Measures Management System
Information Session

# Strategies & Efforts for Improving Diagnostic Performance

**Presenters:**

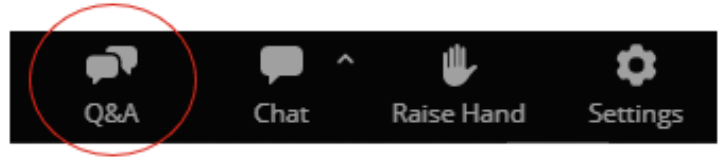
Sumant Ranji, MD
Ben Rosner, MD, PhD

# Want to Ask a Question?

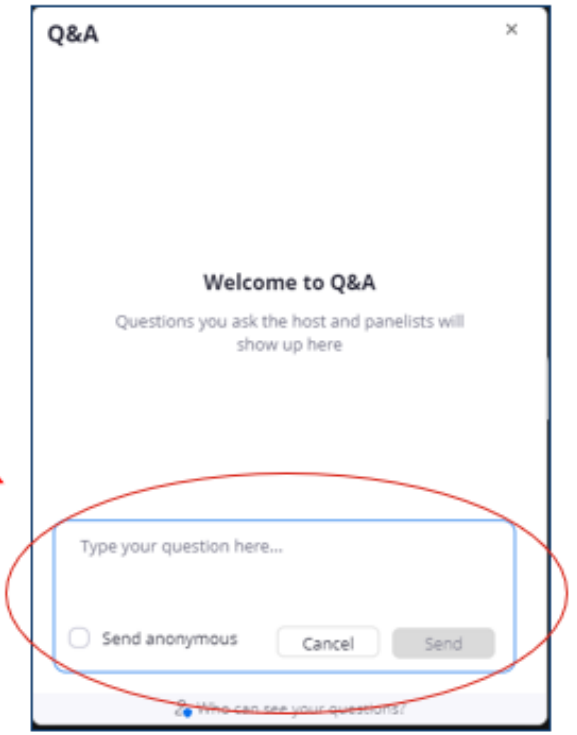
- Audience questions will be answered during the Q&A session at the end of the presentation.
- Instructions on how to submit questions:
  - Zoom Q&A Function
    - Please feel free to submit questions throughout the presentation.
- Note: If your question is not answered during the live Q&A, we will post FAQs to the CMS MMS Hub in a few weeks!

# Want to Ask a Question?
# Use the Zoom Q&A Function

Open the Zoom Q&A function




- Type your **question** into the question box
- Press **send** to submit



UCSF CODEX

# *UCSF Coordinating Center for Diagnostic Excellence*

Sumant Ranji, MD

Ben Rosner, MD, PhD

# *Diagnostic Excellence happens when clinicians and healthcare systems:*

- *Partner with patients to make accurate, timely, and equitable diagnoses;*
- *Use resources efficiently;*
- *And manage uncertainty.*

# *Diagnostic error in different clinical settings*

- *Outpatient: Estimated 1 in 20 adults will experience a diagnostic error (1)*
- *Inpatient: 23% of adult patients who were transferred to the ICU experienced a missed or delayed diagnosis (2)*
- *Emergency Department: 2.5 million misdiagnosis-related harms (3)*
- *Understudied: pediatrics, rural care, non-English speakers*

# *What is CODEX?*

# About CODEX

## Mission

CODEX will lead change in the field of diagnostic excellence by facilitating activities that result in measurable improvement in diagnostic quality, safety, and equity.

## Vision

CODEX will be a leader in transforming the health care system into one where all patients receive an accurate and timely diagnosis.

