Artificial Intelligence (AI) in Action: Generating Claims about Measure Properties

Jeffrey Geppert,
Battelle
April 2, 2025

Consensus-Based Entity AI Pilot

- Goal: to assess the potential for artificial intelligence to support human review of clinical quality measures in CBE processes (E&M, PRMR, MSR)
- Evaluating Claims about Measure Properties
  - PQM Educational Webinar
  - April 2025
- **Generating Claims about Measure Properties**
  - **MMS Information Session**
  - **April 2025**
- Evaluating Importance (Impact) Claims about Measure Properties
  - Forum TBD
  - Date TBD

CBE Strategy: Vision

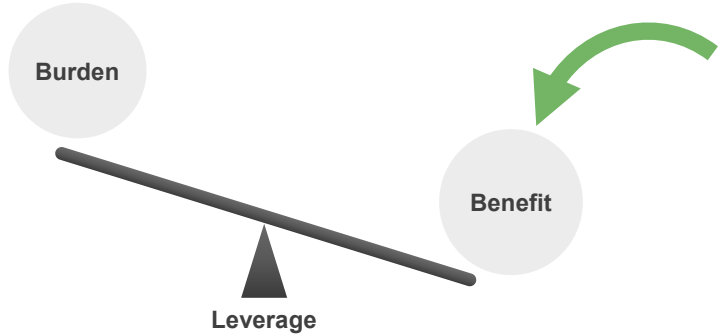
> *When [evidence-based] health and health care policies and programs designed to improve outcomes are not driven by community interests, concerns, **assets**, and needs, these efforts remain disconnected from the people they intend to serve. This disconnect ultimately limits the influence and effectiveness of interventions, policies, and programs.*
>
> *– National Academies of Medicine (NAM), February 14, 2022*

Vision: The vision for the CBE is to realize **health care system change** through the integration of quality measurement and quality improvement processes, and to align the principles of **evidence-based policies and programs** and **meaningful community engagement**.

CBE Strategy: Critical Obstacle

Focus quality measurement where there is

the most benefit **for health care system change**



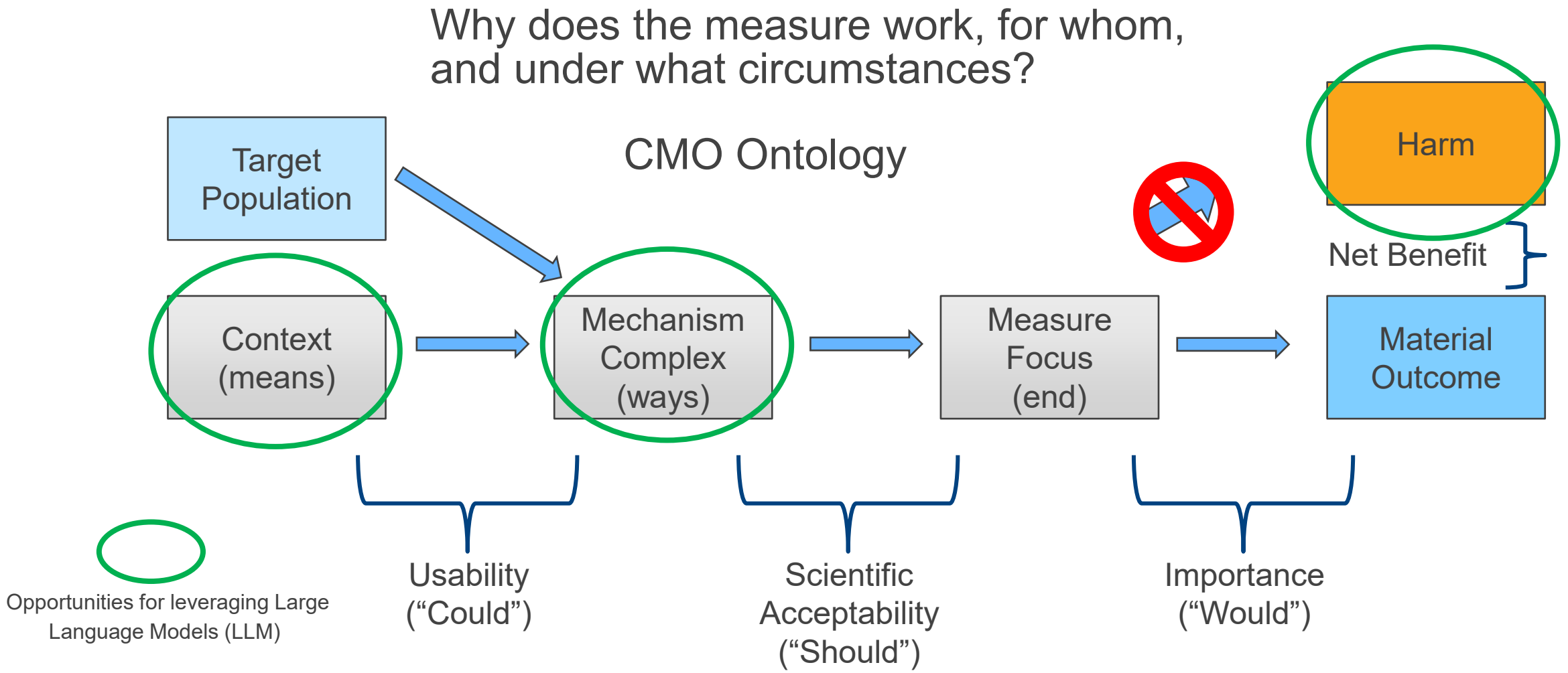
RISK

IMPACT

Impact of measurement	Risk of measurement: **Low uncertainty** (Mechanisms are systemic and persistent; evidence is mature)	Risk of measurement: **High uncertainty** (Mechanisms are not systemic and persistent; evidence is not mature)
Low (few persons and entities) (Magnitude of improvement to benchmark is low; magnitude of mechanism effect is low)	Do not measure (**accept** the risk of low quality)	Quality improvement (**transfer** the risk of low quality)
High (many persons and entities) (Magnitude of improvement to the benchmark is high; magnitude of mechanism effect is high)	Mitigation or monitoring (**control** the risk of low quality)	Quality measurement (**avoid** the risk of low quality)

BATTELLE

Purpose: What change/transformation are you trying to make?

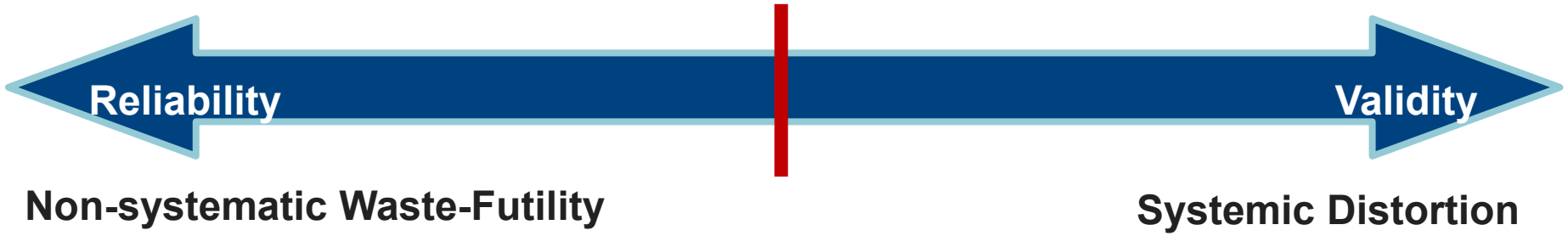


BATTELLE

What is Validity?

"Validity is measuring the right thing; reliability is measuring the thing right"
– Thissen (2001)

Validity is "an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy of appropriate interpretations and actions on the basis of [the measure]"
– Messick (1989); Standards for Educational and Psychological Testing (2014)



BATTELLE

Measure Evaluation – Substantiating Claims

Measure developers and/or measure stewards make certain explicit or *implicit* assertions or claims about the potential benefits and risks/harms associated with measure use (net benefit).

In general, there are three top-level claims related to measure properties necessary for a measure to yield positive net benefit to persons and entities:

***Would* claim:**	Person or entity *would* make decisions based on the measure because the measure focus is associated with a material outcome (end/importance).
***Should* claim:**	There are known and effective ways of selection or choice that the person or entity *should* use (ways/scientific acceptability). ▪ Known: mechanism complex; Effective: causal
***Could* claim:**	Any barriers or facilitators to whether the person or entity *could* use those ways are known and addressed (means/usability).

