Physician-Focused Payment Model Technical Advisory Committee
Public Meeting Minutes

March 25, 2024
9:30 a.m. – 4:49 p.m. EDT
Hubert H. Humphrey Building
200 Independence Avenue, SW
Washington, DC 20201

Attendance

Physician-Focused Payment Model Technical Advisory Committee (PTAC) Members

Lauran Hardin, MSN, FAAN, PTAC Co-Chair (Chief Integration Officer, HC^2 Strategies)
Angelo Sinopoli, MD, PTAC Co-Chair (Executive Vice President, Value-Based Care, Cone Health)
Lindsay K. Botsford, MD, MBA (Market Medical Director, One Medical)
Lawrence R. Kosinski, MD, MBA (Independent Consultant)
Walter Lin, MD, MBA (Chief Executive Officer, Generation Clinical Partners)
Terry L. Mills Jr., MD, MMM (Independent Consultant)
James Walton, DO, MBA (President, JWalton, LLC)
Jennifer L. Wiler, MD, MBA (Chief Quality Officer, UCHealth Denver Metro, and Professor of Emergency Medicine, University of Colorado School of Medicine)

PTAC Members in Partial Attendance

Joshua M. Liao, MD, MSc (Professor and Chief, Division of General Internal Medicine, Department of Medicine, The University of Texas Southwestern Medical Center)*

PTAC Members Not in Attendance

Jay S. Feldstein, DO (President and Chief Executive Officer, Philadelphia College of Osteopathic Medicine)
Soujanya R. Pulluru, MD (Independent Consultant)

Department of Health and Human Services (HHS) Guest Speaker

Elizabeth (Liz) Fowler, JD, PhD (Deputy Administrator, Centers for Medicare & Medicaid Services [CMS] and Director, Center for Medicare and Medicaid Innovation [CMMI])
Susannah Bernheim, MD, MHS (Chief Quality Officer and Acting Chief Medical Officer, CMS CMMI)

Office of the Assistant Secretary for Planning and Evaluation (ASPE) Staff

Lisa Shats, PTAC Designated Federal Officer *
Audrey McDowell
Steven Sheingold, PhD

Via Zoom

List of Speakers and Handouts

1. **PCDT Presentation: Developing and Implementing Performance Measures for PB-TCOC Models**
   Jennifer L. Wiler, MD, MBA, Preliminary Comments Development Team (PCDT) Lead

   Handouts
   - Public Meeting Agenda
   - PCDT Presentation Slides
   - Environmental Scan on Developing and Implementing Performance Measures for Population-Based Total Cost of Care (PB-TCOC) Models
   - Overview of Current Performance Measures Included in Selected Medicare Payment Programs

2. **Panel Discussion: Developing Objectives for Performance Measurement for PB-TCOC Models**
   Cheryl L. Damberg, PhD, MPH, Director, RAND Center of Excellence on Health System Performance*
   Helen Burstin, MD, MPH, Chief Executive Officer, Council of Medical Specialty Societies (CMSS)
   John B. Bulger, DO, MBA, Chief Medical Officer Insurance Operations and Strategic Partnerships, Geisinger Health Plan*
   Eric C. Schneider, MD, MSc, Executive Vice President, Quality Measurement and Research, National Committee for Quality Assurance (NCQA) (*The "Medical Neighborhood" Advanced Alternative Payment Model [AAPM] [Revised Version]* proposal)

   Handouts
   - Panel Discussion Day 1 Panelists' Biographies
   - Panel Discussion Day 1 Introduction Slides
   - Panel Discussion Day 1 Discussion Guide

3. **Listening Session 1: What Do We Want to Measure in PB-TCOC Models, and How?**
   Thomas Sequist, MD, MPH, Chief Medical Officer, Mass General Brigham*
   David Meltzer, PhD, MD, Chief of the Section of Hospital Medicine, Director, Center for Health and the Social Sciences, and Chair, Committee on Clinical and Translational Science, University of Chicago; and Fanny L. Pritzker Professor of Medicine, Department of Medicine, University of Chicago Harris School of Public Policy and the Department of Economics (*Comprehensive Care Physician Payment Model (CCP-PM)* proposal)*
   Franklin Gaylis, MD, FACS, Chief Scientific Officer, Genesis Healthcare Partners; Executive Medical Director, Unio Health Partners; and voluntary Professor, Urology, University of California San Diego*

   Handouts
   - Listening Session 1 Day 1 Presenters' Biographies
   - Listening Session 1 Day 1 Presentation Slides
   - Listening Session 1 Day 1 Facilitation Questions

4. **Listening Session 2: Issues Related to Selecting and Designing Measures for PB-TCOC Models**
   Krishna G. Ramachandran, MBA, MS, Senior Vice President, Health Transformation and Provider Adoption, Blue Shield of California*
   Dana Gelb Safran, ScD, President and Chief Executive Officer, National Quality Forum*
   Vivek Garg, MD, MBA, Chief Medical Officer, Primary Care, Humana*

Sai Ma, PhD, MPA, Director, Enterprise Clinical Quality, Elevance Health*

Handouts

- Listening Session 2 Day 1 Presenters' Biographies
- Listening Session 2 Day 1 Presentation Slides
- Listening Session 2 Day 1 Facilitation Questions

***Via Zoom**

[NOTE: A transcript of all statements made by PTAC members and public commenters at this meeting is available on the ASPE PTAC website located at: https://aspe.hhs.gov/ptac-physician-focused-payment-model-technical-advisory-committee].

The ASPE PTAC website also includes copies of the presentation slides and other handouts and a video recording of the March 25 PTAC public meeting.

Welcome and Co-Chair Update

Angelo Sinopoli, PTAC Co-Chair, welcomed the Committee and members of the public to the March 25–26, 2024 public meeting. He explained that the Committee has been exploring themes that have emerged from proposals submitted by the public to PTAC. PTAC is also releasing a public report to the Secretary of Health and Human Services (HHS) with its findings on each theme. Co-Chair Sinopoli noted that PTAC recently released the June 2023 Report to the Secretary on Improving Management of Care Transitions in Population-Based Models. He shared that PTAC will release the September 2023 Report to the Secretary on Encouraging Rural Participation in Population-Based Total Cost of Care (PB-TCOC) Models soon. Co-Chair Sinopoli noted that previous public meeting discussions and proposals have shown that providers face challenges with implementing performance measures, particularly in TCOC models. He explained that this topic is also of interest to the Center for Medicare and Medicaid Innovation (CMMI or the Innovation Center) and introduced Dr. Elizabeth (Liz) Fowler, the Director of CMMI.

Dr. Fowler started by noting the importance of PTAC and the impact the Committee has had on CMMI models. She noted that prior PTAC proposals have influenced the design of several CMMI models. Primary Care First was influenced by the PTAC proposals submitted by the American Academy of Family Physicians (AAFP) and University of Chicago; the Oncology Care Model and its successor, the Enhancing Oncology Model, were both influenced by PTAC submissions from Hackensack Meridian Health, the Community Oncology Alliance (COA), and the American Society of Clinical Oncology (ASCO); and the Kidney Care Choices Model was influenced by the PTAC proposal submitted by the Renal Physicians Association. Dr. Fowler mentioned that PTAC's theme-based meetings have been helpful for CMMI as the challenges associated with moving to value-based care have become more complex over time. Dr. Fowler noted that the CMMI models continue evolving to address the challenges.

Dr. Fowler provided a preview of an upcoming panel discussion with the Centers for Medicare & Medicaid Services' (CMS') quality leadership. During the panel discussion, leaders across CMS will discuss the current state of performance measurement and where CMS and CMMI hope performance measurement will go moving forward.

Dr. Fowler introduced Dr. Susannah Bernheim, acting Chief Medical Officer and Chief Quality Officer at CMMI. Dr. Bernheim shared that the agency is working to improve the alignment of measures and reduce the burden associated with implementing measures. She indicated that CMMI is focusing its new strategy on person-centered outcomes and patient experience. Dr. Bernheim stated that CMMI is

broadening its definition of success beyond reducing costs and is seeking paths to improve quality across models. Dr. Bernheim noted that the CMS Quality Conference in April 2024 will include discussion of CMMI's new quality pathway.

Co-Chair Sinopoli described the agenda and session topics for the March 2024 public meeting. He noted that the discussion materials and public comments from the March public meeting will be incorporated into a Report to the Secretary delivered to HHS on developing and implementing performance measures for TCOC models. Co-Chair Sinopoli then invited Committee members to introduce themselves and describe their experience with performance measurement.

