>
- ⁴⁴⁰ MedPAC. *March 2022 Report to the Congress: Medicare Payment Policy – MedPAC; 2022*. <https://www.medpac.gov/document/march-2022-report-to-the-congress-medicare-payment-policy/>
- ⁴⁴¹ MedPAC. *June 2018 Report to the Congress: Medicare and the Health Care Delivery System*. 2018. <https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-jun18-medpacreporttocongress-rev-nov2019-note-sec-pdf/>
- ⁴⁴² Assistant Secretary for Planning and Evaluation. *Second Report to Congress on Social Risk and Medicare’s Value-Based Purchasing Programs*. ASPE; 2020. June 2020. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicares-value-based-purchasing-programs>
- ⁴⁴³ Machta R, Peterson G, Rotter J, et al. *Evaluation of the Maryland Total Cost of Care Model: Implementation Report*. Mathematica; 2021. <https://www.mathematica.org/publications/evaluation-of-the-maryland-total-cost-of-care-model-implementation-report>
- ⁴⁴⁴ Medicare and Medicaid Programs; Policy and Regulatory Revisions in Response to the COVID-19 Public Health Emergency. *Federal Register*. Published April 6, 2020. <https://www.federalregister.gov/documents/2020/04/06/2020-06990/medicare-and-medicaid-programs-policy-and-regulatory-revisions-in-response-to-the-covid-19-public>
- ⁴⁴⁵ Centers for Medicare & Medicaid Services. New to Measures. October 2023. <https://mmshub.cms.gov/about-quality/new-to-measures/what-is-a-measure>
- ⁴⁴⁶ Agency for Healthcare Research and Quality. Types of Healthcare Quality Measures. July 2015 <https://www.ahrq.gov/talkingquality/measures/types.html>
- ⁴⁴⁷ National Quality Forum. Improving Healthcare Quality. N.D. https://www.qualityforum.org/Setting_Priorities/Improving_Healthcare_Quality.aspx
- ⁴⁴⁸ Partnership for Quality Management. Endorsement & Maintenance. N.D. <https://p4qm.org/EM>
- ⁴⁴⁹ American Medical Association. Performance Improvement Continuing Medical Education. N.D. <https://www.ama-assn.org/education/ama-pra-credit-system/performance-improvement-continuing-medical-education-pi-cme>

-
- ⁴⁵⁰ Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q.* 2005;83(4):691-729. doi: 10.1111/j.1468-0009.2005.00397.x. PMID: 16279964; PMCID: PMC2690293.
- ⁴⁵¹ Agency for Healthcare Research and Quality. Types of Health Care Quality Measures. July 2015. <https://www.ahrq.gov/talkingquality/measures/types.html>
- ⁴⁵² Centers for Medicare & Medicaid Services. Quality Measurement at CMS. December 2022. <https://mmshub.cms.gov/about-quality/quality-at-CMS/theory#:~:text=At%20a%20high%20level%2C%20CMS,the%20physician%20who%20developed%20it>.
- ⁴⁵³ National Academy of Medicine. Patient-Centered, Integrated Health Care Quality Measures Could Improve Health Literacy, Language Access, and Cultural Competence. February 2019. <https://nam.edu/patient-centered-integrated-health-care-quality-measures-could-improve-health-literacy-language-access-and-cultural-competence/>
- ⁴⁵⁴ Centers for Medicare & Medicaid Services. New to Measures. October 2023. <https://mmshub.cms.gov/about-quality/new-to-measures/what-is-a-measure>
- ⁴⁵⁵ Agency for Healthcare Research and Quality. Understanding Quality Measurement. <https://www.ahrq.gov/patient-safety/quality-resources/tools/chtoolbox/understand/index.html>
- ⁴⁵⁶ Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q.* 2005;83(4):691-729. doi: 10.1111/j.1468-0009.2005.00397.x. PMID: 16279964; PMCID: PMC2690293.
- ⁴⁵⁷ Centers for Medicare & Medicaid Services. Measures Management System. February 2018. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Structural-Measures.pdf>
- ⁴⁵⁸ Agency for Healthcare Research and Quality. Types of Health Care Quality Measures. July 2015. <https://www.ahrq.gov/talkingquality/measures/types.html>;
- ⁴⁵⁹ Centers for Medicare & Medicaid Services. Measures Management System. February 2018. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Structural-Measures.pdf>
- ⁴⁶⁰ Le Corroller AG, Bonastre J. Patient-reported measures: how useful in health economics? *Eur J Health Econ.* 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z
- ⁴⁶¹ American Hospital Association. Total cost of care: Key considerations. <https://trustees.aha.org/articles/1339-total-cost-of-care-key-considerations>
- ⁴⁶² National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Health Care Utilization and Adults with Disabilities. Health-Care Utilization as a Proxy in Disability Determination. Washington (DC): National Academies Press (US); 2018 Mar 1. 2, Factors That Affect Health-Care Utilization. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK500097/>
- ⁴⁶³ Centers for Medicare & Medicaid Services. Cost and Resource Use Measures. March 2023. <https://mmshub.cms.gov/sites/default/files/Cost-and-Resource-Use-Measures.pdf>
- ⁴⁶⁴ Agency for Healthcare Research and Quality. Public Reporting of Cost Measures in Health. March 2015. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/public-reporting-cost-measures_technical-brief.pdf
- ⁴⁶⁵ Le Corroller AG, Bonastre J. Patient-reported measures: how useful in health economics? *Eur J Health Econ.* 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z