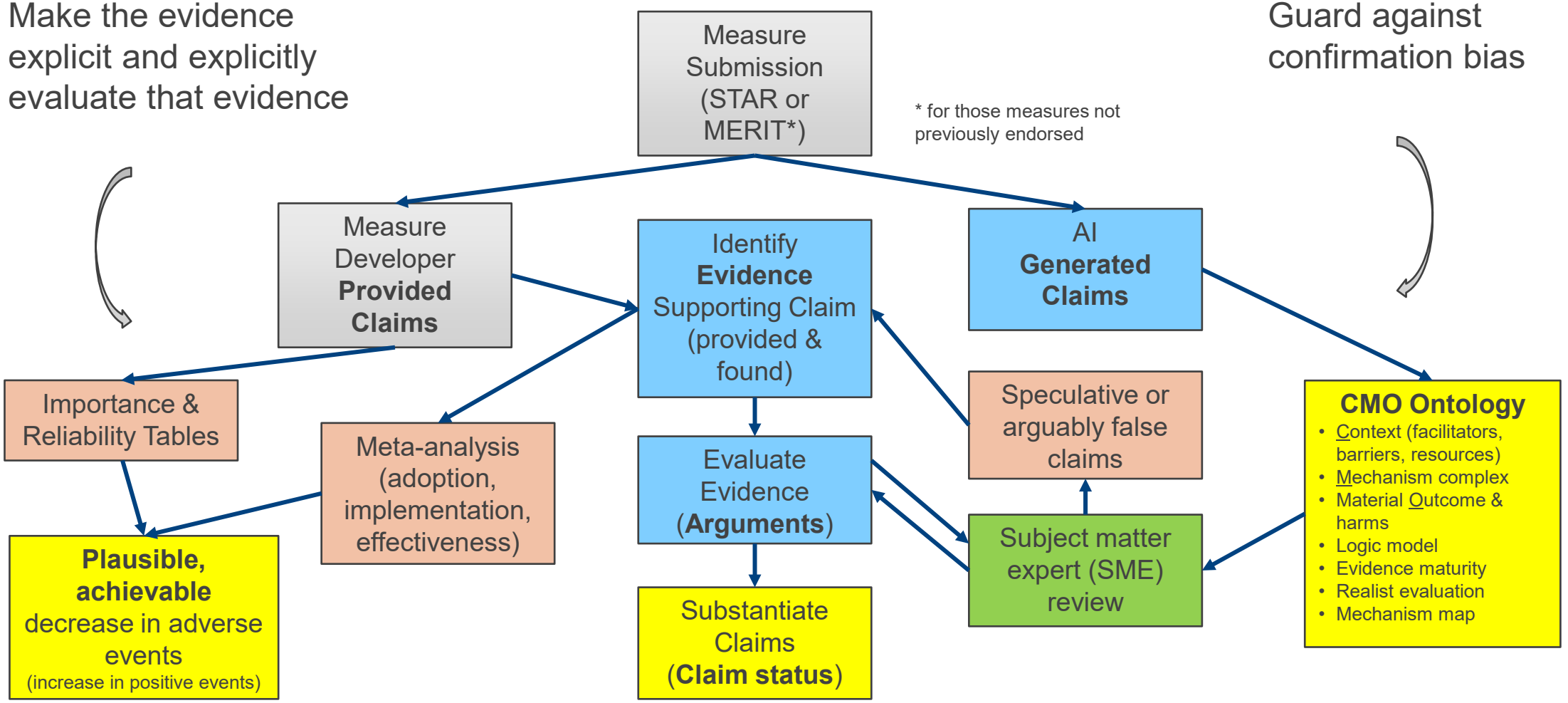
Measure Evaluation – Assurance Cases (CAE)

- Goal: assure trustworthy clinical quality measures
  - Make the evidence explicit and explicitly evaluate that evidence
    - Identifying, assessing, and summarizing literature is time and resource intensive
  - Guard against the potential for “confirmation bias”
    - The tendency to process and interpreting information in a manner that is consistent with existing beliefs
- Approach: “assurance cases” – Claim-argument-evidence (NIST; ISO)
  - Claim (property of measure)
    - Provided by measure developer or generated by AI (ontology, persona, and context)
  - Evidence (in support of claim)
    - Including expertise, experience, logic, empirical, computational, simulation, engineering
  - Argument (why evidence supports claim)
    - Logical inference: deduction, induction, or abduction (inference to best explanation)

Measure Evaluation – Assurance Cases (CAE)

Make the evidence explicit and explicitly evaluate that evidence

Guard against confirmation bias

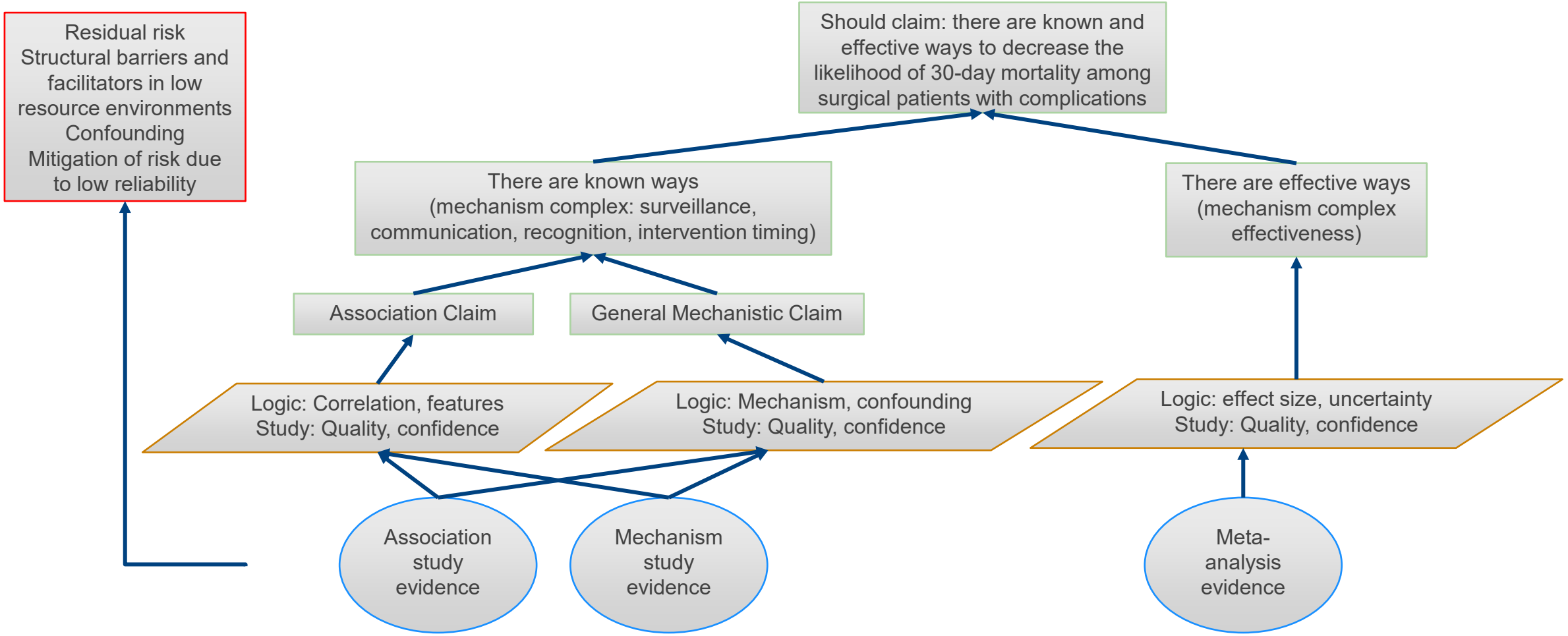


BATTELLE

Measure Evaluation – Claim and Sub-claim Types

Importance	Would claim: Person or entity would make decisions based on the measure because the measure focus is associated with a material outcome
Validity	Should claim: There are known and effective ways of selection and choice that the person or entity should use
-Association	There is an association between the person or entity response to the measure and the measure focus
-Mechanism	There is an explicit articulation of the mechanisms (resources and response to those resources) responsible for the association
Usability	Could claim: Any barriers or facilitators to whether the person or entity could use those ways are known and addressed

Measure Evaluation - Assurance Cases (FTR)



BATTELLE

Measure Evaluation – AI Pilot

- What does the AI Pilot do?
  - For each claim, identifies and assesses evidence, and summarizes arguments
    - uses natural language processing (NLP) to identify evidence that is related to the claim
    - uses a large language model (LLM)-powered AI agent (Agentic AI) to assess evidence and summarize arguments
    - **uses an LLM (e.g., ChatGPT) to generate claims based on the context-mechanism-outcome (CMO) ontology**
- When to use AI Pilot?
  - To evaluate evidence from published sources provided in measure submissions or other documentation
  - **To identify evidence from published sources for an otherwise unsubstantiated claim**
  - Not to evaluate clinical practice guidelines or systematic reviews (i.e., have a GRADE assigned)
  - Not grey literature (book chapters, reports, etc.)

AI Pilot Study Process

1. Measure developer provides a full measure submission (FMS) for the measure
2. CBE staff review the measure submission and manually extract the **measure developer** provided claims using an Evidence Table template
3. AI generates additional claims based on the **CMO ontology**
4. The AI Agent performs the following tasks for the **measure developer** provided claims or AI generated **CMO ontology** claims:
   a. Identifies the published abstract for the evidence cited in support of the claim
   b. Identifies additional published abstracts for evidence relevant to the claim (up to 5)
   c. Determines whether the evidence agrees, disagrees, or is neutral toward the claim
   d. Evaluates the quality level and confidence level of the evidence
   e. Determine the status of the claim (established, speculative, or ruled-out)
   f. Combines the three top level claims to determine the likelihood of endorsement (endorsable, potentially endorsable, unlikely endurable)

AI Pilot Study: AI Generated CMO Ontology Claims

1. CBE staff and the measure developer evaluate the performance for the AI generated **CMO ontology** claims
   a. The proportion of claims that are established, speculative, or rule-out
   b. The proportion of claims that are asserted by the **CMO ontology** (relative to the **measure developer** provided claims)
   c. How plausible are the justifications?
   d. Identification of evidence for selected **CMO ontology** claims
   e. The proportion of claims that are established, speculative, or rule-out with the additional evidence

AI Pilot Study – SME Review

- Documents for SME review
  - AI Generated CMO Ontology Claims
    - Context, mechanism, and material outcomes
    - Evidence maturity
    - Logic model, mechanism map
    - Realist evaluation (why does the measure work, for whom, under what circumstances?)
  - Measure Developer Claims
    - Evidence Table Export (provided and found evidence)
    - Document: endorsement status
    - Claims: claim type, claim status (e.g., established), GRADE, justification
    - Claims-Evidence: title, author, abstract, source (provided, found), study type, quality level, confidence level, justification, claim, agreement (agree, disagree, neither), justification agreement, journal, date, URL, PMID

AI Pilot Study – Sample Results

- A sample of results selected from among the following typical CQM:
  - CBE2023-4440e. Percent of hospitalized pneumonia patients with chest imaging confirmation (Developer: the University of Utah)
  - CBE2020-0071. Persistence of Beta-Blocker Treatment after a Heart Attack (Developer: National Committee for Quality Assurance (NCQA))
  - CBE2019-3512. Knee Arthroplasty Cost Measure (Developer: Centers for Medicare & Medicaid Services)
  - CBE2019-0753. 30-Day Post-Operative Colon Surgery (COLO) Surgical Site Infection (SSI) Standardized Infection Ratio (SIR) (Developer: Centers for Disease Control (CDC))
- Reach out to PQMSupport@battelle.org if you would like to work with us to apply these AI tools to a CQM of interest

AI Pilot Study – LLM: Persona and Context

(Persona): You are an evaluator of clinical quality measures. Your role is to understand the context, mechanisms, and outcomes that explain how a measure works, for whom, and in what circumstances.