Following Committee member introductions, Co-Chair Sinopoli pointed out the four PTAC members who served on the Preliminary Comments Development Team (PCDT) for the meeting: Jennifer Wiler (Lead), Lawrence Kosinski, Soujanya Pulluru, and James Walton. Co-Chair Sinopoli introduced Dr. Wiler, who presented the PCDT's findings included in the background materials.

PCDT Presentation: Developing and Implementing Performance Measures for PB-TCOC Models

Dr. Wiler delivered the PCDT presentation. For additional details, please see the presentation slides, transcript, and meeting recording (18:11-48:19).

- Dr. Wiler summarized the objectives for the theme-based meeting, including discussing the challenges associated with quality measurement development and implementation, as well as understanding how quality measurement development and implementation affect PB-TCOC models and care delivery.
- Dr. Wiler noted that PTAC has received 35 proposals for physician-focused payment models (PFPMs). At least 60% of the proposed models met Criterion 2 (Quality and Cost) and Criterion 4 (Value over Volume) of the Secretary's 10 regulatory criteria.
- Dr. Wiler provided the working definition PTAC is using for performance measures: Performance measures assess and monitor all aspects of participants' performance in models including quality (e.g., process and structure), outcomes, cost, and utilization.
- Dr. Wiler introduced a diagram depicting the relationship between guiding principles and the types of performance measures that should be included in PB-TCOC models. On the outside of the wheel are the guiding principles that serve as the driving forces to move the four gears inside the wheel. The guiding principles include patient engagement; care transitions and coordination; equity; efficiency; and proactive patient-centered, high-touch care. The four gears inside the wheel represent the four types of performance measures used to evaluate the guiding principles. The performance measures include quality, outcomes, cost, and utilization. Patient experience and care delivery team effectiveness are at the core of the wheel.
- Dr. Wiler described findings from an analysis conducted on 24 CMS payment programs and CMMI models using information in the CMS Measure Inventory Tool (CMIT). The analysis generated the following results:
  - There were 618 current performance measures for 24 programs/models in the CMIT.
  - 61% of the measures were used by only one Medicare program or model. These findings suggest that it may be challenging to scale the measures to different groups that participate in multiple Medicare programs.
  - 59% of the measures were not endorsed by a CMS consensus-based entity.
  - Half of the measures were contained in the Merit-based Incentive Payment System (MIPS) Program. MIPS participants select at least six measures from the full set of 309 measures.
  - 52% of the measures were process measures. 39% of the measures were outcome measures.

    - Nearly all measures were linked to payment. However, the types of links to payment varied from pay-for-reporting to pay-for-performance (P4P).
    - There was no clear association between the number of performance measures and the percentage of financial risk across the 24 Medicare payment programs and models.
- Dr. Wiler described several challenges related to developing and implementing performance measures, including the following:
    - There is little evidence that public reporting of measures is linked to improved overall quality of care. Provider scores on performance measures are not necessarily associated with patient outcomes.
    - Measure development requires multiple steps that can take, on average, five to six years to complete. The endorsement process can take up to six months to complete. Additional time and resources are required to adapt measures for use in value-based care programs.
    - Quality reporting places an administrative burden on physicians and staff. Further, the inconsistency in measures used across programs can place an additional administrative burden on physicians and staff.
    - 54% of current performance measures in the CMIT are from electronic sources. There may be additional direct and indirect costs of participating in a registry and providing oversight to collect and submit data. 40% of measures in the CMIT use multiple data sources that must be aggregated, which can add administrative burden for providers.
    - It can take five to six months after a health care event to finalize Medicare administrative claims data. Utilization and cost data from the Healthcare Cost and Utilization Project (HCUP) are available approximately 18 months after the end of the year.
- Dr. Wiler described several challenges related to linking performance measures to payment, including the following:
    - Although evidence is mixed, some types of financial incentives may impact performance more than other types of financial incentives. Pay-for-performance (P4P) incentives, larger incentives (although evidence is mixed), more timely incentives, and financial penalties may impact performance.
    - Although P4P programs might impact performance, they may also widen disparities because they may disproportionately penalize providers who serve lower socioeconomic status (SES) or minority patients.
    - P4P may unintentionally create perverse actions, including decreasing the focus on individual patient concerns; diverting attention from important areas of clinical care that are not subject to P4P incentives; and avoiding treating patients who are disadvantaged, underserved, or high cost.
    - Risk adjustment should account for underlying differences in patient populations. 86% of 14 selected CMMI models use a risk-adjustment method, and approximately 30% of those models use the CMS hierarchical condition categories (HCC) risk scores.
    - National benchmarks do not account for geographic differences in patient populations and may unfairly penalize certain types of providers (e.g., rural providers). Different performance thresholds may have different impacts on performance. Absolute thresholds are consistent and transparent for all providers but may not promote improvement for providers who already meet the thresholds. Relative thresholds promote continuous improvement of organizations but may reduce collaboration between high- versus low-performance groups and create persistent gaps.
- Dr. Wiler concluded the presentation by providing a preview of the session topics planned for the theme-based meeting.

Co-Chair Sinopoli invited Committee members to ask questions and comment on the PCDT presentation. Committee members discussed the following topics. For more details on the discussion, see the transcript and meeting recording (48:19-51:45).

- Solutions must be identified to shorten the amount of time required to implement measures.
- It is critical to ensure that process measures relate directly to outcomes.
- Measures must be relevant for both primary care providers (PCPs) and specialty care providers.
- Use of patient-reported outcomes is critical.
- It is critical to reevaluate developed measures over time to ensure that the measures produce the outcomes we intend to produce.
- The substantial administrative burden placed on providers should be recognized when reporting quality measures.
- Using the Area Deprivation Index (ADI) and other measures to index the environment in which physicians are working to improve patient-reported outcomes should be considered.

Panel Discussion: Developing Objectives for Performance Measurement for PB-TCOC Models

SMEs

- Cheryl L. Damberg, PhD, MPH, Director, RAND Center of Excellence on Health System Performance
- Helen Burstin, MD, MPH, Chief Executive Officer, Council of Medical Specialty Societies (CMSS)
- John B. Bulger, DO, MBA, Chief Medical Officer Insurance Operations and Strategic Partnerships, Geisinger Health Plan

Previous Submitter

- Eric C. Schneider, MD, MSc, Executive Vice President, Quality Measurement and Research, National Committee for Quality Assurance (NCQA) (*The "Medical Neighborhood" Advanced Alternative Payment Model [AAPM] [Revised Version]* proposal)

Lauran Hardin, PTAC Co-Chair, moderated the panel discussion with four subject matter experts (SMEs) offering their perspectives on developing objectives for performance measurement for PB-TCOC models. For additional details, please see the transcript and meeting recording (00:10-1:37:35).

Panelists introduced themselves and provided background on their respective organizations. Full biographies and panelist introduction slides are available.

- Cheryl Damberg introduced herself as the Director of the RAND Center of Excellence on Health System Performance.
    - Dr. Damberg has been engaged in performance measurement throughout her career, during which she has worked on the Healthy People Objectives, led performance measurement and the shift to value-based payment in the private sector, developed and applied performance measures in practice, and conducted applied studies about performance measurement. She shared that the core objective of performance measurement is to encourage health systems to build measurement infrastructures that allow them to internally monitor and improve their performance. Dr. Damberg noted that relying on macro-level measures, such as total cost of care, may drive delivery system transformation with less burden than using more granular measures.
    - Performance measurement dashboards should use a broad set of measures, including health equity measures and patient-reported outcomes; these measures should be tied to payment and accountability to facilitate improved outcomes.
    - Dr. Damberg explained that, beyond use of performance measurement, other strategies and tools need to be deployed to drive change. Payment reform, such as value-based

insurance design, is needed to incentivize care redesign and innovation. In the future, processes for performance measure construction and reporting should leverage electronic health records and artificial intelligence, and performance measure improvement initiatives should target organizations rather than individual physicians.

