

# AI AND GOVERNANCE: THE JOHNS HOPKINS EXPERIENCE

Kimberly Peairs, MD, FACP  
Vice Chair Clinical Affairs, Dept. Medicine  
Co-chair, AI Oversight Council, JHM

Dr. Kimberly Peairs is the Vice Chair for Clinical Affairs for the Johns Hopkins Department of Medicine. She is also the Medical Director for Primary Care Value and Innovation in the Office of Johns Hopkins Physicians. She has an active primary care clinical practice for more than 25 years.

Dr. Peairs graduated from the Johns Hopkins School of Medicine and completed her residency in internal medicine on the Johns Hopkins Medicine Osler Service. After serving as a Chief Resident, she joined the full time Johns Hopkins School of Medicine faculty in General Internal Medicine.

She has been involved in numerous operational and quality improvement initiatives in the Department of Medicine and Johns Hopkins Medicine.

Over the last two years she has been involved in the development and piloting of several Generative AI initiatives for clinical care delivery at Johns Hopkins. As part of the operational AI leadership team, comprised of IT and Digital health leaders, she has helped structure the Johns Hopkins AI governance roadmap.

# What problems are we solving for? What are risks of AI?

- Example Goals:
  - Improve:
    - Efficiency
    - Accuracy
    - Provider/Patient Interactions
  - Decrease:
    - Provider Burnout
    - Excessive documentation burden
- Assessing Risk:
  - What “pre-AI” present state of work flow (accuracy, safety, reproducibility, etc.)?
  - What are unintended consequences?

# Case Example

Categorization of MyChart "Inbasket" messages sent to providers/clinic

- Reason to focus on this?
- Development of Tool
- Pilot phase
- Assessment/Heuristic evaluation with Quality/Safety team

# JHM Clinical AI In-Flight



## Imaging – Mammography (Screen Point)

Provides scores (elevated, intermediate, or low) and markings on the images to assist radiologist in early breast cancer detection. **Live since February 2024 & well-received by breast radiologists.**



## Virtual AI Scribes (Abridge)

Clinical documentation solution combining conversational, ambient AI with advanced generative AI. **System wide go lives over last three months: Ambulatory (12/10), ER (1/14), Inpatient (1/28)**



## Automatically Generates Draft Response (Epic In Basket)

Drafts response to patient message based on information from the patient's record, such as current prescriptions and recent results, which clinicians review before sending. **Current pilots aimed at support staff – low uptake. Epic will have significant updates in the next year (tone/style).**



## Categorize In Basket Messages (Epic In Basket)

Leverage NLP and Generative AI to apply categories to messages to support message management and workflow. **Custom built tool initially deployed in collaboration with Microsoft, now piloting an Epic-built tool with wider capabilities.**



## Healthcare General Reasoner (VTE Decision Support)

Co-developing exploratory models to tackle complex medical questions, such as Venous Thromboembolism (VTE). **Minimal Viable Product (MVP) will go up with a small pilot in April.**

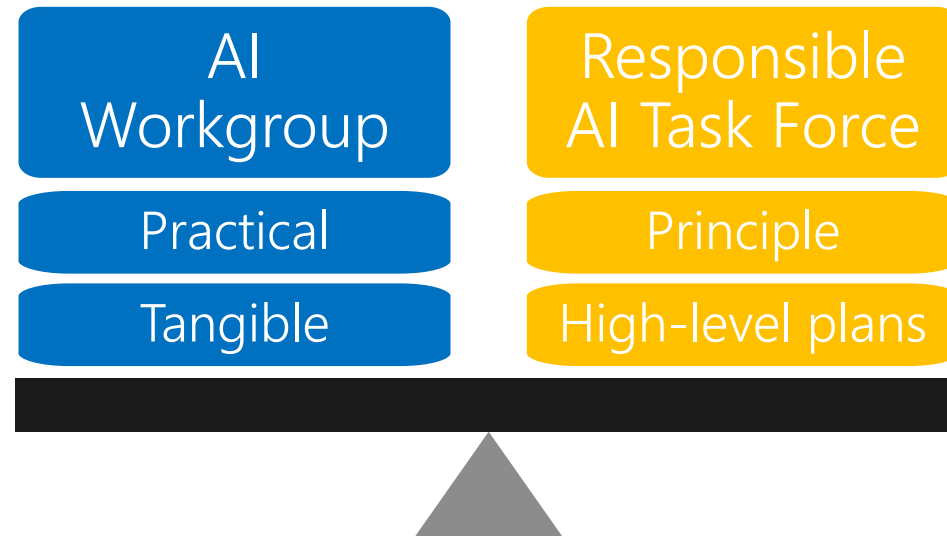


## Chart Summarization (Epic Inpatient)

Early-stage development to summarize specific areas of the medical record, starting with hospital course notes for admitted patients. **Tried & failed. Re-trialing new Epic version.**

**We need to move fast.**

**We need to move carefully.**



# Consortiums, Coalitions, & Core Principles



## **CHAI**

(Coalition For Health AI)

Developing guidelines and guardrails to drive high-quality health care by promoting the adoption of credible, fair, and transparent health AI systems