(Context): Consider a **Measure of Interest** of the risk-adjusted standardized infection ratio (SIR) of observed over predicted deep incisional primary and organ/space surgical site infections (SSIs), over a 30-day post-operative surveillance period, among hospitalized adults who are >=18 year of age with a date of admission and date of discharge that are different calendar days, and the patient underwent a colon surgery (COLO) at an acute care hospital or oncology hospital. The 30-day postoperative surveillance period includes SSIs detected upon admission to the facility or a readmission to the same facility or a different facility (other than where the procedure was performed) and via post-discharge surveillance.

(Event): The **Measure Focus** for this Measure of Interest is development of a deep incisional primary or organ/space surgical site infection (SSI) within the 30-day postoperative surveillance period. The 30-day postoperative surveillance period includes SSIs detected upon admission to the facility or a readmission to the same facility or a different facility (other than where the procedure was performed) and via post-discharge surveillance.

(Person experiencing event): The **Target Population** for this Measure of Interest is persons aged 18 years and older with a date of admission and date of discharge on different calendar days, and with a procedure for colon surgery (COLO).

The **Entity of Interest** is the facility (acute inpatient hospital or oncology hospital).
The Person of Interest is a person in the Target Population.

AI Pilot Study – LLM: CMO Ontology

- Material Outcome

Structure, process or intermediate outcome: Among the Target Population would you explain the association between the Measure Focus and a material health outcome?

Outcome: Among the Target Population would you explain the rational for considering the Measure Focus a material health outcome?

Are there contexts where better performance on [increasing, decreasing] the likelihood of the Measure Focus might result in harm to persons among the Target Population?

Would you explain under what conditions the following claim might be true or not true: "better performance on the mechanism complex is causally associated with better performance on [increasing, decreasing] the likelihood of the Measure Focus.“

AI Pilot Study – LLM: CMO Ontology

- Mechanism

Would you explain a mechanism complex responsible for [increasing, decreasing] the likelihood of the Measure Focus among the Target Population?

Would you explain how contextual mechanisms may work to reinforce or counter-act this mechanism complex either in whole or in part?

Would you describe a logic model for [increasing, decreasing] the likelihood of the Measure Focus among the Target Population, including inputs, activities, outputs, short-, intermediate-, and long-term outcomes?

Does the logic model include any broad, systemic changes, feedback mechanisms, assumptions, and external factors?

Would you describe and draw a mechanism map for the mechanism complex?

AI Pilot Study – LLM: CMO Ontology

- Context

Are there resources to support implementation of the mechanism complex for [increasing, decreasing] the likelihood of the Measure Focus among the Target Population?

Are there barriers or facilitators to implementing these resources?

In conducting a realist evaluation, we are interested in why a measure works, for whom, and in what circumstances.

- ***A measure might work by enabling or blocking choice-making by an entity. Similarly, a measure might work by enabling or blocking reasoning by an entity (why).***
- ***However, not every entity has the capacity or resources to make that choice. Similarly, not every entity has the capacity or resources to conduct that reasoning (by whom).***
- ***Finally, the entity might operate in a cultural or physical or social context that activates or does not activate the operation of the choice-making. Similarly, the entity might operate in a cultural or physical or social context that activates or does not activate the operation of the reasoning (in what circumstances).***

For the measure of interest, conduct a realist evaluation.

Material Outcome: CBE2024-4440e

Inputs

- Bedside X-ray
- Portable CT scanners
- Radiology department
- Clinician training
- EHR / CDS / ordering
- Risk stratification tools (CURB-65)
- Patient consent
- Patient education
- POC Ultrasound

Mediators

- Age-perceived risk
- Comorbidities (COPD, HF)-differential DX
- Severe symptoms (hypoxia, fever, mental status, dementia)-perceived urgency
- Delayed presentation
- Incomplete medical history

	Benefits*	Harms*
Pneumonia (high risk) (actively treated for suspected infection)	• Risk stratification (correct management-antimicrobial selection, dose, duration) • Reduced complications (pleural effusion, abscess) • Reduce mortality (severe pneumonia, sepsis, RF)	• Antimicrobial resistance • Incidental findings • Treatment delays • Reduced clinical judgement
Not pneumonia (low risk) (acute bronchitis)	• Decrease in unnecessary antimicrobial use • Decrease in antimicrobial-related complications • Diagnosis of true condition	• Increased costs • Exposure to Rx • Resource diversion • Missed DX for atypical presentations

Avoidable utilization

- Imaging
- Readmissions
- Antimicrobial stewardship (duration)

Moderators

- Image timing (within 48 hours)
- Radiological expertise
- Specialty tele-consult
- Volume (capacity/resource constraints)
- Delayed interpretation (off-site)

BATTELLE

Material Outcome: CBE2019-0071

Mechanism Complex

- Guideline-Concordant Prescribing
- Risk Stratification
- Discharge Planning and Care Transitions
- Health Literacy and Medication Understanding
- Behavioral Cues and Habit Formation
- Perceived Benefits vs. Side Effects
- Provider education

Mediators

- Age (frailty)
- Comorbidities (e.g., heart failure, COPD/ obstructive chronic bronchitis, diabetes)
- Contraindications (e.g., bradycardia, hypotension, or asthma)
- Intolerance or allergy
- Chronic respiratory conditions due to fumes and vapors
- Genetic and biological variability
- Socioeconomic factors (income, transportation, social support)
- Health literacy or cognitive impairment (adherence)
- Cultural preference for non-pharmacologic approaches

Benefits*	Harms*
• Reduced all-cause and cardiovascular mortality • Lower rates of recurrent myocardial infarction • Improves preservation of left ventricular function	• Clinical Harm from Inappropriate Beta-Blocker Use • Harm from Over-Adherence or Lack of Individualization • Psychosocial Harm Related to Medication Burden • Systemic Harm from Measure-Driven Practices • Delayed Discontinuation in Palliative or End-of-Life Care • Differential Impact on Vulnerable Populations

Avoidable utilization

- Reduced risk of heart failure hospitalizations

Moderators

- Medication coverage policies
- Care coordination programs
- Medication management programs (pharmacy led)
- Patient education-engagement
- Referral to cardiac rehabilitation programs
- Health IT and data analytics
- Access to cardiologists, pharmacies
- Resource constraints (workforce)

BATTELLE

Material Outcome: CBE2019-3512

REMINDER: All AI generated CMO ontology claims are considered "speculative" until associated with evidence (including SME review) and argument

Potential Benefits	Potential Harms

• Decrease in out-of-pocket costs
  • Less delayed or foregone care, non-adherence to rehabilitation or follow-up treatments, financial strain impacting overall wellbeing
• Decrease in complications, readmissions, additional procedures (revisions), prolonged recovery, disability
• Decrease in fragmented or poorly coordinated care
  • Less redundant testing, avoidable ED visits
• Indirect health impacts
  • Less stress, mental health strain, reduced engagement with future healthcare needs
• Indirect financial impacts
  • Less time away from work
	• Undue pressure to reduce costs leading to decrease in necessary care
	  • More postoperative complications, suboptimal functional recovery, reoperations, chronic pain
	• Inappropriate patient selection (risk avoidance)
	  • More delayed or denied access, worse functional limitations, pain, decreased quality of life
	• Excessive substitution of lower-cost care settings
	  • More risk of falls, infections, or thrombotic events
	• Overemphasis on short-term costs over long-term value (costly revisions, chronic functional deficits)
	• Psychological burden on patients
	• Misalignment between cost and individual patient needs (housing insecurity, extended rehabilitation)

Material Outcome: CBE2019-0753

Inputs
- Infection prevention protocols
- Patient education (adherence)
- Antibiotic stewardship program
- Antiseptic agents
- Sterile equipment and supplies
- Staffing (infection prevention, nurses, pharmacists)

Mediators
- Comorbidities (diabetes, obesity, malnutrition, immunosuppression)
- Access to transportation, home health care
- Access to supplies (clean dressings)
- Local resistance patterns

	Benefits*	Harms*
Colon Surgery (high risk) (surgical complexity, disruption of gastrointestinal flora, intraoperative contamination)	• Decreased risk of prolonged wound healing/recovery, pain, functional impairment, and reduced quality of life • Decreased risk of complications (e.g., sepsis) and mortality	• Invasive diagnostic procedures (wound sampling, imaging) • Resource diversion from other critical post-operative needs
(low risk)		• Antibiotic resistance (C.Diff) • Adverse drug reactions

Avoidable utilization
- Length of stay
- Readmissions
- Cost of care (e.g., antibiotics)
- Duration of care

Moderators
- Volume (resources/capacity)
- Remote monitoring
- Infection tracking system
- Infection control teams
- Advance sterilization equipment
- EHR triggers (e.g. prophylaxis)