- For additional details on Dr. Damberg's background and organization, see the panelist introduction slides (slides 2-5).
- Helen Burstin introduced herself as the Chief Executive Officer of the Council of Medical Specialty Societies, an organization that advances the expertise and collective voice of specialty societies. Dr. Burstin noted that specialty societies develop and test quality measures, and many have clinical registries. She has ample expertise in quality and equity measurement, including serving as the former Chief Scientific Officer at the National Quality Forum. She shared that little has changed in equity, measurement, assessment, and improvement over her decades of experience in quality and equity measurement. Dr. Burstin stated that specialists should be accountable for measures specifically attributable to their performance, as well as measures that reflect appropriateness, shared decision-making, and patient-reported measures to support collaboration across time, clinicians, and settings in PB-TCOC models. Measures should be developed for and by clinicians and derived from rich clinical data, such as clinical registries as well as digital quality registries. She also noted the importance of using measures that are meaningful to both physicians and patients. For additional details on Dr. Burstin's background and organization, see the panelist introduction slides (slides 6-8).
- John Bulger introduced himself as Chief Medical Officer Insurance Operations and Strategic Partnerships at Geisinger Health Plan, where he helps lead performance measurement and quality measures for the Keystone Accountable Care Organization (ACO) and other CMS demonstration projects. Dr. Bulger was responsible for all quality reporting across inpatient and outpatient settings as the former Chief Quality Officer of Geisinger Health, and he conducted all measurement reviews as the former Chair of the National Quality Forum Consensus Standards Approval Committee (CSAC). Dr. Bulger shared that quality measures should be simple, focus on outcomes, and recognize equity. He also noted that the goal of measurement should be to protect the public, highlighting that some payment approaches can incentivize providers to focus on a narrow set of things, or create a risk for gaming the system. For additional details on Dr. Bulger's background and organization, see the panelist introduction slides (slides 9-11).
- Eric Schneider introduced himself as the Executive Vice President of Quality Measurement and Research for the National Committee for Quality Assurance (NCQA), an organization that collects Healthcare Effectiveness Data and Information Set (HEDIS) data from health insurance plans representing over 203 million Americans (61% of the population). While PB-TCOC models have potential to reduce costs while improving the health of the population, there are concerns that this approach might lead to stinting on care and worsening health outcomes if providers focus primarily on financial risk. There are poor health outcomes related to equity in the United States, and patient cost burden erodes trust among patients which impacts access to timely care. For organizations to take on PB-TCOC payments, they should be empowered to take responsibility for all attributed members and be able to coordinate care effectively across primary care, specialty care, behavioral health, and community-based organizations. Dr. Schneider also shared four priorities for performance measurement, in order of importance: 1) equity, including reducing disparities and addressing unmet social needs; 2) access to care, including availability and timeliness; 3) experience and outcomes of care, including communication and person-centered outcomes; and 4) effectiveness of clinical services, including evidence-based care, cost-effectiveness, reliability, and safety. He emphasized the importance of moving beyond claims data and investing in a health data infrastructure to create clinically relevant measures. For

additional details on Dr. Schneider's background and organization, see the panelist introduction slides (slides 12-31).

Panelists discussed the main goals of performance measurement for TCOC organizations.

- The same measure can be used across different levels of a system. Larger organizations with larger patient populations can manage greater financial risk than individual physicians or small practices. Nuanced data are needed to understand the risks and health needs of populations. There are small practices doing great work; however, they operate without a network of support.
- The definition of a provider is unclear, with definitions ranging from individual health care professionals to health systems. PB-TCOC models must integrate the role of the clinical care team within the broader vision of performance measurement and reporting at the level of a health care system. Clinically relevant measures and measures that reflect actionable performance are important, although actionable measures for the clinical care team may differ from actionable measures at the system level.
- The primary goal for performance measurement is to produce information for driving change in a health system, produce better care for patients, and help patients achieve better outcomes. Providers struggle to meet the reporting needs of different stakeholders, including producing information that consumers can use to make informed health care decisions. Financial incentives garner providers' attention, but the main objective for performance measurement should be to help providers understand how they are performing and how to improve.
- Performance measurement serves as a "gating mechanism" that holds TCOC organizations accountable and prevents providers from gaming the system. Sharing performance measurement information with the public can help improve quality across the system. Not all measures are reported, such as electronic health record (EHR) measures and patient-reported measures, but these measures can still be used by organizations to improve their performance. Geisinger Health Plan uses an integrated database of clinical and claims data to improve its performance, although it is not held financially accountable for all of these measures.

Panelists discussed types of performance measures for measuring organizations' performance relative to the desired characteristics of PB-TCOC models.

- TCOC is one of the most important performance measures to include in PB-TCOC models, along with access to care, timeliness of care, and denials of care. There has been a desire to reduce low-value care; however, measures of low-value care can create unnecessary friction without producing additional benefit. By listening to patients and understanding their care experiences, the health care system can become more patient-centered, leading to better outcomes for patients managing complex conditions. Building trust and connection between patients and providers can help with co-management of health care conditions and address underlying barriers to patient success, such as socioeconomic factors. Equity measures must also remain a focus.
- A balanced portfolio of measures is crucial. The ideal system measures TCOC with fail-safes to ensure that physicians do not game the system. Prior authorizations cause waste; ideally, providers would determine treatments for patients solely through patient-provider conversations. Providers can use claims-based utilization measures to track primary care, emergency room, and inpatient visits. Under a TCOC model, providers are incentivized to increase primary care visits and decrease utilization in more expensive settings (for example, emergency room, inpatient visits), which tends to reduce TCOC. Currently, different provider organizations bill procedures differently, making it difficult to compare performance across providers.

- Existing measures would benefit from better data. Clinical and public health systems should focus on measuring important health threats — such as diabetes, blood pressure, depression, and maternal outcomes — to reduce morbidity and mortality over time. The current health data exchange infrastructure cannot support the ideal measurement system. It is critical to understand the relative spending and mix of services provided across settings, such as excessive emergency room use. It can take multiple years to see shifts in performance, but most studies on the return on investment of payment system models do not allow enough time for interventions to demonstrate their effects.
- TCOC measures can lead to unintended consequences, such as stinting of patient care and care denials. Prior authorizations continue to burden physicians. Since TCOC is often measured at the system level, clinicians may not feel a part of that actionable process. Measures of shared decision-making and appropriateness may help physicians understand their role in TCOC. Similarly, it is critical for quality measures to help providers across different specialties and settings understand their role in the broader system of care.

Panelists discussed the differences between ideal performance measures for PB-TCOC models and current measures used in Medicare value-based payment programs and other Alternative Payment Models.

- The slow implementation of TCOC models hinders both providers and patients. TCOC models aim to create long-term change that short-term measures (for example, month-to-month or year-to-year measures) cannot capture. Leading and lagging indicators may be necessary. It is important to use measures that capture change over time because meaningful change in overall population health can take several years to realize. Shorter-term measures are still valuable to ensure that providers are moving in the right direction.
- Risk adjustment is a critical issue in PB-TCOC models. Netherlands' system reallocates funds among insurers based on population risk, which reduces the risk of providers gaming the system by pushing less healthy people out of the model and bringing in healthier people. Alternatively, a cross-sectional approach would follow patients through their care journey to see if risk selection occurs by tracking when patients disenroll and move to different care settings. Under a risk-based payment model, it is easier to cherry-pick to a lower-cost model.
- Team-based measures are important. Measures should reflect collaboration, communication, access, innovations (for example, telehealth), and meaningful outcomes. Clinical outcomes are often included in clinical registries, but registries are difficult to use. There is a lack of trust related to the appropriate use of data. Measures and benchmark data can help clinicians improve; for example, specialty societies can contact clinicians with low performance measure scores to ask about their plan for improvement. Clinicians can be motivated both intrinsically (i.e., by showing them how their performance compares with their peers) and extrinsically (i.e., financially). Many quality measures have been measured consistently over the years without much improvement. In the future, specialty societies may be able to help illuminate other important aspects of care for measurement purposes.
- Ideal performance measures may not differ from current performance measures. However, the current use of that information is not ideal. When the unit of accountability is the individual hospital or provider (for example, in MIPS), organizations do not always use this information to drive system improvement. PB-TCOC models have a less siloed approach. They focus on how organizations can respond to transform care delivery. Other programs allow physicians to self-select their measures, which provides opportunities to game the system. Approximately half of physicians are employed by larger health or hospital systems with resources and supports to help them improve their performance. Physicians working in small practices, especially in rural

environments, may not be able to improve even if they are aware of their performance. The ability of providers to act on performance measurement and make change differs substantially across settings.

Panelists discussed how the public and private payers can encourage health systems to build their own measurement infrastructure.

- The government does not need to be overly prescriptive. Commercial and government payers have provided performance measures over the past two decades, which encouraged health systems to invest in building their own measurement infrastructure. Smaller practices have struggled to build such an infrastructure, while larger systems with more resources have built robust dashboards spanning primary care and specialty care settings. However, specialty care measures are the weakest aspect of current measurement infrastructures and performance dashboards. If a better set of specialty measures were developed, systems would incorporate these specialty measures into their measurement infrastructure to receive the financial incentives. With the right directional signals, the government does not need to be overly prescriptive.
- Health data exchange is central to performance measurement. Advanced health technologies could support the creation of health data standards and data exchange capabilities at low cost. The technology industry may lead these advancements, rather than government mandates. For example, Trusted Exchange Framework and Common Agreement (TEFCA) allows organizations to use cloud computing power for health data exchange without investing in their own data exchange systems. Performance measurement strategies must consider recent changes in technology when building toward the future.
- Measurement dashboards should go beyond patient care to include measures of data collection burden, staff burnout, and staff turnover to capture a fuller perspective of health system performance.