# *Our Goals*

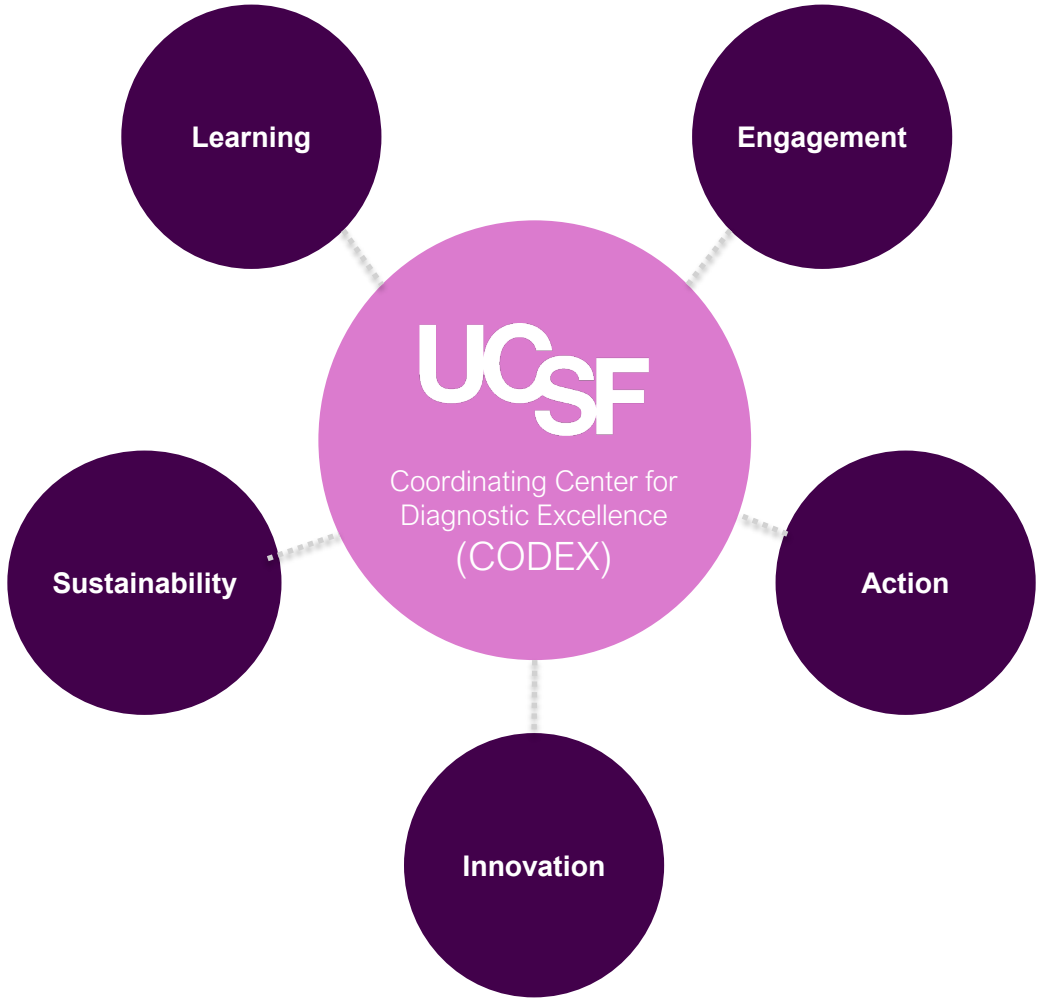
**01: Learning**
Become the leading source for objective information on advances in the diagnostic excellence field

**02: Engagement**
Build a community of diagnostic excellence scholars at UCSF and nationwide, informed by and responsive to patients' and other stakeholders' perspectives

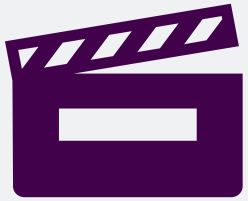
**03: Action**
Measurably improve diagnosis by convening committed stakeholders to target areas ripe for diagnostic improvement on a defined timeline

**04: Innovation**
Identify and facilitate novel programs to address existing and emerging challenges in the field

**05: Sustainability**
Secure funding to support the center's ongoing operations



# *CODEX Principal Domains –*
# *Work to Date*

| LEARNING HUB | ENGAGEMENT HUB | ACTION INCUBATOR |
|---|---|---|
|  |  |  |
| Synthesize and disseminate major advancements in diagnostic excellence research and clinical practice. | Convene key members of the national community for field-building and dissemination purposes. | Launch "action roundtables" and drive measurable improvements across groups of committed stakeholders. |

# *Learning Hub – Led by Anjana Sharma*

*1*

## CODEX Digest

*Spend less time searching and more time learning*

Stay current with the CODEX Digest, which cuts through the noise with a list of recent, must-read publications handpicked by us. Subscribe to receive the Digest directly in your inbox every Thursday.