- ⁴⁶⁶ Le Corroller AG, Bonastre J. Patient-reported measures: how useful in health economics? *Eur J Health Econ.* 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z
- ⁴⁶⁷ Le Corroller AG, Bonastre J. Patient-reported measures: how useful in health economics? *Eur J Health Econ.* 2023;24(1):1-4. doi:10.1007/s10198-022-01524-z
- ⁴⁶⁸ The Agency for Healthcare Research and Quality. What is Patient Experience? September 2023. <https://www.ahrq.gov/cahps/about-cahps/patient-experience/index.html>
- ⁴⁶⁹ Centers for Medicare & Medicaid Services. Measures Management System. February 2018. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Access-Measures.pdf>

-
- ⁴⁷⁰ Office of the Assistant Secretary for Planning and Evaluation (ASPE). Developing Health Equity Measures. May 2021. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//200651/developing-health-equity-measures.pdf
- ⁴⁷¹ Centers for Medicare & Medicaid Services. Comprehensive End-Stage Renal Disease (ESRD) Care (CEC) Model Performance Year 5 Annual Evaluation Report Findings at a Glance. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/cec-annrpt-py5-fg>
- ⁴⁷² Centers for Medicare & Medicaid Services. Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Fourth Annual Evaluation Report. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/cpc-plus-fourth-annual-eval-report>
- ⁴⁷³ Centers for Medicare & Medicaid Services. End-Stage Renal Disease (ESRD) Treatment Choices (ETC) Model at a Glance. <https://www.cms.gov/priorities/innovation/media/document/etc-my2-aag>
- ⁴⁷⁴ Centers for Medicare & Medicaid Services. Center for Medicare and Medicaid Innovation (CMMI). Global and Professional Direct Contracting (GDPC) Model. <https://www.cms.gov/priorities/innovation/innovation-models/gdpc-model>
- ⁴⁷⁵ Centers for Medicare & Medicaid Services. Center for Medicare and Medicaid Innovation (CMMI). Home Health Value-Based Purchasing Model. <https://www.cms.gov/priorities/innovation/innovation-models/home-health-value-based-purchasing-model>
- ⁴⁷⁶ Centers for Medicare & Medicaid Services. Evaluation of the Independence at Home Demonstration <https://www.cms.gov/priorities/innovation/data-and-reports/2023/iah-year7-eval-report>
- ⁴⁷⁷ Centers for Medicare & Medicaid Services. Evaluation of the Next Generation Accountable Care Organization (NGACO) Model Evaluation. <https://www.cms.gov/priorities/innovation/data-and-reports/2022/nextgenaco-fifthevalrpt>
- ⁴⁷⁸ Centers for Medicare & Medicaid Services. Evaluation of the Primary Care First (PCF) Model Findings at a Glance. <https://innovation.cms.gov/data-and-reports/2022/pcf-first-eval-aag-rpt>
- ⁴⁷⁹ Revere L, Langland-Orban B, Large J, Yang Y. Evaluating the robustness of the CMS Hospital Value-Based Purchasing measurement system. *Health Serv Res.* 2021. doi: 10.1111/1475-6773.13608
- ⁴⁸⁰ Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. McKinsey & Company. July 2018. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>
- ⁴⁸¹ Eggleton K, Liaw W, Bazemore A. Impact of Gaps in Merit-Based Incentive Payment System Measures on Marginalized Populations. The University of Auckland. May 2017. <https://www.annfammed.org/content/annalsfm/15/3/255.full.pdf>
- ⁴⁸² Young-Rock H, Ngyuen O, Yadav S, Etzold E, Song J, Duncan R, Turner K. Early Performance of Hospital Value-based Purchasing Program in Medicare. *Medical Care.* 2020. <https://doi.org/10.1097/MLR.0000000000001354>
- ⁴⁸³ Beauvais B, Whitaker Z, Kim F, Anderson B. Is the Hospital Value-Based Purchasing Program Associated with Reduced Hospital Readmissions? *Dovepress.* May 2022. <https://doi.org/10.2147/JMDH.S358733>
- ⁴⁸⁴ Banjee S, McCormick D, Paasche-Orlow M, Lin M, Hanchate A. Association between degree of exposure to the Hospital Value Based Purchasing Program and 30-day mortality. *BMC Health Services Research.* December 2019. <https://link.springer.com/article/10.1186/s12913-019-4562-7>
- ⁴⁸⁵ Izon G, Pardini C. association Between Medicare's Mandatory Hospital Value-Based Purchasing Program and Cost Inefficiency. *Applied Health Economics and Health Policy.* October 2017. <https://doi.org/10.1007/s40258-017-0357-3>
- ⁴⁸⁶ Jamieson D, Machado-Pereira M, Carlton S, Repasky C. Assessing the Medicare Advantage Star Ratings. McKinsey & Company. July 2018. <https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-medicare-advantage-stars-ratings>
- ⁴⁸⁷ Skopec L, Berenson R. The Medicare Advantage Quality Bonus Program. The Urban Institute. June 2023. <https://www.urban.org/sites/default/files/2023-06/The%20Medicare%20Advantage%20Quality%20Bonus%20Program.pdf>
- ⁴⁸⁸ Apathy N, Everson J. High Rates of Partial Participation in the First Year of the Merit-Based Incentive Payment System. *Health Affairs.* September 2020 doi: 10.1377/hlthaff.2019.01648
- ⁴⁸⁹ Nguyen O, Turner K, Parekh A, Tabriz A, Hanna K, Merlo L. Merit-based incentive payment system participation and after-hours documentation among US office-based physicians. *Journal of Evaluation in Clinical Practice.* November 2022. <https://doi.org/10.1111/jep.13796>