Panelists discussed the measurement-related initiatives that they would make mandatory.

- Multi-payer participation arrangements would be valuable. The Office of the National Coordinator for Health Information Technology (ONC) has some mandates related to data exchange. EHRs create the capability to participate in multi-payer environments. A trusted exchange framework could support performance measurement and reduce administrative burden.
- Mandates should increase relevancy and decrease burden, such as by eliminating measures that do not add value and providing easier, more consistent measure collection methods.
- Multi-payer participation aligns what providers are being asked to pay attention to across different areas: interoperability and health information exchange systems, the potential to help coordinate care across settings, enhancing care delivery, and reducing low-value care. In addition, better care coordination measures are needed because patients, especially older patients, receive care across many different settings.
- A mandatory glide path toward TCOC models could facilitate the transition away from the fee-for-service (FFS) system. In the past, glide paths to mandatory bundles led to a focus on quality assessment. Quality measurement differs under TCOC models and FFS systems; under FFS systems, providers take a transactional approach to quality measurement. Requiring TCOC participation would encourage providers to focus on measures that improve the health of the population.

Panelists discussed the ideal mixture of quality, outcome, patient experience, cost, and utilization measures.

- Ideal measures focus on outcomes and equity more than process measures. Care model innovations (for example, lifestyle coaching, community health workers, peer counselors) have the potential to improve care and care coordination, but these innovations are not well-integrated into the health care system. Furthermore, process measures can impose unnecessary burden for providers who adopt new care model innovations.
- Equity and care coordination measures are critical, as well as measures of clinician health and burden. Ideal measures provide relevant and actionable information for clinicians, thereby capitalizing on clinicians' intrinsic motivation to improve care for their patients.
- Accountability measures are less important for transformation in TCOC models; in TCOC models, providers are incentivized to make decisions that benefit patients while also reducing TCOC (for example, home-based care). A broad portfolio of measures, including equity and patient experience measures, is important; however, there is not a need to use mandates because TCOC environments naturally lead to innovation related to measurement.

Panelists discussed what measures should be included to ensure a balanced portfolio of measures.

- There are useful safety and reliability measures for hospital settings, but we need better safety and reliability measures for ambulatory settings (for example, home-based care). A balanced portfolio should also incorporate measures related to staff and the workforce.
- To reduce low-value care and increase high-value care, a balanced portfolio of measures includes measures related to patient harm, equity, coordination, and appropriateness.
- A balanced portfolio of measures should consider access issues (for example, denials and wait times), as well as which settings are generating the most revenue (for example, primary care or inpatient). Provider burnout, clinician supply, and adequacy of provider networks can be measured to assess access issues in TCOC models.
- Measurement of patient churn across providers could help with understanding of whether providers are gaming or cherry-picking, how patients move across settings, and if patients cannot access care.

Listening Session 1: What Do We Want to Measure in PB-TCOC Models, and How?

SMEs

- Thomas Sequist, MD, MPH, Chief Medical Officer, Mass General Brigham
- Franklin Gaylis, MD, FACS, Chief Scientific Officer, Genesis Healthcare Partners; Executive Medical Director, Unio Health Partners; and voluntary Professor, Urology, University of California San Diego

Previous Submitter

- David Meltzer, PhD, MD, Chief of the Section of Hospital Medicine, Director, Center for Health and the Social Sciences, and Chair, Committee on Clinical and Translational Science, University of Chicago; and Fanny L. Pritzker Professor of Medicine, Department of Medicine, University of Chicago Harris School of Public Policy and the Department of Economics (*Comprehensive Care Physician Payment Model [CCP-PM]* proposal)

Co-Chair Sinopoli moderated the listening session with three SMEs on what do we want to measure in PB-TCOC models, and how. Full biographies and presentations are available.

Thomas Sequist presented on strategies and considerations to improve quality measurement in PB-TCOC models.

- Goals of measuring quality are to improve patient outcomes, patient experience, and equity throughout health care delivery, while also cutting costs.

  - Over the past 20 years, there have been slow improvements in patient outcomes through delivering evidence-based care, especially for conditions that contribute significantly to morbidity and mortality, such as cardiovascular disease and the substance use disorder epidemic.
  - There have been limited improvements in the coordination of care for patients, patient access to a baseline level of care, and the delivery of care through an empathetic lens.
  - Inequities in patient health outcomes, recruitment and training of providers, and in provider workforce diversity have persisted and worsened.
- A challenge to achieving these outcomes, experience, equity, and total cost of care goals is the lack of collaboration and alignment of priorities among ACOs. For example, at a health system level, the focus on specific metrics required for ACOs participating in various programs makes it difficult to coordinate the work to measure and improve quality.
- A second challenge is the misalignment of payment with an outcomes-focused orientation. For example, long-term clinical goals, such as reducing cardiovascular mortality, require a 5- to 10-year timeline, and the finances for pay-for-performance programs do not support that long of a timeline.
  - Overall, these two challenges create confusion around the direction clinicians and clinical management teams should follow for incentive programs. While there are many quality measures, the larger challenge is alignment and coordination of these measures.
- Dr. Sequist introduced two ways of thinking about quality: through content areas (i.e., the Institute of Medicine [IOM] model) and through structure, process, and outcomes (i.e., the Donabedian Model of Quality).
  - He noted that outcome quality measures are focused on process and structure measures even though improving patient outcomes should be the ultimate goal. For example, more than 50% of quality measures in CMS programs fall under process and structure.
- Dr. Sequist provided the following suggestions on how to improve patient outcomes in PB-TCOC models:
  - Strongly evaluate measures used in programs and prioritize outcome measures over process and structure measures. Move process and structure measures along a pathway that lead to outcome measures.
  - Clarify the definition of a quality and outcome measure as distinct from a utilization or access measure.
  - Synchronize metrics across hospital and ambulatory programs. For example, a CMS hospital value-based purchasing program can directly influence ambulatory HEDIS metrics.
  - Increase the number of quality measures for ambulatory care, which is often delivered in the specialty care space as opposed to the primary care setting, which has a long history of quality measures.
- To improve the patient experience in PB-TCOC models, Dr. Sequist suggested measuring communication, coordination, and empathy as outcome measures independent of a specific clinical outcome and doing so through objective reports of care rather than subjective ratings of care (i.e., instead of asking for general hospital ratings, ask patients about specific care activities, such as if medications were explained at discharge).
- To improve equity in PB-TCOC models, Dr. Sequist offered the following suggestions: prioritize closing the equity gap; expand the measurement of social risk factors that predict clinical outcomes beyond race, ethnicity, and language; focus on outcome measures rather than

structural measures; and acknowledge the role of risk adjustment in perpetuating existing inequities and the need for more resources to tackle inequity.

For additional details on Dr. Sequist's presentation, see the presentation slides (pages 3-11), transcript, and meeting recording (0:42-14:35).

David Meltzer presented on the measurement of desired characteristics and outcomes for PB-TCOC models.

- Dr. Meltzer's research found that the introduction of hospitalists or inpatient general medicine, a form of specialized medicine, did not improve patient outcomes. In response to this finding, he developed and submitted the Comprehensive Care Physician Payment Model (CCP-PM), in which primary care physicians provide comprehensive care for patients at increased risk of hospitalization. This model was evaluated through several randomized trials:
    - A randomized clinical trial (RCT) funded by CMMI found a decrease in hospitalizations and increase in primary care provider ratings. However, this decrease was lower among dual-eligible enrollees due to adverse selection of high-risk patients into the treatment arm of the study. This finding highlighted how unmet social needs influence the effectiveness of the model.
    - An RCT funded by the Patient-Centered Outcomes Research Institute (PCORI) compared the effectiveness of three programs or models in reducing hospitalization: the CCP-PM model; Robert Wood Johnson Foundation's Comprehensive Care, Community and Culture Program (C4P); and the Partners-like Care Coordination Program. Interim results showed that the C4P was most effective at reducing hospitalizations for dual-eligible enrollees.
- Dr. Meltzer discussed goals of performance measurement, including measuring both outcome and care process (how care is provided). He emphasized that process measures have advantages over outcome measures as use of process measures may discourage the avoidance of high-risk/cost patients, increase the likelihood that care guidelines that improve outcomes are followed, and encourage studies that examine how to improve care. He noted that the effectiveness of PB-TCOC models to improve care and reduce costs should be compared to alternatives, such as FFS reform and competition.
- Dr. Meltzer discussed opportunities for improvement with measuring patient experience, outcomes, and costs. Opportunities include measuring a health system's success in retaining vulnerable populations, measuring outcomes in patients who transition out of care, and appropriately defining the denominator for the population being served by the health system. He discussed challenges specific to patient experience (minimal vs. aspirational coding), population health measurement (general health vs. disease-specific measures), and costs (include costs to non-medical stakeholders that provide social services). He noted the importance of also measuring how the work life of health care providers impacts their relationships within the health care system and how this affects continuity of care.