*2*

## Editor's Picks

*A standout article moving diagnostic excellence forward*

These pieces offer meaningful, patient-centered insights, use innovative approaches, and speak to the needs of patients, clinicians, researchers, and decision-makers alike.

*3*

## DxEx Primers

*3-part introduction to the field of diagnostic excellence*

These primers were created to address common misconceptions about diagnostic error, create a strong knowledge foundation, and provide essential tools for enhancing diagnostic accuracy.

*4*

## Monthly Webinars

*Innovations advancing real-world improvements in diagnosis*

Hear from researchers, providers, and the next generation of diagnostic excellence leaders about the path toward meaningful improvements for health care delivery and patient care outcomes.

Read More

http://tiny.ucsf.edu/BOC5i7

# CODEX Digest: October 2, 2025

Spend less time searching and more time learning.

codex.ucsf.edu/news/digest

This week's digest features research on the fatal consequences of misdiagnosing neonatal sepsis, challenges in reliably measuring misdiagnosis in pediatric emergency departments, and barriers to implementing clinical decision support tools in primary care. Also highlighted are studies using AI tools for assessing patient-submitted surgical images, training nursing students in diagnostic reasoning, and supporting pathologists in clinical assessments.

Explore the digest below or view it on our website.

*Titles link to the PubMed record or free-to-access sites with full text availability unless otherwise noted.*

**Diagnostic excellence: turning to diagnostic performance improvement.** (subscription required)

Auerbach A, Raffel K, Rasooly IR, et al. Diagnosis (Berl). Epub 2025 Sep 16

Organizational commitment to diagnostic improvement is necessary to measure error and achieve excellence. This commentary reviews the importance of Diagnostic Excellence Programs, shifting from an epidemiologic approach to targeting overall performance. System-focused initiatives, individual clinician behavior enhancements, patient engagement strategies, and cautious AI implementation are recommended for diagnostic excellence program development.

**Recognizing atrial flutter in the emergency department: challenges in diagnosis.** (subscription required)

Avidan Y, Aker A, Sliman H, et al. Am J Emerg Med. 2025;96:224-229.

ECG misinterpretation can lead to cardiovascular disorder misdiagnoses. This retrospective Israeli single-center analysis looked at 2,003 emergency department ECGs of patients discharged with atrial fibrillation (AF) or atrial flutter (AFL) diagnoses to assess accuracy of non-computer-aided conclusions. Results showed 44% of the 209 AFL cases to be misdiagnosed as AF cases. The authors found factors associated with misdiagnosis to be older age and prior AF diagnosis and call for targeted education rather than improving computerized accuracy.

**Exploring the risks of over-reliance on AI in diagnostic pathology. What lessons can be learned to support the training of young pathologists?**

Bellahsen-Harrar Y, Lubrano M, Lépine C, et al. PLoS ONE. 2025;20(8):e0323270.

Independent pathologist judgement is a valuable skill when using AI for slide assessments. This French study contrasted the diagnostic accuracy of eight pathologists with varying levels of experience in cases with and without AI assistance. AI-supported pathologists achieved higher accuracy, but there was some evidence of overreliance and automation bias with AI in less-experienced pathologists who ignored lower-confidence ratings coming from the AI model.

**Implementation of a quality improvement and clinical decision support tool for cancer diagnosis in primary care: process evaluation.**

Chima S, Hunter B, Martinez-Gutierrez J, et al. JMIR Cancer. 2025;11:e65461.