Mechanism: CBE2024-4440e Logic Model

Inputs (Resources-Means)	Activities (What the program does-Ways)	Outputs (Direct results of the activities)	Outcomes	Impact (Broad, systemic changes influenced by the quality program):
1. Staff and Expertise: - Radiologists, imaging technicians, and clinicians trained in pneumonia diagnosis and management. - Quality improvement (QI) teams and clinical leaders. 2. Infrastructure: - Imaging equipment (e.g., X-rays, CT scanners, portable units). - Electronic Health Record (EHR) systems with integrated clinical decision support (CDS) tools. 3. Financial Resources: - Funding for equipment, staff training, and operational support. - Grants or reimbursement incentives tied to pneumonia care measures.	1. Protocol Development: - Develop or refine protocols requiring timely chest imaging for suspected pneumonia cases. 2. Training and Education: - Conduct training sessions for clinicians on evidence-based pneumonia diagnosis, imaging indications, and antimicrobial stewardship. - Educate patients about the role of imaging in pneumonia care. 3. Technology Integration: - Deploy CDS tools in the EHR to prompt imaging orders based on clinical criteria. 4. Quality Improvement (QI) Initiatives: - Launch PDSA (Plan-Do-Study-Act) cycles to test and refine interventions aimed at increasing imaging adherence.	1. Protocols Implemented: - Standardized imaging protocols for pneumonia patients adopted across facilities. 2. Trained Staff: - Clinicians, radiologists, and QI teams trained in pneumonia care and imaging processes. 3. Functional Technology: - CDS tools operational in EHR systems, prompting timely imaging. 4. Engaged Patients: - Increased patient understanding of imaging's role in pneumonia care. 5. Data Availability: - Regular performance reports on imaging rates and adherence to protocols.	Short-term (Changes resulting from the outputs) 1. Increased adherence to imaging protocols for pneumonia diagnosis. 2. Improved clinician awareness and confidence in pneumonia care guidelines. 3. Higher rates of timely chest imaging among the Target Population. 4. Enhanced patient acceptance of imaging recommendations. Intermediate term (Effects observed as the program matures) 1. Clinical Process Improvements: - Reduced variability in imaging practices across clinicians and facilities. - More accurate pneumonia diagnoses due to consistent imaging use. 2. Resource Optimization: - Efficient use of imaging equipment and staff.	1. Diagnostic Standardization Across Facilities: - Implementation of uniform imaging protocols for pneumonia diagnosis within healthcare networks, improving consistency and reducing variability in care. 2. Integrated Care Models: - Enhanced coordination between primary care, emergency departments, radiology, and inpatient teams to streamline diagnostic workflows and improve efficiency. 3. Focus on Value-Based Care: - Shift toward reimbursement models that emphasize quality and outcomes (e.g., reduced mortality, fewer readmissions) rather than volume of services, aligning incentives with better adherence to the Measure Focus. 4. Expansion of Telehealth and Remote Diagnostics: - Increased use of tele-radiology and mobile imaging to improve access in underserved or rural areas. 5. Data-Driven Healthcare:

BATTELLE

Mechanism: CBE2019-0071 Logic Model

Inputs (Resources-Means)	Activities (What the program does-Ways)	Outputs (Direct results of the activities)	Outcomes	Impact (Broad, systemic changes influenced by the quality program):
• **Clinical Guidelines:** ACC/AHA recommendations for post-AMI care. • **Health IT Systems:** EHRs with clinical decision support (CDS), pharmacy databases, adherence monitoring tools. • **Healthcare Workforce:** Physicians, nurses, pharmacists, case managers, and care coordinators. • **Patient Resources:** Education materials, medication adherence apps, pill organizers.	• **Provider-Focused Activities:** o Training on guideline-based prescribing. o Integrating CDS tools into EHRs for beta-blocker reminders. o Regular performance feedback and audit reports. • **Patient-Focused Activities:** o Medication counseling at discharge and follow-up visits. o Providing adherence tools (pillboxes, mHealth apps). o Motivational interviewing to address barriers to adherence. • **System-Level Activities:** o Standardizing discharge protocols (e.g., Project RED, BOOST). o Care coordination between inpatient and outpatient providers.	• **Quantitative Outputs:** o Number of providers trained on beta-blocker guidelines. o Number of patients receiving discharge medication counseling. o Number of CDS alerts generated and acted upon. o Percentage of eligible patients enrolled in adherence programs. • **Qualitative Outputs:** o Improved provider knowledge and confidence in managing post-AMI medications. o Enhanced patient understanding of the importance of beta-blocker therapy.	Short-term (Changes resulting from the outputs): • **Increased Provider Awareness:** Greater adherence to guidelines for prescribing beta-blockers at discharge. • **Improved Patient Knowledge:** Patients understand the role of beta-blockers in preventing future cardiac events. • **Enhanced Medication Access:** Reduced financial barriers through insurance coverage and prescription assistance programs. • **Initial Adherence:** Higher rates of prescription fills within the first 30 days post-discharge. Intermediate term (Effects observed as the program matures) • **Sustained Medication Adherence:** Increased percentage of patients meeting the persistence	• **Transition to Value-Based Care:** o Movement from fee-for-service models to value-based payment systems (e.g., Medicare Advantage Star Ratings, ACOs) incentivizes medication adherence as a quality metric. o **Impact:** Aligns financial incentives with persistent beta-blocker use, encouraging health plans and providers to invest in adherence programs. • **Expansion of Health IT Infrastructure:** o Nationwide adoption of electronic health records (EHRs) with clinical decision support (CDS) tools, health information exchanges (HIEs), and integrated pharmacy data systems. o **Impact:** Facilitates real-time monitoring, adherence tracking, and

BATTELLE

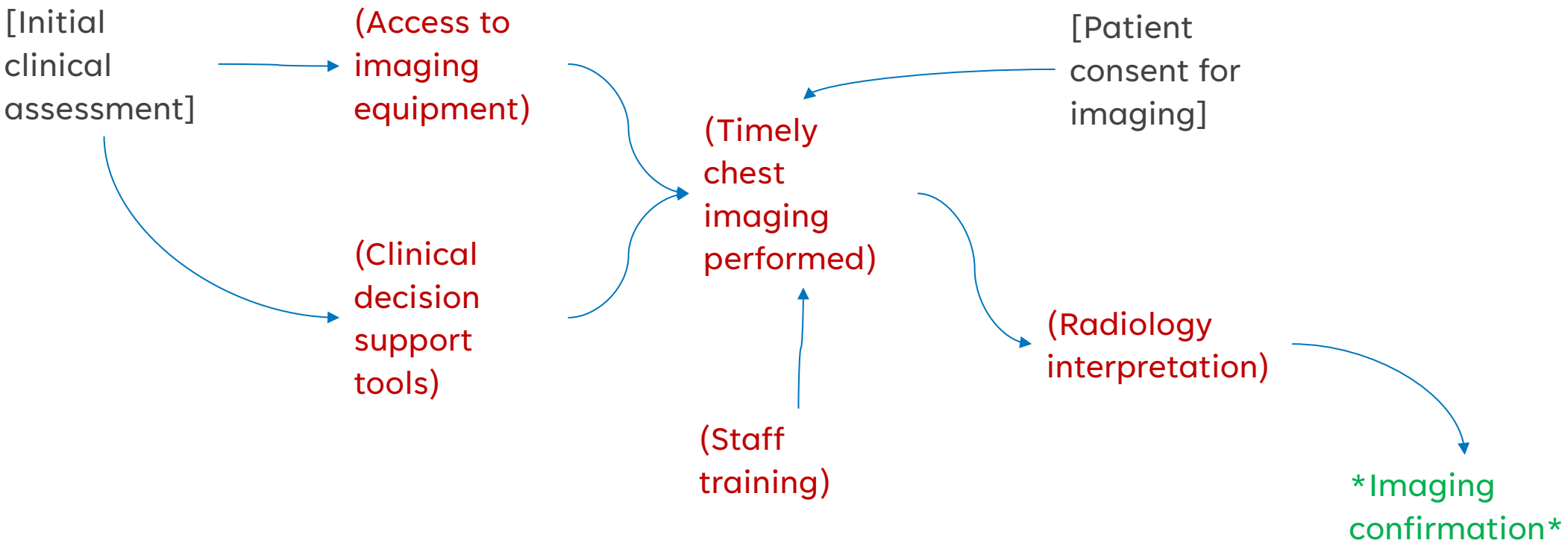
Mechanism: CBE2019-3512 Logic Model

Inputs (Resources-Means)	Activities (What the program does-Ways)	Outputs (Direct results of the activities)	Outcomes	Impact (Broad, systemic changes influenced by the quality program):
• **Clinical Guidelines & Toolkits:** AAOS guidelines, ERAS protocols, BPCI/CJR resources. • **Clinical Staff:** Surgeons, anesthesiologists, nurses, physical therapists, case managers. • **Care Coordination Staff:** Care navigators, discharge planners, social workers. • **Health IT Systems:** Integrated EHRs, predictive risk tools, data tracking systems. • **Financial Support:** Value-based payment models (e.g., Bundled Payments), organizational leadership commitment. • **Patient Education Tools:** Shared decision aids, prehabilitation materials, discharge instructions.	• **Preoperative Optimization & Risk Stratification (30 Days Before Surgery):** ○ Comprehensive patient risk assessment (e.g., BMI, diabetes control, frailty screening). ○ Prehabilitation (prehab) involving physical therapy, nutrition optimization, smoking cessation. ○ Patient education and shared decision-making to align expectations. ○ Care planning for patients with complex needs (e.g., SDOH support). • **Standardized Surgical & Anesthesia Protocols (During Surgery):**	• **Increased Preoperative Risk Assessments & Prehab Completion:** % of patients receiving pre-surgical risk screening and prehab. • **Standardized Intraoperative Practices:** % adherence to ERAS protocols and multimodal pain management. • **Timely Discharge & Recovery Plans:** % of patients discharged home with clear instructions and support. • **Patient Engagement:** % of patients reporting understanding their care plan. • **Postoperative Follow-up:** % of patients receiving timely follow-up and navigation support.	Short-term (Changes resulting from the outputs): • **Fewer Preventable Complications:** Reduced rates of surgical site infections, venous thromboembolism (VTE), and other common post-surgical complications. • **Reduced Length of Stay (LOS):** More patients safely discharged on the day of or day after surgery. • **Improved Patient Activation:** Patients more informed and confident in managing their recovery. • **Improved Care Transitions:** More patients experience seamless handoffs from surgery to home rehabilitation. Intermediate term (Effects observed as the program matures) • **Lower Readmission Rates:** Fewer patients returning to the hospital for preventable issues. • **Reduced Post-Acute Facility Use:** More patients recovering safely at home rather than	• **Normalization of Value-Based Care**: Care coordination, ERAS pathways, and episode management become standard practice across orthopedic surgery. • **Shift Toward Multidisciplinary Surgical Teams**: Surgeons, anesthesiologists, physical therapists, and care navigators routinely co-manage surgical patients, increasing team-based care culture. • **Greater Emphasis on Patient Activation**: Systems increasingly invest in tools and processes to involve patients in their preoperative preparation and postoperative recovery. • **Expansion of Preoperative Optimization as Routine**: Prehabilitation and risk stratification protocols become a default part of surgical preparation for other procedures beyond knee arthroplasty. • **Health Equity Integration**: Organizations address social