For additional details on Dr. Meltzer's presentation, see the presentation slides (pages 12-19), transcript, and meeting recording (14:44-28:01).

Franklin Gaylis presented on the implementation of a pay-for-performance quality improvement payment model at Genesis Healthcare Partners (GHP) medical group.

- GHP's quality improvement model focused on identifying meaningful organization-level and provider-level specialty-related performance measures to improve treatment of low-risk prostate cancer and to provide feedback on provider performance.

- Dr. Gaylis cited a recent article in the *Journal of the American Medical Association (JAMA)* on the role of implementation science in accelerating the time it takes to translate evidence into clinical practice. GHP's project takes an implementation science approach. He summarized the history of GHP's quality improvement interventions and highlighted metrics and methods that were effective in improving physician adherence to performance measures for and adoption of conservative management with low-risk prostate cancer. These interventions included transparent reporting (physician audit and feedback) and a pay-for-performance value-based model (payment incentive). He noted that the payment incentive determination depended on whether the entire group met all four quality measure benchmarks, and the incentive was paid to the group not to individual physicians.
- The projected cost savings of improving the adoption of conservative management for low-risk prostate cancer are estimated at $150 to $200 million over three years. Upfront investment to build an automated electronic data capture and analytics system cost roughly $222,090.
- Challenges to implementing performance measures within a health system or medical group include agreement on the relevance of measures, their definitions, and thresholds; commitment of leadership to ease the burden of implementation on physicians/providers; and the cost of investing in an electronic health data system.
- In urology, cross-cutting measures are the most commonly reported, rather than measures specific to quality of care, such as a measure of active surveillance of prostate cancer, similar to the one used in GHP's intervention.
- Dr. Gaylis concluded with challenges for implementation of a quality improvement program, including encouraging participation by medical groups and payers and securing funding for start-up expenses. Based on GHP's experience, he suggested a pay-for-reporting approach. He noted that measurement and reporting often leads to the Hawthorne effect whereby one's work product improves due to being monitored.

For additional details on Dr. Gaylis' presentation, see the presentation slides (pages 20-37), transcript, and meeting recording (28:06-43:08).

Following the presentations, Committee members asked questions of the presenters. For more details on this discussion, see the transcript and meeting recording (43:12-1:29:00).

Presenters discussed recommendations on what to measure in TCOC models to meet the quadruple aim – the measurement of outcomes, patient experience, work life of health care providers, and costs.

- Dr. Gaylis noted that measurement needs to be disease-specific. In his medical group's quality improvement model, the goal was to mitigate the harms of overtreatment of low-risk prostate cancer. The outcome measured was active surveillance. Patient experience was measured via the rater8 system which includes the net promoter score. Cost benefits were also measured. One additional consideration is the protection of patients with disease progression and ensuring intervention in a timely fashion to provide curative therapy.
- Everything should be measured, not just outcomes. The path from process to structure to outcomes is important, and a balanced approach is needed. Disease-specific measures are cleaner but narrow. Cross-cutting measures should be considered because they can be synchronized across care settings.
- There are essentially two approaches being discussed: 1) broad-based process or structure measures that theoretically will improve care and desired outcomes across the board; or 2) outcomes measures that are specific to a patient population under treatment and reflect what contributes most to morbidity and mortality in that patient population. Neither approach is

clearly the right answer, but it is important to avoid having PB-TCOC programs with different priorities within a health care delivery system.

Presenters discussed the use of a balanced portfolio approach and recommendations for the appropriate mixture and volume of outcomes, patient experience, and process measures to directly measure systems-level change.

- The appropriate approach depends on the characteristics of the health care delivery system and the goals of measurement. Measuring process allows for testing of how care delivery is correlated with outcomes. Our confidence in outcomes and process measures can guide which are prioritized and which are over- or under-weighted.
- There is no one correct approach or percentage of process, structure, and outcomes measures that should be in a portfolio. We often are stuck in the process and structure measures and need to ensure that each process or structure measure is tied to an outcome measure over time. The link between process measures and outcomes needs to be validated repeatedly over time.
- There is no "one-size-fits-all" approach. The appropriate measures vary by specialty and patient population and need to be customized. There are still no standardized measures uniformly applied to urology and, thus, we see reporting on cross-cutting measures.

Presenters discussed how performance measurement in value-based care models differs from what it needs to be in a TCOC model.

- The issue is not so much the difference in quality measures between value-based care and TCOC programs, but rather, the path forward for quality measurement in general. First, the design of TCOC programs is important to prevent gaming the system in a way that adversely affects underserved populations (i.e., through avoidance of high-risk patients). Guardrails are needed wherever measuring cost; this has been a historical problem in value-based programs. Second, it is important to have selection of one approach to quality management and clear messaging of this approach to providers on the ground so they can understand how they are being incentivized to improve care.
- The principles of quality management can be applied across all disciplines and subspecialities; however, the measures that are important may be very specific to a discipline or subspecialty.
- PB-TCOC models differ from traditional value-based purchasing programs in how the population is defined (i.e., whether the denominator is fluid or fixed) and how cost is measured (i.e., the market structure which affects gaming likelihood). For socially vulnerable populations, the availability of social services and its interaction with medical utilization are important. A more integrated view of quality measurement is needed. For example, management of prostate cancer in mid-life can affect urinary incontinence later in life, which in turn affects ability to live independently, and ultimately the cost to Medicare.

Presenters discussed patterns or archetypes seen in subgroups of high-risk populations and how this informs selection of measures that would advance health equity in TCOC models.

- Dr. Meltzer shared that in his program, young Black men with end-stage renal disease (ESRD) have high costs of care and poor outcomes; they present with a history of neglected care and tremendous unmet social needs. This subgroup comprises 6% of the program's total patient population and accounts for a large proportion of morbidity and cost. Alternative Payment Models (APMs) for ESRD were active in the period when Dr. Meltzer's study began, and unfortunately there were incentives to not select these patients into those programs. Dr. Meltzer noted that this condition is extremely difficult to study using RCTs with adequate scale.

- Dr. Gaylis commented that the Prostate Cancer Active Surveillance Project (PCASP), a national consortium of academic and community urologists, has had difficulty conducting RCTs examining implementation science approaches due to insufficient resources and the lack of preliminary data. Conducting retrospective analysis of EHR data is not feasible due to poor physician documentation. Dr. Gaylis' medical group invested in an IT system to collect and pull the data without compensation.
- Additional challenging subgroups of patients are those struggling with substance use disorders and housing/food insecurity. Many of these patients seek care in emergency departments, which increases costs. Additionally, patients whose primary language is not English struggle in a care delivery system that is designed for English speakers.
- Dr. Meltzer noted that many of their patients with ESRD are dual-eligible patients and have tremendous unmet social needs. The high-need, high-cost subpopulations being discussed overlap profoundly.
- Black men tend to present with more advanced and aggressive disease. The predominant driver of this inequity is access to care.

Presenters discussed the importance of a patient caregiver, such as a family member, and patient-reported outcomes.

- Patient-reported outcome measures (PROMs) can apply across conditions or be disease-specific. PROMs are useful for improving and expanding upon other outcomes measures to give insight into critical outcomes of care. For a patient who is hospitalized for a knee replacement, outcomes measures typically may include hospital infection rates and 30-day readmission rates, but not the true reason the patient underwent the surgery (such as inability to climb the stairs). PROMs have tremendous potential to fill this gap and produce better outcome measures. The logistics of collecting this data from patients are a challenge, and needs additional work.
- In the context of a university, students are engaged in data collection from patients. Often, technology is not accessible for vulnerable populations, and other methods of collecting data are needed. Patient-engaged research and the use of mixed methods are helpful ways to make sure necessary data are captured, especially in vulnerable populations.
- Validated questionnaires are needed in private practice settings, as well as resources to collect, manage, and analyze the data. Reimbursement of these activities does not yet exist in our health care system, so this work is challenging in small and private practice settings.
- PROMs are one of the most expensive and resource-intensive, but useful, types of measures collected by our health care system.