Clinical decision support (CDS) is an important contributor to diagnostic excellence. This mixed-method analysis surfaced successes and barriers to the use of Future Health Today, a CDS tool with quality improvement support, supporting timely diagnosis of cancer in 21 Australian primary care practices. Uptake of the tool was low with variable numbers of patients flagged for cancer. The results indicate that time, resources, individual practice characteristics, and cumbersome application support were barriers to implementation of a primary care CDS tool.

# Engagement Hub

*Supporting the growing DxEx field*

*Direct support for:*

- *DEX Annual Conferences*
- *Patients for Patient Safety US*
- *NASEM Forum Advancing Diagnostic Excellence*
- *NAM/CMSS Scholars in Diagnostic Excellences Fellowship*




# *Current State of Measurement of Diagnostic Excellence*

*1*

## *CMS*

2025 Measures Under Consideration (MUC) include 2 diagnostic excellence measures

*2*

## *Leapfrog*

Leapfrog has added questions to its ratings survey to assess what actions hospitals are taking to reduce harm to patients from diagnostic errors, including delayed, wrong, or missed diagnoses, and diagnoses not communicated to the patient.

*3*

## *AHRQ*

14 DxEx measures under consideration for AHRQ Quality Indicators Program

# *Current State of Measurement of Diagnostic Excellence*

*Gordon and Betty Moore Foundation Measurement Portfolio*

Pre-Rulemaking and Use in Federal Quality Programs and Consensus-Based Entity Endorsement:

| Measure Title | Measure Developer | Current Status |
|---|---|---|
| CVD Risk Assessment Measure | University of California, Irvine | Implemented in MIPS |
| Appropriate Germline Testing for Ovarian Cancer Patients | American Society of Clinical Oncology | Implemented in MIPS |
| Delay of Venous Thromboembolism in Primary Care | Brigham and Women's Hospital | Included in the proposed rule for use in the MIPS program |
| Follow-Up on Abnormal Screening Mammograms for Breast Cancer | Brigham and Women's Hospital | Included on 2025 MUC List |
| Follow-Up on Positive Stool-based Tests for Colorectal Cancer | Brigham and Women's Hospital | Included on 2025 MUC List |

| CBE ID | Measure Title | Measure Developer | Endorsement Status |
|---|---|---|---|
| 4440e | % hospitalized pneumonia patients with chest imaging confirmation | University of Utah | Endorsed |
| 3671/3690 | Inappropriate dx of CAP<br>Inappropriate dx of UTI | University of Michigan | Endorsed |
| 3617 | Provider Level Continuity of Care | American Board of Family Medicine | Endorsed |
| 3749e | Delay of Venous Thromboembolism in Primary Care | Brigham and Women's Hospital | Endorsed |

# *Current State of Measurement of Diagnostic Excellence*

*Gordon and Betty Moore Foundation Measurement Portfolio*

American College of Emergency Physicians CEDR Registry Use (2026 Implementation):

| Measure Title | Measure Developer |
|---|---|
| Diagnostic Utilization – Proportion of all ED visits with chest CT ordered to investigate PE | Yale University, ACEP |
| Diagnostic Yield – Proportion of chest CTs ordered with Acute PE | Yale University, ACEP |
| Diagnostic Opportunity – Proportion of ED diagnostic opportunities 0-7 days prior to ED PE with PE diagnosis | Yale University, ACEP |

# *AI and Diagnostic Excellence*

## Benefits

- Predictive analytics from processing large amounts of data across multiple sources (e.g., EHRs, medical databases)
- Image analysis and decision support in radiology, pathology, ophthalmology, and dermatology
- Digital diagnostic "wayfinding"
- AI-powered clinical decision support in non-image-based field
- More opportunities for patient engagement and partnership in the diagnostic process
- Optimize clinician cognitive load by taking on administrative tasks (e.g., patient portal messages, clinical documentation)
- Potential for increased access and more personalized care

## Risks and unintended consequences

- Cognitive deskilling
- Complacency, anchoring, and automation biases
- Overdiagnosis and overtesting
- Perpetuating or amplifying health inequities from biased or insufficient data on racial and ethnic minorities
- Increasing digital divides

# *Action Incubator*

# *How can we ensure that AI achieves diagnostic excellence?*

# *Model Development*

## *Landscape Review and Iteration*

Has anyone done this before? No

- Pulled from other models based on our aims and resources
- Key informant interviews with leaders in diagnosis, AI, policy, and measurement space
- Review and continuous iteration with senior advisors

## *Stakeholder Identification*

Who hasn't been in the same room before?