Mechanism: CBE2019-0753 Logic Model

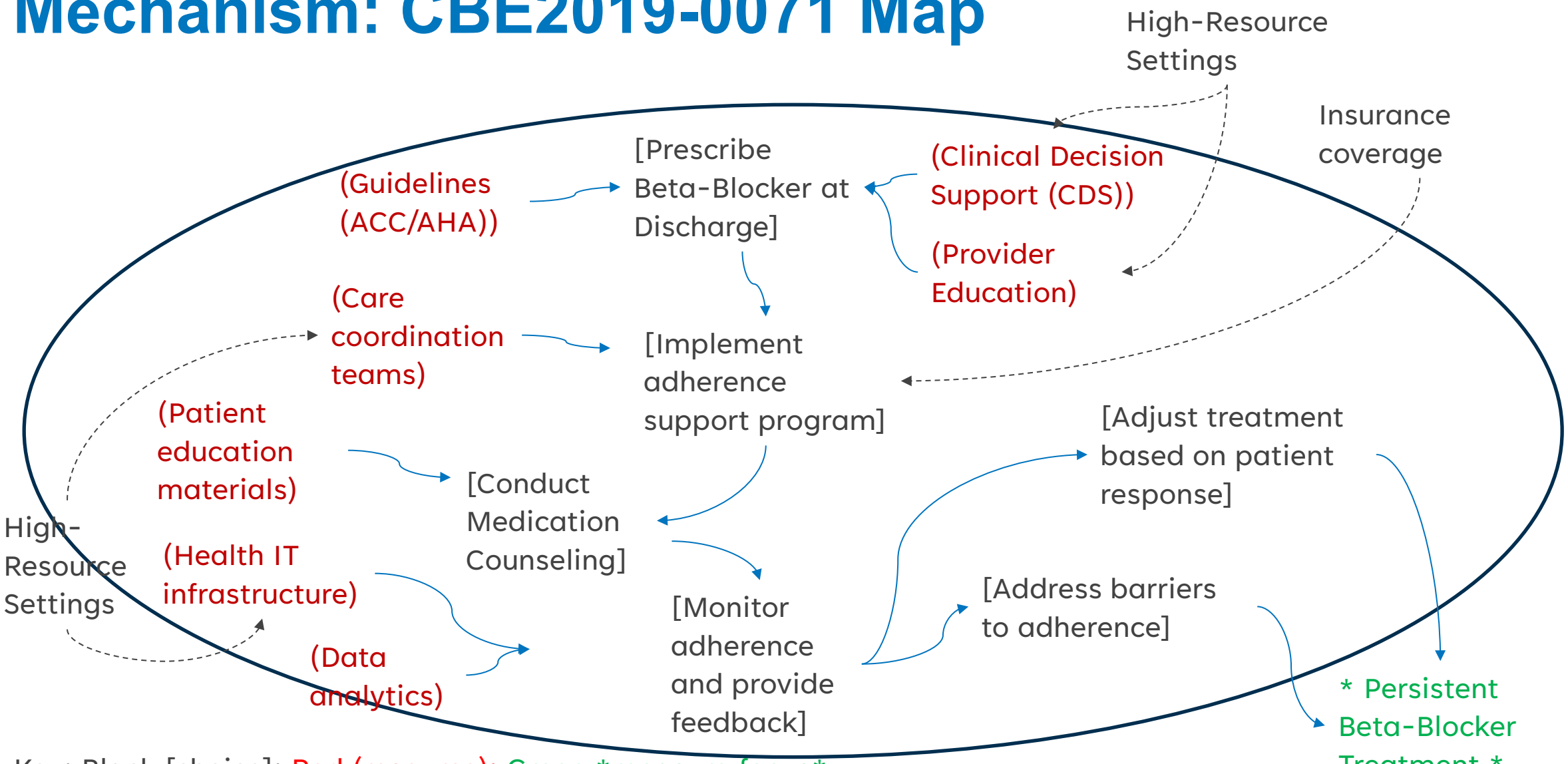
Inputs (Resources-Means)	Activities (What the program does-Ways)	Outputs (Direct results of the activities)	Outcomes	Impact (Broad, systemic changes influenced by the quality program):
• **Human Resources:** ○ Infection preventionists, surgeons, nurses, and pharmacists. ○ Educators for staff and patient training. ○ Administrative support for program oversight. • **Financial Resources:** ○ Funding for training, surveillance systems, and procurement of supplies. • **Infrastructure and Tools:** ○ Sterile surgical equipment and supplies (antiseptics, gowns, drapes). ○ Technology (e.g., electronic health records [EHRs] with alerts, infection tracking tools). ○ Access to evidence-based guidelines (CDC, WHO).	• **Training and Education:** ○ Regular training sessions on SSI prevention for healthcare staff. ○ Patient education on preoperative preparation and post-discharge wound care. • **Preoperative Interventions:** ○ Risk stratification for high-risk patients. ○ Implementation of evidence-based antisepsis and bowel preparation protocols. ○ Timely administration of prophylactic antibiotics. • **Intraoperative Practices:** ○ Adherence to sterile surgical techniques. ○ Maintaining optimal operating room conditions (e.g., airflow, sterilization).	• **Short-Term Outputs:** ○ Number of staff trained in SSI prevention protocols. ○ Number of patients educated on wound care and follow-up. ○ Implementation of standardized preoperative, intraoperative, and postoperative protocols. ○ EHR-enabled reminders and documentation for key interventions. • **Intermediate Outputs:** ○ Increased adherence to infection prevention guidelines. ○ Improved communication and collaboration among multidisciplinary teams.	<u>Short-term</u> (Changes resulting from the outputs): • Improved compliance with preoperative antibiotic prophylaxis and sterile technique. • Early detection and intervention for postoperative wound issues. • Increased staff knowledge and confidence in SSI prevention. <u>Intermediate term</u> (Effects observed as the program matures) • Reduction in the rate of SSIs identified during initial hospitalization or post-discharge. • Improved patient satisfaction and trust in surgical care.	• **Cultural Shifts:** ○ Promoting a culture of safety and accountability in healthcare settings. ○ Encouraging a mindset of continuous quality improvement among staff. • **Institutionalization of Protocols:** ○ Embedding standardized SSI prevention protocols into routine clinical workflows. ○ Establishing infection prevention as a core component of surgical care. • **Health Equity Improvements:** ○ Addressing disparities in infection prevention resources and access to care. ○ Tailoring interventions to underserved populations. • **Policy and Incentive Structures:** ○ Advocating for alignment of hospital policies with

Mechanism: CBE2024-4440e Map



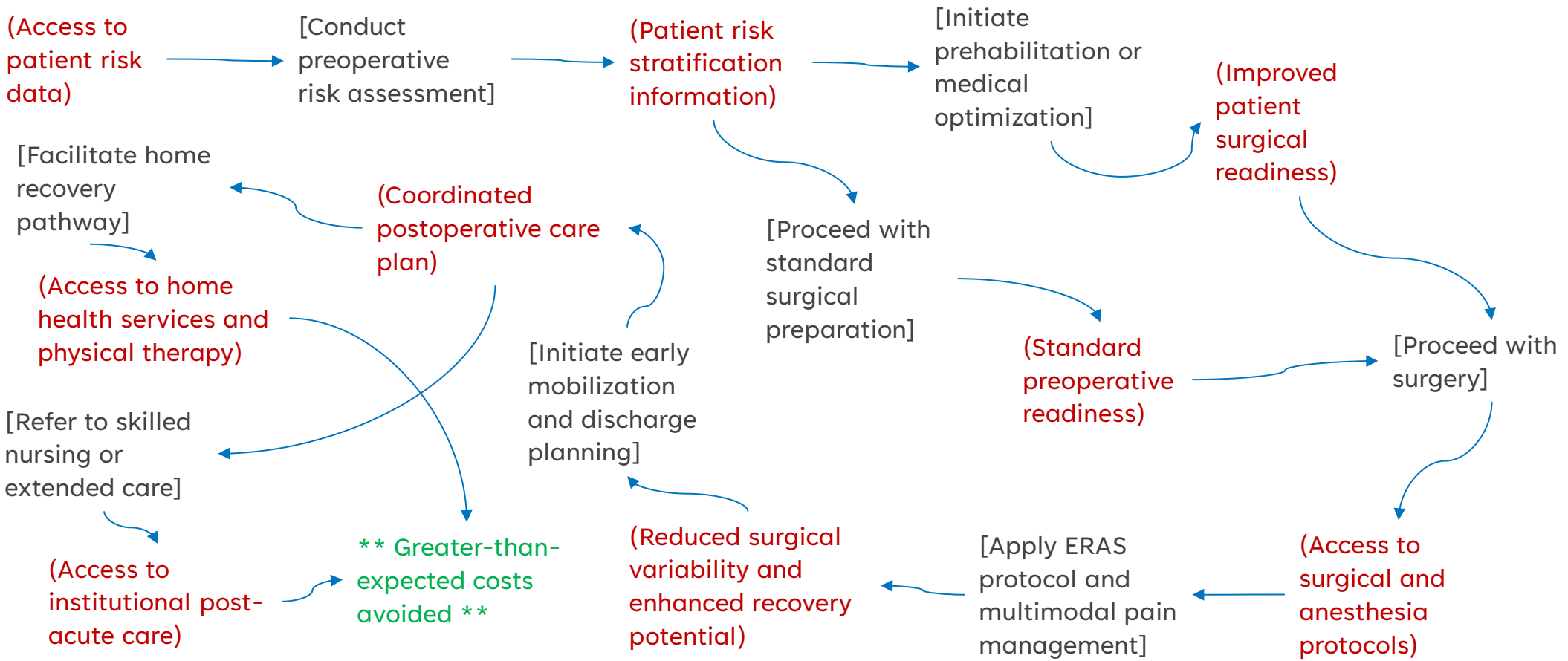
Key: Black [choice]; Red (resource); Green *measure focus*