Presenters discussed how to integrate or structure metrics for specialists and specialty services into TCOC models in a broader network.

- In the urology discipline, there are many subspecialties that make bundling difficult. Different specialties would need very different measures; bundling should start organically at the specialty level.
- Bringing specialty services into TCOC models is essential. The question is whether to measure TCOC across an entire system or at the individual physician level. Care is often spread across different providers, and there is confusion to whom the bundle of care should be assigned. Having a whole system perspective is easier and facilitates measuring performance; for example, the cost of treating a hip fracture from start to finish, at the system level without becoming mired in provider-level detail. An additional consideration is how the start and end of a medical episode is defined as this will influence the inclusion of specialty care in the TCOC.

- Integrating medical specialty services with primary care in terms of how to measure quality is a better approach than trying to create measures for each specialty service. Quality measures should be broad and interconnected across systems, and infrastructure should be built recognizing this synergy.

Presenters discussed what publicly funded entities such as CMS can do to reduce confusion on the ground about the direction of incentive programs.

- The incoming generations are faced with challenges in funding for clinician investigators. Resources for academic medicine are insufficient to build and grow the next generation of investigators. CMS, the National Institutes of Health (NIH), and other funding agencies need larger budgets to fund the next generation of research and development in this area.
- Given the magnitude of this problem, CMS should prioritize creating a clear plan for addressing health equity with a goal of being given successful in improving the quality and cost of care. Given the evidence showing that unmet social needs increase the cost of health care, CMS can fund social risk factors in a way that would push commercial payers to move in the direction of funding social services. CMS can encourage EHR vendors to develop products and services that support vulnerable populations. Finally, CMS can be a leader in expanding language access throughout the health care delivery system.
- The government should engage more physicians from community and private practice in reforming the payment system and conducting research on performance measurement. The government needs to invest in the infrastructure as private practices cannot absorb these costs.

Listening Session 2: Issues Related to Selecting and Designing Measures for PB-TCOC Models

SMEs

- Krishna G. Ramachandran, MBA, MS, Senior Vice President, Health Transformation and Provider Adoption, Blue Shield of California
- Dana Gelb Safran, ScD, President and Chief Executive Officer, National Quality Forum
- Vivek Garg, MD, MBA, Chief Medical Officer, Primary Care, Humana
- Sai Ma, PhD, MPA, Director, Enterprise Clinical Quality, Elevance Health

Co-Chair Hardin moderated the listening session with four SMEs on issues related to selecting and designing measures for PB-TCOC models. Full biographies and presentations are available.

Krishna Ramachandran presented on the pay-for-value strategy used by Blue Shield of California, including the challenges its providers face in measuring performance and the strategies used to overcome these challenges.

- Blue Shield of California is a nonprofit, tax-paying health plan, and an independent member of the Blue Shield Association. The pay-for-value strategy is focused on APMs that deliver high-quality care, lower costs, create an exceptional member and provider experience, and ultimately achieve optimal health and well-being for all Californians.
- Providers face many challenges in measuring and improving performance, including: 1) volume and variability of measures; 2) engaging specialists; 3) timely, accurate, and actionable analytics; and 4) patient attribution and risk stratification.
- Mr. Ramachandran highlighted some of the approaches Blue Shield of California has taken to overcome these challenges, including:
    - Partnering with many stakeholders, such as purchasers, providers, and payers, on harmonizing measures.
    - Collaborating with specialty associations to ensure that their perspectives are incorporated into the Blue Shield of California model.

    - Investing in technologies to manage data and create actionable analytics.
    - Embedding analytics into provider workflows without the added burden.
- The California Advanced Primary Care Initiative is one example of the partnership between Blue Shield of California and other stakeholders to harmonize measures.
- Mr. Ramachandran provided examples of collaborative efforts between Blue Shield of California and specialty associations and medical societies, including the American College of Cardiology, the California Orthopaedic Association, the California Medical Association, and the National Quality Forum.
- He shared an overview of the investments Blue Shield of California is making in data and technology to support data sharing and exchange, and internal integration of the data collected to improve workflows.
- He presented three key takeaways from the efforts undertaken at Blue Shield of California:
    - Harmonizing measures requires collaboration with purchasers, providers, and payers.
    - Collaboration helps ensure that the right measures are being utilized by specialty care providers.
    - Investing in actionable analytics allows providers to focus on delivering healthcare.

For additional details on Mr. Ramachandran's presentation, see the presentation slides (pages 38-48), transcript, and meeting recording (0:55-10:28).

Dana Gelb Safran presented on the challenges of developing measures for value-based payment and the methods being implemented by the National Quality Forum (NQF), including the Aligned Innovation initiative, to develop quality measures that reduce provider burden.

- She presented findings from a recent study conducted by NQF that sought to identify the perspectives of stakeholder groups (e.g., payer, provider, purchaser, patient advocate, and policymakers) on the state of quality measurement and its uses today.
    - The findings were not surprising and highlight the challenges that NQF is focused on solving, particularly the cost and length of time it takes to develop new measures, the number of measures available, and the burden of implementing measures without added benefit.
- Dr. Safran shared a visual that was developed by a subcommittee from the Health Care Payment Learning & Action Network (HCPLAN) in 2016 showing that value-based payment demands a shift from the "little dot" measures that are used now (i.e., the transactions or processes of care which are a by-product of an FFS model) to the more appropriate "big dot" measures that are needed under a value-based payment model.
    - Moving to "big dot" measures addresses parsimony because fewer measures are needed, and more importantly, identifies the outcomes that the value-based payment model is looking to achieve while leaving the process and how to achieve those outcomes to the providers.
- Dr. Safran presented the Alternative Quality Contract measure set that she developed in 2007. Measures included process, outcome, and patient experience for both ambulatory and hospital settings.
    - Value-based payment measure sets today look nearly the same as in 2007 despite over a decade of agreement and consensus regarding the need for outcome-oriented, "big dot" measures that are more in keeping with the goals of value-based payment.
- The challenges related to measures and methods that need to be addressed to optimize value-based payment results include:
    - Measures representing outcomes that matter.

- Data sources that increase the clinical value of the information while reducing burden.
    - Units of measurement that support accountability and improvement.
    - Alignment of measures, measure sets, and methods within and across payers.
    - Incentive structures that enable multi-year goal setting and motivate ongoing improvement.
- The Aligned Innovation initiative is designed to address these challenges to accelerate progress toward the next generation of measures for value-based payment that fill high-priority gaps and represent the outcomes that matter most to patients and clinicians.
    - The four differentiating factors of the Aligned Innovation initiative compared to traditional measure development include prospective alignment, patient-centered outcomes, broad diverse provider involvement, and timeframe.
    - The Aligned Innovation initiative also involves a multistakeholder advisory council (MAC) with five pillars of participation and representation from stakeholder groups, including accreditors and policy groups, payers and purchasers, patient and consumer advocacy groups, health information technology, and professional societies with expertise relevant to the topic of focus in a given cycle.
        - The MAC is representative of the end users and enablers of the measures being developed and ensures that their input is included throughout every step of the process, including piloting developed measures.
- An additional area of focus for NQF is ensuring that the data infrastructure in our country enables the goal of richer clinical information without added burden.
    - Providers today have an increased burden with EHRs compared to claims-based measures used previously.
- Dr. Safran highlighted four areas that NQF is focused to advance clinically-sourced measures while reducing burden:
    - Ensuring that the data elements needed for quality measurement and interoperability for clinical care are prioritized in the data standards by supporting ONC's continued evolution of the United States Core Data for Interoperability (USCDI) and USCDI Plus (USCDI+).
    - Leveraging artificial intelligence (AI) methods, including natural language processing (NLP), for quality measurement to enable continued use of EHR workflows that involve a combination of narrative entries and structured fields.
    - Advancing the integration of standardized PROMs into EHRs with automated longitudinal tracking and clinically useful information displays.
    - Pioneering standards by which to evaluate quality measures derived from AI/NLP methods.
- Dr. Safran shared seven key elements that are important design features for developing value-based payment incentive models that have been proven to enable ongoing performance improvement on outcomes and TCOC. Models should:
    - Offer a continuum of performance targets rather than a single cutoff or "cliffs."
    - Set absolute, not relative, benchmarks.
    - Set benchmarks for a multi-year period to allow for planning.
    - Ensure that quality earning potential is enough to be "worth it."
    - Include efficiency-tinged quality measures (this may be worthwhile even with shared savings).
    - Align measure sets across providers, payers, and programs.
    - Include real-time ability to track performance against targets.