- Health system representatives committed to implementing selected measures in-kind

# *UCSF CODEX's Inaugural Action Incubator: Conceptual model*

*1*

**Identify key stakeholders**

-6 months

*2*

**In-person meeting**

Month 0

*3*

**Develop actionable measurement framework**

Month 6

*4*

**Pilot framework across participating health systems**

Months 6 -18

# *In-Person Meeting*

September 11-12, 2025

*How do we measure diagnostic excellence in the context of ambient scribes and EHR-embedded AI tools?*

## *Participants*

Participants included 30 healthcare leaders spanning:

- Industry: insurance, malpractice, vendors
- Health systems: community-based, AMCs, integrated
- Patients and patient advocates
- Federal payors

## *Format*

- Facilitator to ensure engagement, shared decision-making, and consensus-building
- Framing the challenge to cover current issues with measuring diagnostic excellence
- Presentations from key experts on current landscape and future state of AI
- Breakout sessions including representatives from healthcare systems and clinicians from ambulatory, emergency department, and acute care
- Group discussions to share out key insights

# *In-Person Meeting*

September 11-12, 2025

## Key takeaways

- Widespread AI implementation in radiology has improved productivity – and introduced workflow changes and biases
- Implementation of AI tools for ambulatory clinicians has improved satisfaction but had mixed effects on cognitive load
- Measures of AI for DxEx should build on existing measure sets whenever possible
- As patients increasingly use AI, metrics need to incorporate patient engagement in the diagnostic process

# *Evaluation in the Context of Current State of Widely Deployed AI*

Traditional Care

AI Scribe

**Issue Brief 17**

## Cognitive Load Theory and Its Impact on Diagnostic Accuracy

Cognitive load is an independent variable that affects diagnostic accuracy and should be accounted for when designing clinician workforce structures.

# *Identifying candidate measures through a consensus process*

*1*

**Synthesize insights into an actionable measurement framework**
*Refine insights gathered from in-person meeting*

*2*

**Continuous participant feedback**
*Regularly convene webinars to continue follow-up discussion and refine measures*

*3*

**Modified Delphi panels**
*Rank candidate metrics based on feasibility, size of impact, causal association with AI, and low burden on health system participants*

# *In-progress: impact of AI scribes on clinical follow-up*

## Current state – AI scribes

- AHRQ's quality indicator of timely follow-up on abnormal FIT test or mammography
- Time to follow up on abnormal test results
- Patient-facing survey question (under development)

## Future state - EHR-embedded, AI-based diagnostic decision support

- Scaling eTriggers of diagnostic excellence by automated LLM chart review/adjudication
- LLM as a tool to assess quality of diagnostic decision-making in encounter notes
- Measuring proportion of clinician listening time