Mechanism: CBE2019-0071 Map



Key: Black [choice]; Red (resource); Green *measure focus*

Mechanism: CBE2019-3512 Map



Key: Black [choice]; Red (resource); Green *measure focus*

BATTELLE

Mechanism Complex Complexity Assessment

Complexity Domain	Cochrane Intervention Complexity Assessment Tool for Systematic Reviews (iCAT_SR)
Number of Components	**High Complexity**: The mechanism complex consists of multiple interdependent components spanning the preoperative, perioperative, and postoperative phases. These include risk assessment, prehabilitation, ERAS protocols, multimodal pain management, discharge planning, and post-acute care coordination.
Degree of Interaction between Components	**High Complexity**: The components interact dynamically—preoperative optimization affects surgical outcomes, which in turn influences postoperative recovery and discharge choices. Failures in one component (e.g., inadequate prehab) can cascade into downstream complications and increased costs.
Number and Variability of Outcomes	**Moderate Complexity**: While the primary focus is on cost reduction (Measure Focus), the mechanism complex also affects clinical outcomes (e.g., complications, readmissions, functional recovery) and patient experience (e.g., pain management, satisfaction).
Degree of Flexibility or Tailoring Allowed	**High Complexity**: The mechanism complex requires tailoring based on patient characteristics (e.g., age, comorbidities, functional status) and local resource availability (e.g., home health services in rural areas). Clinicians adapt pathways based on individual patient needs.
Context Dependency	**High Complexity**: Effectiveness is highly context-sensitive. Urban hospitals with integrated care teams and home health support may see cost reductions, whereas rural settings with post-acute care limitations may struggle to avoid SNF stays and higher costs.
Nature of the Behaviors Required by Those Delivering the Intervention	**High Complexity**: Successful implementation requires multidisciplinary teamwork across surgeons, anesthesiologists, physical therapists, case managers, and home health providers. It demands behavior change toward standardized protocols while allowing flexibility for patient-specific needs.
Nature of the Behaviors Required by Those Receiving the Intervention	**Moderate Complexity**: Patients are expected to engage in prehabilitation, adhere to recovery plans, and manage their rehabilitation at home. Older adults or patients with social barriers (e.g., housing insecurity, low literacy) may struggle with home recovery expectations, complicating adherence.
Potential for Spillover Effects	**Moderate Complexity**: Improvements in preoperative optimization and ERAS protocols can spill over into other surgical procedures, enhancing overall perioperative care standards. However, cost containment pressures can also spill over, leading to undercare in other patient populations.

Source: https://methods.cochrane.org/sites/methods.cochrane.org/files/uploads/icat_sr_additional_file_4_2016_12_27.pdf

Context: CBE2024-4440e Realist Evaluation

Why Does the Measure Work?	For Whom Does the Measure Work?	In What Circumstances Does the Measure Work?
- The measure works by enabling or blocking choice-making related to whether the entity (facility) performs chest imaging for pneumonia patients. - It incentivizes adherence to evidence-based diagnostic protocols, prompting facilities to prioritize timely imaging as part of clinical workflows. - Choices include: - Ordering imaging based on clinical presentation. - Allocating resources to ensure imaging availability. - Incorporating imaging adherence into performance metrics. - The measure works by enabling or blocking reasoning regarding the appropriateness of chest imaging for pneumonia patients. - Reasoning processes include: - Assessing clinical presentation and determining the likelihood of pneumonia. - Evaluating risks and benefits of imaging for individual patients. - Balancing resource utilization against diagnostic accuracy.	- Not all entities have the capacity to make this choice due to resource constraints: - Well-Resourced Facilities: Large urban hospitals with advanced imaging technology and sufficient staffing are more likely to make this choice. - Resource-Limited Facilities: Rural or underfunded hospitals may lack imaging equipment or trained radiologists, limiting their ability to choose imaging as a diagnostic tool. - Capacity also depends on: - Staff training and clinical expertise. - Decision-making autonomy (e.g., influenced by institutional policies or payer constraints). - Not all entities have the resources or capacity to conduct this reasoning effectively: - Clinician Expertise: Facilities with experienced providers may perform more nuanced reasoning, leading to appropriate imaging decisions. - Technology Support: Entities equipped with clinical decision support (CDS) tools in electronic health records (EHRs) can enhance reasoning by automating guideline adherence and risk stratification. - Administrative Support: Hospitals with robust quality improvement systems may use performance data to refine reasoning processes over time. - Barriers to Reasoning: - High patient volumes or time pressures may limit the ability of clinicians to engage in careful reasoning. - Limited access to data or evidence-based guidelines can hinder informed decision-making.	- The cultural, physical, or social contexts that activate or inhibit choice-making include: - Cultural Context: A culture of defensive medicine may encourage overuse of imaging, while a culture of evidence-based practice ensures its appropriate use. - Physical Context: Geographic isolation or lack of infrastructure may prevent timely imaging, even if clinically indicated. - Social Context: Public health policies and payer systems that incentivize imaging adherence can activate the choice-making process, while lack of reimbursement or systemic inequities may inhibit it. - Reasoning is activated or inhibited by contextual factors: - Cultural Context: A culture emphasizing clinician autonomy may encourage reasoning based on experience, while strict adherence to protocols may discourage individualized reasoning. - Physical Context: Availability of imaging technology and rapid turnaround times can influence whether reasoning leads to imaging orders. - Social Context: Social norms or patient expectations may shape reasoning. For example: - Patients demanding imaging may pressure clinicians to comply. - Disparities in access to care or bias may result in differential reasoning for patients from underserved populations.

Context: CBE2019-0071 Realist Evaluation

Why Does the Measure Work?	For Whom Does the Measure Work?	In What Circumstances Does the Measure Work?
• **How the Measure Works (Why):** The measure works by **enabling or blocking choice-making** by health plans, healthcare systems, and providers. This includes decisions related to: ○ **Prescribing behavior:** Whether to initiate beta-blocker therapy at discharge. ○ **Adherence support:** Whether to implement programs that encourage patient persistence with medication. ○ **Resource allocation:** Whether to invest in health IT, care coordination, and pharmacy services to improve adherence rates. • **Key Dynamics:** ○ **Enablers:** Clinical guidelines (e.g., ACC/AHA recommendations), performance incentives (e.g., HEDIS measures), and decision support systems. ○ **Blockers:** Lack of provider awareness, competing clinical priorities, and absence of financial incentives for adherence-focused interventions. • **How the Measure Works (Why):** The measure also works by influencing the **reasoning processes** of entities (health plans, providers): ○ **Clinical reasoning:** Evaluating the risk-benefit profile of beta-blocker therapy for each patient. ○ **Organizational reasoning:** Assessing the importance of medication adherence in quality improvement strategies. ○ **Policy reasoning:** Determining how performance on this measure impacts reimbursement, accreditation, or public reporting.	• **Large Integrated Health Systems:** ○ **Capacity:** Robust health IT infrastructure, dedicated QI teams, and strong care coordination models. ○ **Resources:** Access to comprehensive patient data, pharmacy integration, and financial resources to support adherence programs. • **Health Plans with Value-Based Care Models:** ○ **Capacity:** Sophisticated data analytics capabilities to monitor adherence and implement targeted interventions. ○ **Resources:** Incentive structures tied to quality measures (e.g., Medicare Advantage Star Ratings) that promote persistent medication use. • **Clinicians in Academic Medical Centers:** ○ **Capacity:** High levels of guideline awareness, access to continuing education, and multidisciplinary care teams. ○ **Resources:** Embedded clinical decision support tools and resources for medication counseling. • **Small or Resource-Limited Community Hospitals:** ○ **Capacity Constraints:** Limited health IT capabilities, fewer staff dedicated to QI, and fragmented care transitions. ○ **Resource Gaps:** Lack of integrated pharmacy services and insufficient funding for adherence interventions.	• **Supportive Policy Environments:** ○ **Example:** Health systems operating under value-based payment models that reward medication adherence. ○ **Effect:** Strong alignment between financial incentives and adherence-improvement efforts. • **High-Functioning Care Transitions:** ○ **Example:** Hospitals with robust discharge planning, medication reconciliation, and follow-up care processes. ○ **Effect:** Increased likelihood that patients will receive, understand, and persist with beta-blocker therapy. • **Patient-Centered Cultures:** ○ **Example:** Organizations that prioritize shared decision-making, patient engagement, and culturally competent care. ○ **Effect:** Enhanced patient understanding, motivation, and adherence to prescribed therapies. • **Fragmented Healthcare Systems:** ○ **Example:** Lack of coordination between inpatient and outpatient providers, leading to gaps in medication management. ○ **Effect:** Increased risk of non-adherence due to poor follow-up and lack of continuity in care.