For additional details on Dr. Safran's presentation, see the presentation slides (pages 49-59), transcript, and meeting recording (10:32-25:28).

Vivek Garg presented on patient and caregiver experiences in the context of PB-TCOC models.

- He shared a quote from Mark Twain comparing data to garbage, noting that if one collects data and does not know what to do with it, it begins to smell, particularly to clinicians. Therefore, the ways in which data are used to drive insight and action, and ultimately patient and caregiver experience, is most important.
- Dr. Garg presented an example of a balanced scorecard similar to the one used at CenterWell & Conviva Primary Care, a senior-focused primary care group under the umbrella of Humana. This scorecard is used by many value-based care practices to focus on panel management and population impact. The balanced scorecard includes a set of metrics and goals that cover broad domains of population management (e.g., engagement, patient experience and satisfaction, clinical quality, population outcomes and cost, and panel size and productivity).
    - Patient experience is just one important component, equally rated, but it is not the only component.
    - Additionally, groups such as CenterWell & Conviva Primary Care often create bonus programs to incentivize clinicians to complete balanced scorecards. The incentive is somewhere between 10-25% of clinicians' base salary and is needed to motivate clinicians to participate.
- Providers can use tools developed for the customer service industry to gain knowledge and insights into patient and caregiver experiences, including provider practice online reviews (e.g., Google), practice net promoter score (NPS), and call experience.
    - If a provider practice site is listed in Google and available on Google Maps, it is open to receive reviews. Therefore, practice sites should ensure that there is comprehensive and accurate information about their practice online (e.g., phone number, hours of operation, virtual care access, updated website, location) and curate, monitor, and respond to patient feedback.
    - Provider practices can partner with third-party platform tooling systems such as the National Research Corporation (NRC) Health to conduct NPS surveys and use those scores to examine trends over time, across clinics, and against benchmarks. Scores and patient comments can also be delivered to clinicians monthly.
    - Provider practices can listen to and monitor the quality of patient calls. This can include call handling (e.g., service level, abandonment rate, and transfer rate), patient service quality (e.g., NPS, resolution, agent satisfaction, quality), and center experience (e.g., patient scheduling, escalation turnaround time).
- Although the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey delivers crucial information, it is not timely or specific enough from a medical group perspective to drive the type of practice-level action needed to fully understand patient and caregiver experiences.
- Given the various measurement issues that currently exist (e.g., too many metrics, too much variation, and selectivity), CMS has promoted adoption of the Universal Foundation and quality framework. While there are adult and pediatric versions of the Universal Foundation measures, there is not a version for senior adults (e.g., ages 65 or older). A senior-specific Universal Foundation measure set could incorporate:
    - Barbara Starfield's "4Cs" framework of primary care: comprehension, first contact access, coordination, and continuity. In this framework, the CAHPS survey could be substituted by something such as the Person-Centered Primary Care Measure (PCPCM),

a validated 11-question survey advocated by the American Academy of Family Physicians (AAFP) and others, that assesses a patient's relationship with their primary care clinician and practice and whether they are getting the responsiveness that they need.
    - The 5 "Ms" of geriatric care: mind, mobility, medications, multi-complexity, and matters most. This framework helps to orient measurement about what is most important for patients to drive the experience and health outcomes that they are seeking.
    - PROMs as advocated by CMMI without allowing for too much selectivity across payment models and pilot programs, and instead moving towards national provider practice area-specific metrics and balanced scorecard-type approaches.
- In summary, there is a tremendous opportunity to create national reporting alignment for patient and caregiver experience and to deliver it in real time at the practice level. This would enable medical groups and practices to drive the process and programmatic changes needed to deliver on the modern expectations of patients. Taking a balanced scorecard approach, including consideration of a Universal Foundation measure set for senior adults, can help ensure that the data collected is not garbage, as Mark Twain suggested, rather that it is useful, well placed, and creates real action and uniform movement toward the experience all patients deserve.

For additional details on Dr. Garg's presentation, see the presentation slides (pages 60-68), transcript, and meeting recording (25:30-39:39).

Sai Ma presented on selecting health equity measures for PB-TCOC models.

- Health equity has become a fundamental priority for policymakers and industry leaders. Most efforts have been aimed at stratifying existing measures to identify disparities. For example, the NCQA has implemented race/ethnicity stratification for several HEDIS measures. Additionally, CMS is implementing the Health Equity Index (HEI) for Medicare Advantage programs.
- Methodology is important when it comes to stratification and has implications for how programs are designed and how resources are allocated. Methodological considerations include the quality of risk factor data; the interaction of risk factors, references, or benchmarks; absolute versus relative disparities; and within versus between disparities.
- Dr. Ma emphasized the importance of terminology and described the differences between health equity and health care equity.
    - Health equity means that everyone has a fair and equal opportunity to attain their highest level of health. Equitable health is the result of a broad spectrum of individual and societal factors experienced over one's lifetime.
    - Health care equity more narrowly describes the experience of accessing and interacting with a health care system and other organizations. Health care equity more directly examines whether a patient has equitable access, receives equitable care, and has an equitable experience along the care journey.
- Health care inequities that are measurable at the individual level, proximal to health care outcomes, and actionable are within the purview of health care organizations and should be prioritized by payers and providers.
- Dr. Ma shared a roadmap to help organizations identify the root cause of health inequalities and to determine where resources should be allocated to improve outcomes.
    - To use the roadmap, organizations need to identify an outcome they want to focus on and advance equitably. To diagnose the root cause of the health inequalities related to the outcome of interest, organizations can look at key performance indicators (KPIs) internally or published research to identify where the disparities exist along the care journey (e.g., prevention and access, transitions, quality of care, post-discharge).

- Elevance Health has developed a whole health index (WHI) measure to understand individual needs to better support and improve individuals' health. The WHI is a comprehensive measure that encompasses the major drivers of health and is currently being tested to see if it can be used to evaluate health care organizations' ability to address all major drivers of health.

For additional details on Dr. Ma's presentation, see the presentation slides (pages 69-74), transcript, and meeting recording (39:46-51:38).

Following the presentations, Committee members asked questions of the presenters. For more details on this discussion, see the transcript and meeting recording (51:42-1:26:05).

Presenters discussed the major challenges providers and health care systems experience implementing performance measures in PB-TCOC models.

- One challenge is a lack of alignment and uniformity across different payer models and programs. Clinicians want to look at the metrics across all patients to better understand patients' clinical needs, context, and goals. To overcome this challenge, groups such as CenterWell & Conviva Primary Care built massive data infrastructures to integrate the data as much as possible. A second issue is data explainability and interoperability. Every new metric must be explained to the clinical team so they understand it. Additionally, as new clinicians join the team, they must be trained on the value-based care model, including completing balanced scorecards, reviewing cost and utilization trends of their patient population, or drilling into specialist referrals.
- One challenge is the opportunity to plan for improvement on the measure set for which clinicians are accountable. This requires clinicians to have enough time before measure benchmarks are going to change. Multi-year contracts and models should allow the measure sets and performance targets to be fixed so those who are accountable for the measure sets and performance targets can plan their improvement strategy. In addition, having those performance targets set in absolute rather than relative terms is motivating and promotes best practice sharing because one's success is not impeded by someone else's success. A second challenge is knowing who internally can drive the improvement (e.g., the pharmacists, social workers, nurse practitioners, and physicians). Internal incentive structures and measures need to support external accountability for measures, targets, and improvements. A third challenge is determining whether something is worth it or not. Many of the value-based payment programs today put little money on quality. The quality measures are often just a gate to accessing shared savings. Last, many clinicians have not been in contracts before that made them accountable for TCOC and therefore need help understanding the data to determine where there might be waste and opportunities for shared savings.
- Most providers are using a balanced scorecard approach; however, there is a need to develop a more manageable balanced scorecard that is aligned across different payer models and programs. Clinicians need support in understanding how to bend the cost curve through timely, meaningful, and accurate methods.
- Not only are there too many measures, but there are also different measure specifications required by each regulator, highlighting the need for measurement reduction and alignment.

Presenters discussed clinician burnout and the idea of having clinicians give an NPS for internal systems as an indirect measure of their comfort level with the system and the system's support for them and their patients.