UCSF CODEX

# *Thank you!*

codex.ucsf.edu
codex@ucsf.edu
@ucsfcodex
ucsfcodex.bsky.social

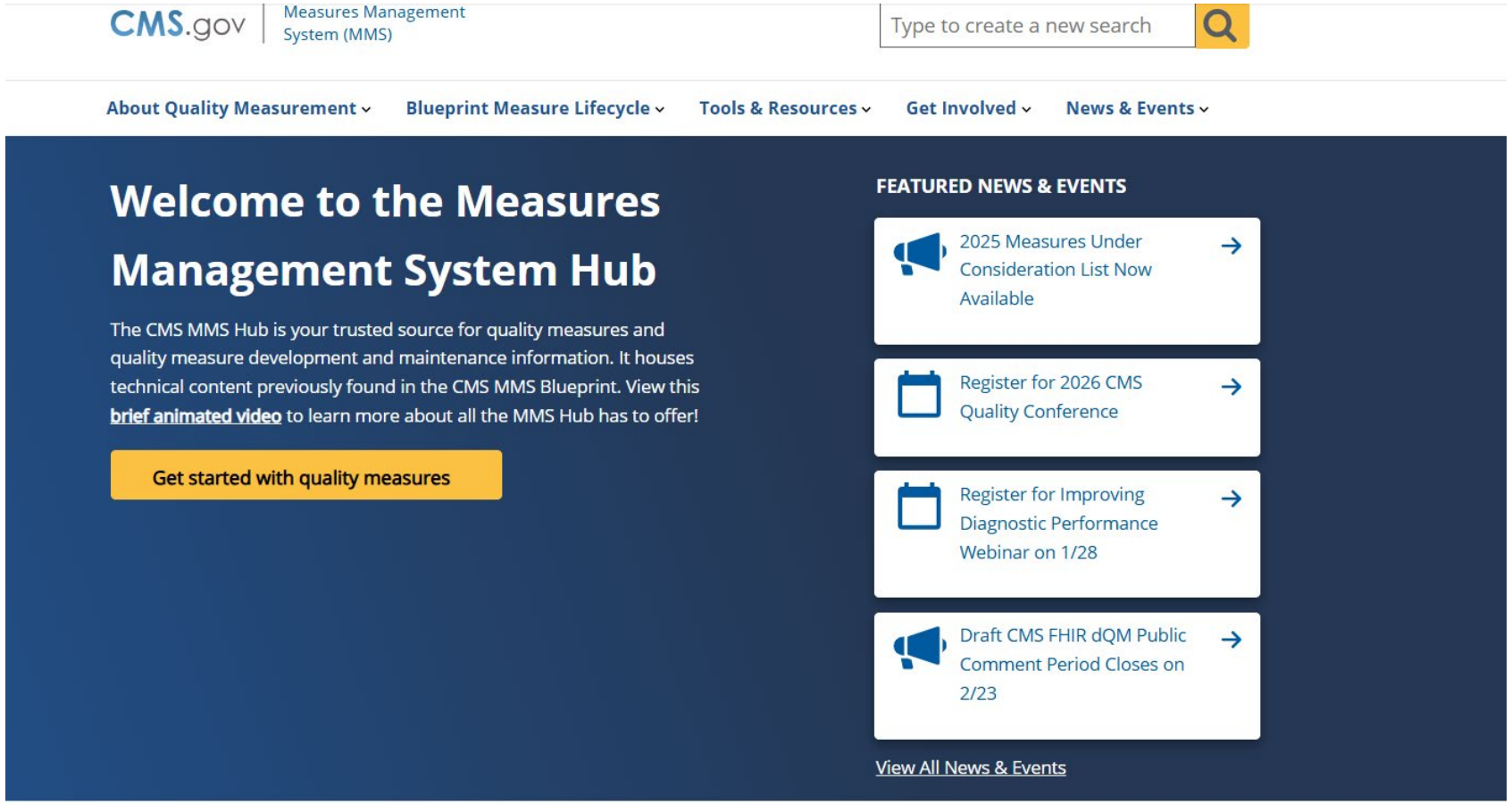

# Questions?

# Want to Ask a Question?

- Audience questions will be answered during the Q&A session at the end of the presentation.
- Instructions on how to submit questions:
  - Zoom Q&A Function
    - Please feel free to submit questions throughout the presentation.
- Note: If your question is not answered during the live Q&A, we will post FAQs to the CMS MMS Hub in a few weeks!