Context: CBE2019-3512 Realist Evaluation

Why Does the Measure Work?	For Whom Does the Measure Work?	In What Circumstances Does the Measure Work?
Why: The Measure Works by Enabling or Blocking Choice-Making The measure influences clinician behavior by making episode costs visible and holding clinicians accountable for the financial efficiency of knee arthroplasty care. • It enables choice-making by highlighting cost drivers across the preoperative, perioperative, and post-acute care phases. • It blocks certain choices by incentivizing the reduction of unnecessary services and complications, discouraging overuse of post-acute facilities or prolonged inpatient stays. **Why: The Measure Works by Enabling or Blocking Reasoning** The measure promotes reasoning by prompting clinicians to evaluate their practice patterns and consider how preoperative preparation, surgical technique, and discharge pathways affect both cost and quality. • Enables reasoning by providing comparative cost performance data and stimulating reflection on variation across patients and providers. • Blocks reasoning when the cost signal is ambiguous or overly punitive, leading to defensive behavior (e.g., avoiding high-risk patients).	**For Whom: Capacity and Resources for Choice-Making** • Works well for: ○ Large, integrated health systems with care coordination capacity, data analytics infrastructure, and standardized surgical pathways. ○ Clinicians with access to home health, outpatient physical therapy, and prehabilitation programs. ○ Providers experienced with bundled payment models, who have support from administrative teams to optimize episode costs. • Works less well for: ○ Small practices and rural hospitals with limited administrative capacity, fewer care coordination resources, and gaps in home health or outpatient rehab. ○ Clinicians in fragmented systems where pre- and post-acute care decisions are made by different, unaligned entities. ○ Surgeons without data feedback or comparative benchmarks may lack the information necessary to make cost-conscious choices. **For Whom: Capacity and Resources for Reasoning** • Works well for:	**In What Circumstances: Context That Activates or Blocks Choice-Making** • Activating Contexts: ○ Bundled payment models (e.g., BPCI, CJR) align financial incentives with the measure, amplifying the choice-making mechanism. ○ Health systems with EHR-integrated pathways and care coordinators facilitating prehabilitation and discharge planning. ○ Professional culture emphasizing value-based care—clinicians view cost-efficiency as part of delivering high-quality care. • Blocking Contexts: ○ Fee-for-service environments encourage volume-driven care, weakening the choice-making mechanism. ○ Social determinants of health (SDOH) challenges (e.g., transportation barriers, housing instability) limit viable discharge choices, narrowing clinicians' options. ○ Regions with limited home health services or post-acute care deserts constrain discharge planning, reducing the effectiveness of cost-conscious decision-making. **In What Circumstances: Context That Activates or Blocks Reasoning** • Activating Contexts: ○ Benchmarking against peers contextualizes cost data, enabling interpretation beyond individual patient variability.

BATTELLE

Context: CBE2019-0753 Realist Evaluation

Why Does the Measure Work?	For Whom Does the Measure Work?	In What Circumstances Does the Measure Work?
• The measure works by **enabling or blocking choice-making** through its emphasis on: ○ **Incentives:** External motivators, such as public reporting and financial penalties, push entities to prioritize SSI reduction. ○ **Defined Options:** The measure outlines specific intervention strategies (e.g., preoperative prophylaxis, sterile techniques), guiding entities in their infection prevention efforts. • The measure works by **enabling or blocking reasoning** through: ○ **Access to Evidence:** Guidelines and best practices provide entities with the knowledge to reason through SSI prevention strategies. ○ **Data Feedback:** Regular surveillance and reporting (e.g., through NHSN) allow entities to analyze trends and identify gaps in performance. ○ **Problem-Solving Frameworks:** Tools like root cause analysis encourage entities to reason about failures and develop corrective actions.	• The measure works for entities that have: ○ **Organizational Capacity:** Adequate staffing levels, including infection preventionists, trained clinicians, and support staff. ○ **Financial Resources:** Funding for training, equipment (e.g., advanced sterilization tools), and technology (e.g., electronic health records). ○ **Leadership Support:** Engagement from hospital leadership to allocate resources and enforce compliance. • **For whom the measure does not work:** ○ Entities lacking financial or human resources may struggle to implement the required practices. ○ Smaller facilities with limited infrastructure, such as rural or community hospitals, may face barriers in adopting the choice-making framework. • The measure works for entities with: ○ **Analytical Capacity:** Access to trained infection preventionists or quality improvement experts who can interpret data. ○ **Data Infrastructure:** Advanced EHR systems or surveillance tools to collect, analyze, and report SSI data. ○ **Commitment to Quality Improvement:** Organizations that embrace continuous learning and adaptive practices. • **For whom the measure does not work:**	• **Cultural Context:** ○ A culture of safety and accountability encourages proactive decision-making. ○ Conversely, a punitive or blame-oriented culture may discourage honest reporting and learning from SSIs. • **Physical Context:** ○ Entities in well-resourced urban centers may find it easier to implement the necessary interventions. ○ Geographic isolation or resource limitations in rural settings can inhibit successful choice-making. • **Social Context:** ○ Collaborative environments, where multidisciplinary teams work together, enhance choice-making. ○ Fragmented systems, with poor communication between surgeons, nurses, and infection control staff, may block effective choices. • **Cultural Context:** ○ An open, learning-oriented culture fosters critical reasoning and adaptation. ○ A rigid or hierarchical culture may limit staff input and creative problem-solving. • **Physical Context:** ○ Entities with centralized data systems and access to real-time feedback are better positioned to reason effectively.

BATTELLE

Building a Validation Roadmap

Validity Claim (A: entity response to measure; B: measure focus)	Association studies (A clinical study for the claim that A is a cause of B repeatedly measures the values of a set of measured variables that includes the variables A and B)	Mechanism studies (A mechanistic study for the claim that A is a cause of B is a study which provides evidence of structure or features of the mechanism (M1) by which A is hypothesized to cause B (M2))
Causal claim (A is a cause of B)	(Prone to bias)	(Prone to complexity)
Correlation claim (A and B are probabilistically dependent conditional on potential confounders C)	LEVEL I Importance Table (Performance score by decile): • Observed association between entity (A) and GTE costs (B) Known Groups: • Observed association between group (A) and GTE costs (B) Related Process or Outcome: • Fewer complications, unnecessary services, optimized recovery, patient activation	LEVEL 3 (arguable true) Effectiveness of the mechanism complex • Association between elements of the mechanism complex and the measure focus Reinforcing Contextual Mechanisms: • Integrated health systems and care continuity Counteracting Contextual Mechanisms: • Fragmented Care Systems and Weak Care Transitions • Resource Constraints in Smaller or Rural Settings
General mechanistic claim (there is a complex of mechanisms that invokes A as partially responsible for B and that can account for the extent of the correlation)	LEVEL 2 (arguable false) Reliability Table (reliability by volume decile): • Proportion of entity-level variation explained by uncertainty and chance Confounders C associated with both A and B: • Proportion of entity-level variation explained by risk-adjustment (comorbidities, frailty, physical) • Social Determinants of Health (SDOH) Challenges (low income, rural)	LEVEL 4 Experimental or quasi-experimental studies establish that implementing the mechanism complex reduces greater-than-expected costs. LEVEL 5 The explicit mechanism complex is widely recognized as a best practice.

Building a Validation Roadmap

Validity Claim	Association studies (A clinical study for the claim that A is a cause of B repeatedly measures the values of a set of measured variables that includes the variables A and B)	Mechanism studies (A mechanistic study for the claim that A is a cause of B is a study which provides evidence of structure or features of the mechanism (M1) by which A is hypothesized to cause B (M2))
Causal claim (A is a cause of B)	(Prone to bias)	(Prone to complexity)
Specific mechanism hypothesis (posit features of such a mechanism complex)		Adoption/implementation (fidelity) of the mechanism complex: • Preoperative Optimization and Risk Stratification • Standardized Surgical and Anesthesia Protocols • Proactive Postoperative Care and Discharge Planning • Longitudinal Care Coordination and Case Management

AI Pilot Study – Building Trust

1. **Best practice design**: CMO ontology, context, persona constrains the input (prompt) and output (response) space for LLM to be contextually appropriate
2. **Assurance cases**: Each claim is assumed unsubstantiated and must be supported by evidence and argument (ground truth)
3. **Expert review**: Arguments must be plausible to subject matter experts (SMEs)
4. **Harms**: CMO ontology intentionally seeks out harms, disadvantaged entities and populations
5. **Transparency**: Prompts and responses (justifications) are transparent and subject to SME, staff, and committee review
6. **Monitoring**. Track key performance indicators (KPIs) for continuous improvement
7. **Cost**: Much less time and resource intensive (hours/days not weeks/months)