- No issue is clearer than when patients and doctors are saying the same thing in their different feedback mechanisms. CenterWell & Conviva Primary Care has an annual team member survey that is more general, although it is looking for survey options that are more clinician-specific to assess clinicians' perspectives on internal systems and level of burnout. Partner organizations

such as NRC Health can provide benchmarking from practices in the same geographic region and area of practice.
- Using a provider NPS as an accountability measure from the payer is concerning because the providers would likely feel a lot of pressure to make their system look good on the measure. The patient NPS would be a good accountability measure and tells a lot about the provider NPS and is much less gameable.

Dr. Safran discussed harmonizing measures through collaboration with key stakeholders and how to involve more stakeholders, specifically medical societies. The four medical societies involved in the Aligned Innovation initiative were explicitly targeted for the first cycle of work which is focused on behavioral health and maternal health measures. For every cycle of work, the clinical focus area shifts, and the appropriate medical societies are engaged. It is critical to ensure that the profession feels that the measures will have value, be fair, and are feasible.

Dr. Safran discussed the importance of analytics in driving change and the cost burden of defining equity, including the required reporting and analytics. There are two types of analytics to consider: 1) analytic tools that we need to inform clinical decision-making in clinical practice; and, 2) population-level analytics.

- The analytics needed to inform clinical decision-making should ideally be the requirement of the EHR vendor, but increasingly other solutions have to be incorporated. For example, the behavioral health measures being developed under the Aligned Innovation initiative (e.g., depression and anxiety for children, adolescents, and adults) are going to be PROMS. PROMS implementation has been one of the greatest pain points for practices. EHR vendors are slowly starting to incorporate PROMs; however, it would happen faster if payers increasingly started to demand the use of PROMS. In the meantime, NQF hired one of several solution vendors that is facilitating making it easier to collect the PROMS and enter the resulting data into the clinical record to trigger the longitudinal follow-up that needs to happen at the right moment. These vendors also have analytics that create a data display of the PROMS for the patient over time with benchmarks about other similar patients and how that score changed due to clinician intervention. This is a powerful tool that motivates clinicians to use PROMS.
- One of the biggest challenges with population-level analytics is when there is misalignment across payers such that different analytics are needed to work with data sets from every payer. Providers want to see measures for their entire population of patients, and then be able to parse by payer. There are efforts underway to create measure model alignment, which is not just alignment on the measures used but alignment on the data being collected, data compilation, analytics, and so forth.

Presenters discussed the essential measures that should be considered for health equity and the measures that have the largest impact on the improvement of health equity.

- Most of the effort to date is on stratifying existing quality and outcome measures. There are some measures in place to directly measure the root causes of needs such as health literacy, food insecurity, and transportation barriers. One measure that is a little further along than others is the social driver screening tool that the NCQA and CMS are pushing forward to allow providers and payers to collect information on health-related social needs (HRSNs). At a minimum, it can help identify patients' social drivers. There are other very specific health equity direct measures, such as access to care, patient engagement, and cultural competency. A formal process is needed to evaluate these and other health equity measures.
- The Universal Foundation includes universal screening for HRSNs. The burden of HRSNs is very high across the Medicare and Medicaid populations. There are a lot of operational challenges to

appropriately stratifying by race and ethnicity. The difference in outcomes stratified by race, ethnicity, and income have been known for decades; however, there is a limit to what medical groups can do to address these issues, some of which may be better addressed at the larger community and societal levels.

- One challenge provider organizations face is how to use data to make patients' health outcomes better. Therefore, it is important to help identify how provider organizations can take those data and connect to the community to help address HRSNs. Second, health equity indices are needed to tell organizations how they are doing. The index along with the granular information on the results is needed to inform improvements. Third, where differences in social risk among populations exist, it is unclear if the measure of performance or the financial payment should be adjusted. For example, these types of adjustments might include adding incentives for providers who take care of populations that have higher risk by awarding them money upfront, acknowledging that it might take something more or different to take care of this population, and rewarding a given level of performance more for providers taking care of a population that is viewed to be more difficult.
- As a payer, where can incentives be aligned to move the equity conversation forward? Are there coverage opportunities that we can just lean into and cover new services? Is there a place where we can add more providers to better reflect the patient population? There needs to be consistency in how data are collected and flexibility in how to solve problems using data to adjust for regional and provider variations.

Committee Discussion

Co-Chair Sinopoli opened the floor to Committee members to reflect on the day's presentations and discussions. The Committee members discussed the following topics. For additional details, please see the transcript and meeting recording (00:27-27:09).

- Health care has had a limited transition to becoming a high-performing industry.
- Performance measures need a clearly specified purpose. For example, process measures can be used for the purpose of testing theories about the link between process and outcomes instead of demonstrating a clear link between process and outcomes.
- Development of person-centered outcomes and patient-reported outcomes is needed.
- There are opportunities to develop CMS-level measures that could cascade to an entity-level, and then specific measures that can cascade to individual physicians depending on their specialty.
- Developing and implementing measures is expensive, and several new measures are still needed. Specifically, the development of specialty care measures, social determinants of health (SDOH) measures, and equity measures is needed. Equity continues to worsen, further increasing the need for risk stratification.
- Populations in TCOC models must be defined.
- We should be more comfortable with examining long-term outcomes instead of focusing on short-term outcomes.
- A shift is needed from using claims data and micro measures toward "big dot" measures. An all-payer model could allow for these efficiencies.
- Achieving harmonization requires collaboration and multi-payer solutions.
- Investments must be made in performance measurement.
- The entity must bear the risk and create the measures. Providers should not be at financial risk and instead be incentivized to meet the measures.

- The downward pressure of financial risk has led to unintended consequences, including eroded outcomes and trust and exacerbated inequities.
- TCOC models should focus on equity. Future PTAC public meetings could focus on definitions of equity, methodology, and risk adjustment, and how to define different subpopulations. It remains unclear who is responsible for care.
- In addition to measures of improvement, we must consider measures of accountability (e.g., safety, appropriateness, reporting) and measures for payment. All measures will likely be different.
- Multi-payer alignment and mandatory participation in programs may be necessary for success. Making program participation mandatory would require a deliberate glide path with an appropriate timeline for engagement and performance evaluation and improvement.
- Investments should be made in research focused on advancing health equity. To continue making improvements in the field in the future, we must also invest in the next generation of clinical scholars.
- "Gaming the system" is a prevalent issue in performance measurement.
- To avoid assuming that something is wrong with a patient who lacks trust in the health care system, we could consider developing a measure of trust. System trustworthiness protects the public and advances health equity at the individual level and community level. We may not be able to influence the community-level metrics, but we can advocate for community-level metrics.
- Providers experience challenges with performance measurement due to the complexity of measures, challenges with implementing measures, and using data. Providers need a clear, multi-year road map.
- We cannot forget about the quality of life of caregivers.
- Equity needs to be improved at a large scale, yet the work needs to be conducted on a patient-to-patient basis to understand how equity can be advanced within one's community.
- The portfolio of measures will likely not be the same across localities. However, all portfolios will include outcome measures, equity and appropriateness of care measures, and access to care measures with decreasing emphasis on process measures.
- There are a large number of measures that overburden providers. CMS should consider focusing on the "big dot" measures (i.e., lower cost, better care, better health).
- Applying "big dot" measures may have different meanings for different populations. We must consider at what level we make recommendations about measures. Although the "big dot" measures may be the same, the "sub-dot" measures will likely differ across populations.
- Capturing PROMs is challenging. PROMs are important to equity.
- We must consider use of AI, NLP, and Fast Healthcare Interoperability Resources (FHIR), and help EHR vendors to collaborate.
- Dashboards and other data visualization can help providers and systems better understand data.
- We must capture data from publicly available sources to help drive change.
- Use of claims-based measures will not drive change. We need to decrease the cost of reporting, decrease the cost of measure development, and rely on digital quality measures.
- We cannot assume that clinicians will know how to succeed using digital quality measures. Education on how to use these measures should be provided to clinicians.
- Gaps still exist in measures designed for specialty care, particularly measures that link specialty care to primary care.

Closing Remarks

Co-Chair Sinopoli adjourned the meeting.

The public meeting adjourned at 4:49 p.m. EDT.

Approved and certified by:

//Lisa Shats//	3/26/2024
____	________________________
Lisa Shats, Designated Federal Officer Physician-Focused Payment Model Technical Advisory Committee	Date
//Lauran Hardin//	3/26/2024
_____________________________________	________________________
Lauran Hardin, MSN, FAAN, Co-Chair Physician-Focused Payment Model Technical Advisory Committee	Date
//Angelo Sinopoli//	3/26/2024
______________________________________	________________________
Angelo Sinopoli, MD, Co-Chair Physician-Focused Payment Model Technical Advisory Committee	Date