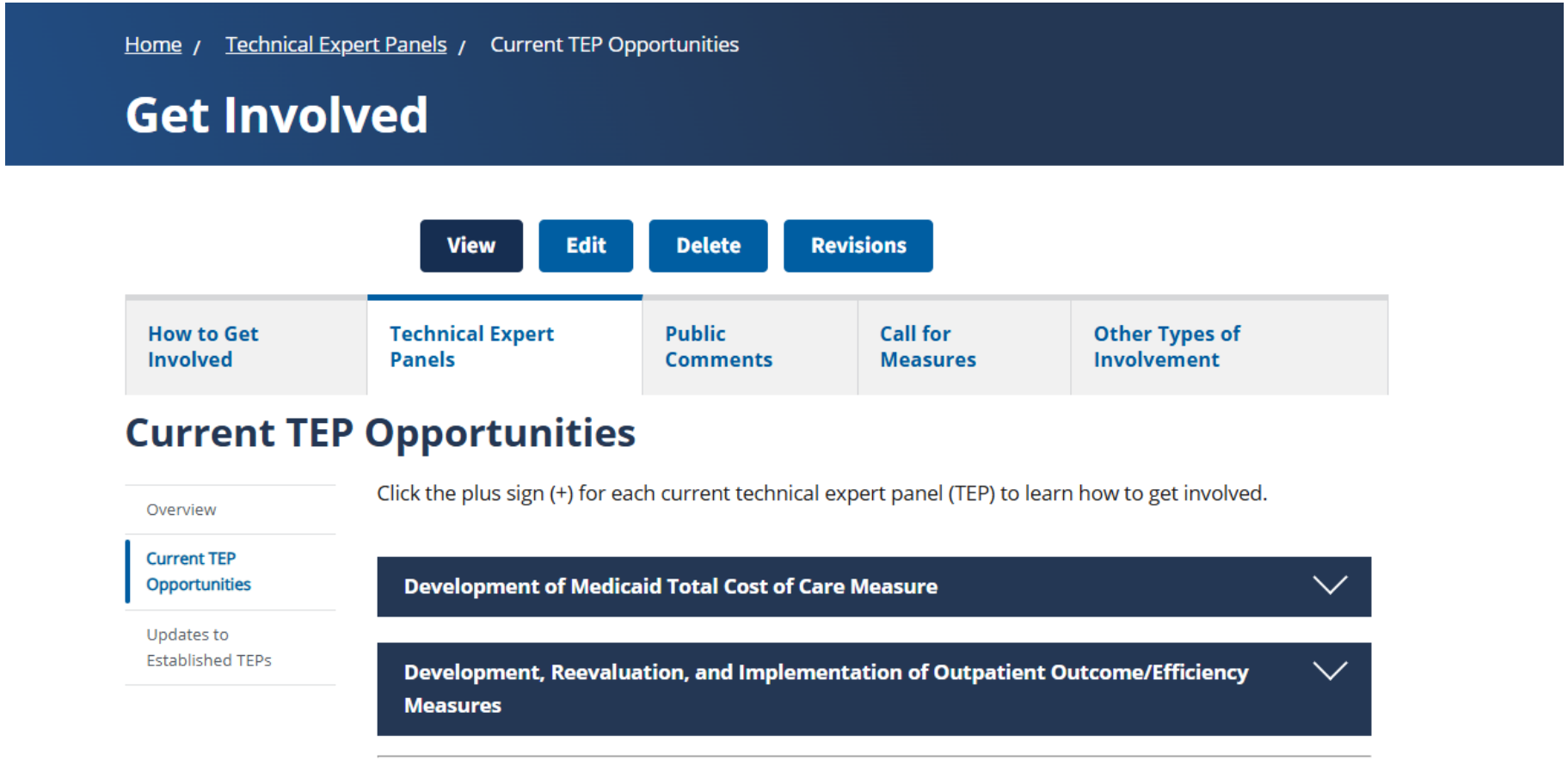

https://www.cmsqualcon.com/

# Get Connected




Visit: https://mmshub.cms.gov/

# Get Involved




https://mmshub.cms.gov/get-involved/technical-expert-panel/current


**Battelle**
**MMSSupport@battelle.org**

**CMS**
**Gequincia Polk**
**gequincia.polk@cms.hhs.gov**