AI Pilot Study – System Safety

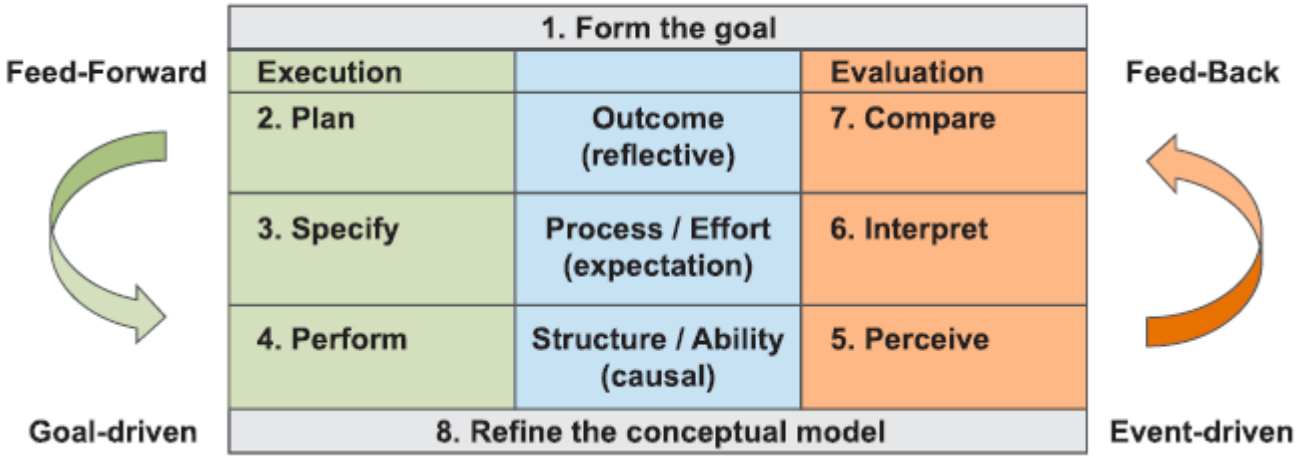
Comparison of the System-Theoretic Process Analysis (STPA) and Context-Mechanism-Outcome (CMO)

Phase	STPA	CMO
Define the purpose of the analysis	1) Identify loses 2) Identify system-level hazards 3) Identify system-level constraints 4) Refine hazards	1) A material outcome 2) Context-mechanism (worse) 3) Context-mechanism (better)
Model the control structure	1) Controller a) control algorithm b) process model 2) Controlled process	Quality Improvement Action Model 1) Agent: goal, authority, accountability a) feed-forward (goal-driven) b) feed-back (even-driven) 2) Action: perform-perceive
Identify unsafe control actions	1) Unsafe control action a) Not providing the control action leads to a hazard b) Providing the control action leads to a hazard c) Providing a potentially safe control action but too early, too late, or in the wrong order d) Control action lasts too long or is stopped too soon (for continuous control actions, not discrete ones) (source, type, action, context, hazard)	1) Context a) individual b) interpersonal c) institutional d) infrastructural (social, etc.) 2) CMO configuration (triggers)
Identify loss scenarios	1) Unsafe controller behavior 2) Inadequate control algorithm 3) Unsafe controller input 4) Inadequate process model	1) Wrong goal 2) Wrong execution/FF 3) Wrong perform-perceive 4) Wrong evaluation/FB

Source: MIT Partnership for Systems Approaches to Safety and Security (PSASS)

AI Pilot Study – Action Model

Figure 1. The Quality Improvement Action Model



Goal formation:

1. **Form the goal**. What does the user want to accomplish? In quality improvement, the goal is generally to improve outcomes and / or reduce costs.

Execution:

2. **Plan the action**. What are the alternative action sequences? Each action sequence or pathway should result in achievement of the outcome.
3. **Specify the action sequence**. What action can I do now? Given the capabilities, resources, authority, accountability or other constraints of the user, which of the projected action sequences is possible to perform?
4. **Perform the action(s)**. How do I perform the action? What physical, human, information, or knowledge structure or standard operating procedures should be used?

Evaluation:

5. **Perceive the actual and expected or intended result**. What happened? May the user perceive whether the prescribed structure or standard operating procedure generated the expected or intended result?
6. **Interpret the actual and expected or intended result**. What does it mean? If the actual result is different than the expected or intended result, why (5 whys)? What was the experience, expertise, or context that generated the actual result?
7. **Compare the actual result with the goal**. Has the user accomplished the goal? Regardless of how the result was achieved, is the result consistent with the goal?

Maturation:

8. **Revise the conceptual model**. What are the systematic or persistent factors that explain any difference between the expected and actual result, and may these factors be incorporated into the conceptual model?

Source: Norman, Don. The Design of Everyday Things: Revised and Expanded Edition. Basic Books, 2013.

DHS Generative AI Public Sector Playbook

Table 1. CBE AI Pilot Deployment Steps

Deployment Step	Description	Current Status
Mission-Enhancing GenAI Use Case	• Public sector organizations must ensure that GenAI deployments align with their mission • Narrowly scoped, mission-enhancing pilots are useful tools for exploring how an organization can use GenAI	
Coalition Building and Effective Governance	• Organizations should cultivate support for GenAI applications from top leadership and across functional teams to give GenAI the greatest chance for successful deployment and effective oversight	
Tools and Infrastructure	• Organizations should evaluate the technical tools and infrastructure they already possess and consider what technical capabilities they require to deploy GenAI applications	
Responsible Use and Trustworthiness Considerations	• From the very beginning, organizations should consider how to make sure GenAI use is responsible and trustworthy and how to address potential risks like privacy, security, bias, and safety	
Measurement and Monitoring	• Teams that are developing GenAI applications should measure progress with appropriate metrics and report on that progress to leadership and other stakeholders	
Training and Talent Acquisition	• Organizations should train their staff on responsible and effective GenAI use and hire skilled employees who can support GenAI development	
Usability Testing and Other Feedback Mechanisms	• Organizations should incorporate iterative feedback from users and other stakeholders to develop and improve GenAI applications	

Source: DHS Generative AI Public Sector Playbook | Homeland Security

DHS Generative AI Public Sector Playbook

Table 1. Key Performance Indicators for AI Pilot

Domain	Metric	
Percentage of claims by source	Number	Percentage
• Provided claims only		
• CMO claims only		
• Both Provided and CMO claims		
• **Total**	Number	Percentage
Status of CMO claims (SME review)		
• Arguable true		
• Speculative or Arguably false		
• **Total**	Number	Percentage
Claim Status Justifications (SME review)		
• Agree		
• Neutral		
• Disagree		
• **Total**		

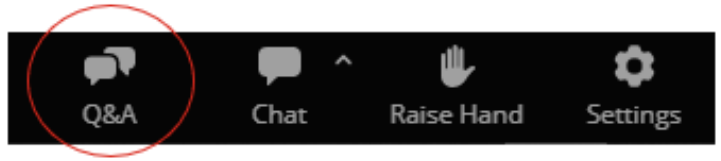
Measure Evaluation – AI Pilot Use Cases

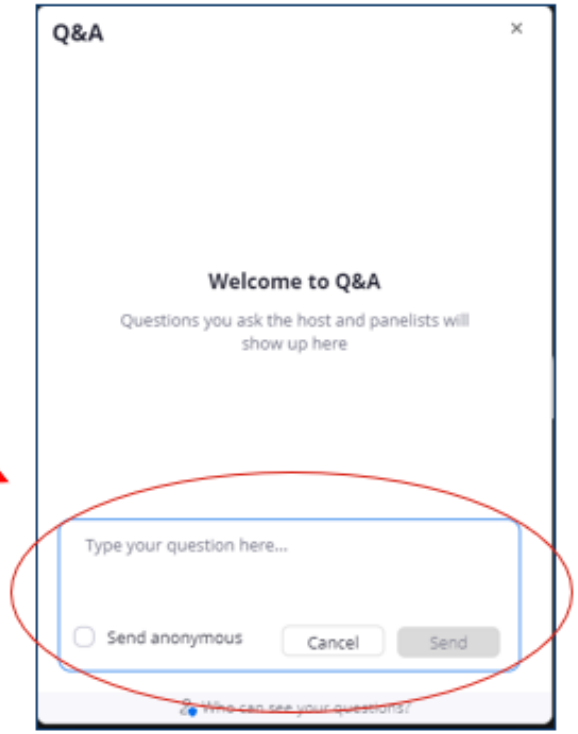
- Use Cases
  - Measure Lifecycle
    - Accelerate conceptualization, specification, testing, implementation, use
  - Measure Evaluation
    - Support efficiency and effectiveness of CBE staff and committee reviews
  - Technical Assistance
    - Reduce burden on measure developers/QCDRs/community developers
  - Endorsement Pathways
    - Enable alternative pathways for commercial plans, state agencies, communities

Questions?

Open the Zoom Q&A function

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





BATTELLE

References

- https://nam.edu/programs/value-science-driven-health-care/assessing-meaningful-community-engagement (CBE Strategy)
- Pawson, R., Tilley, N. (1997). Realistic evaluation. United Kingdom: SAGE Publications (context-mechanism-outcome ontology)
- Thissen D, Wainer H. Test scoring. Lawrence Erlbaum Associates, 2001; Messick, S. (1989b). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New "York: Macmillan; Standards for Educational and Psychological Testing. (2014). United States: American Educational Research Association (modern validity theory)
- National Institute of Standards and Technology (2022) Engineering Trustworthy Secure Systems (NIST SP 800-160 Vol. 1 Rev. 1) https://csrc.nist.gov/pubs/sp/800/160/v1/r1/final; International Organization for Standardization. (2019). Systems and software engineering — Systems and software assurance – Part 1: Concepts and vocabulary (ISO Standard No. 15026-1:2019). https://www.iso.org/standard/73567.html (assurance cases)

References

- Parkkinen, VP, et. al. Evaluating Evidence of Mechanisms in Medicine: Principles and Procedures. Springer International Publishing, 2018; Shan, Y., Williamson, J. (2023). Evidential Pluralism in the Social Sciences. United States: Taylor & Francis

- Parkkinen, VP, et. al. Evaluating Evidence of Mechanisms in Medicine: Principles and Procedures. Springer International Publishing, 2018; (Claim Status, Confidence level of evidence)

- Atkins, D., Best, D., Briss, P. A., Eccles, M., Falck-Ytter, Y., Flottorp, S., et. al. GRADE Working Group (2004). Grading quality of evidence and strength of recommendations. BMJ (Clinical research ed.), 328(7454), 1490. https://doi.org/10.1136/bmj.328.7454.1490 (Quality level of evidence)

BATTELLE

It can be done

800.201.2011 | solutions@battelle.org | www.battelle.org