## **TRAIN**

(Trustworthy & Responsible AI Network)

Aims to operationalize responsible AI principles to improve the quality, safety and trustworthiness of AI in healthcare



## **VALID AI**

(Vision, Alignment, Learning, Implementation, and Dissemination of Validated Generative AI)

Collaborative that will explore uses, pitfalls, and best practices for Gen AI in health care and research, and accelerate execution and real-world evidence

## **AAMC**

(Association of American Medical Colleges)

**Digital Health & AI Learning Collaborative**

## **NAM**

(National Academy of Medicine)

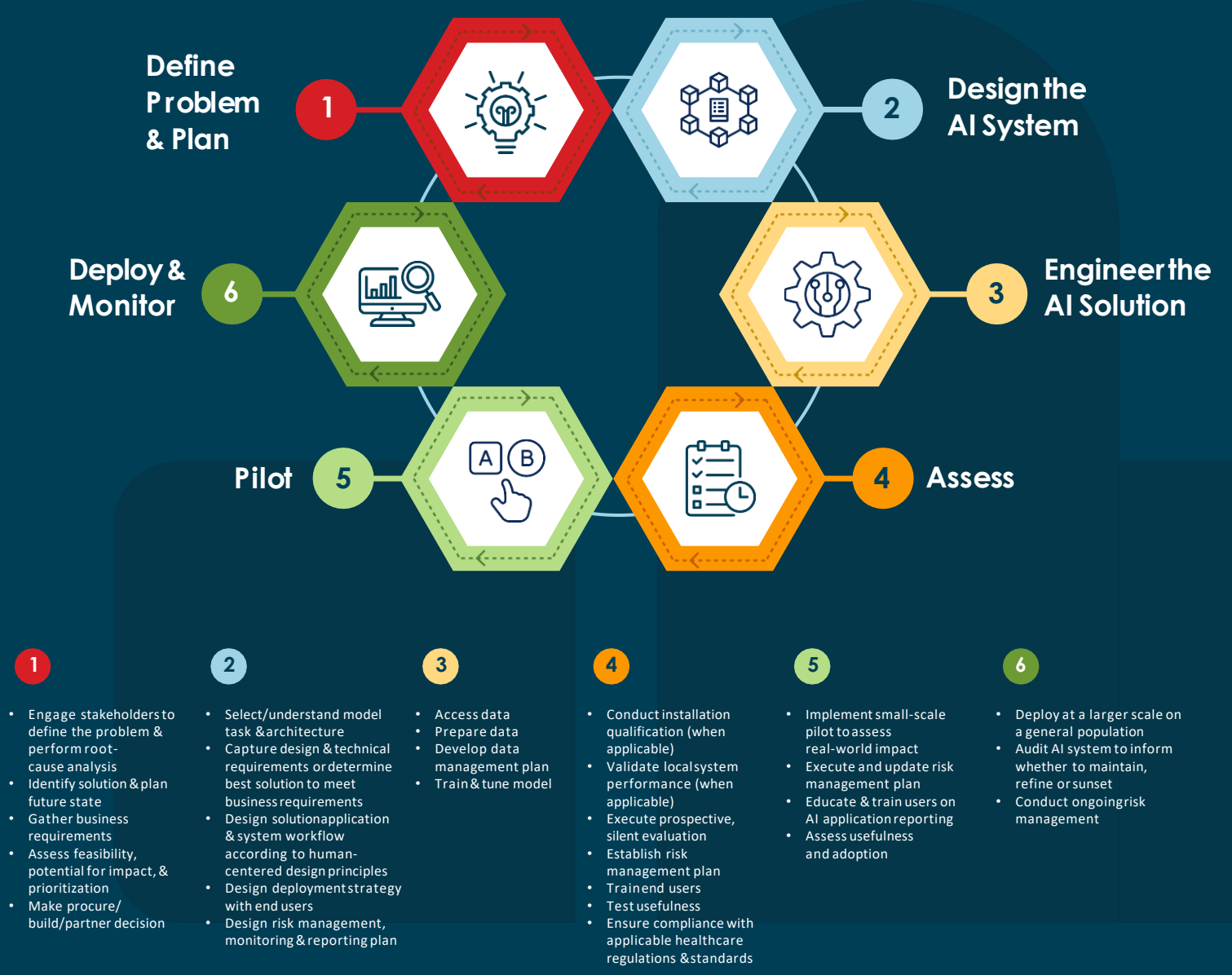
**Healthcare AI Code of Conduct**

**CHIME** (College of Healthcare Information Management Executives)  
**Central AI Principles**

# THE AI LIFECYCLE

The AI lifecycle is central to understanding and implementing CHAI's Assurance Standards in healthcare. The six-step lifecycle outlines the essential stages and processes involved in developing, deploying, and maintaining AI systems.

By systematically addressing each phase of the lifecycle, the framework ensures that AI systems adhere to the highest standards of safety, efficacy, fairness, transparency, and security. This structured approach supports risk mitigation, managing biases, and promotes accountability and trustworthiness in AI applications.



# JHM

## Clinically-led Responsible AI Program



## Multidisciplinary Expertise

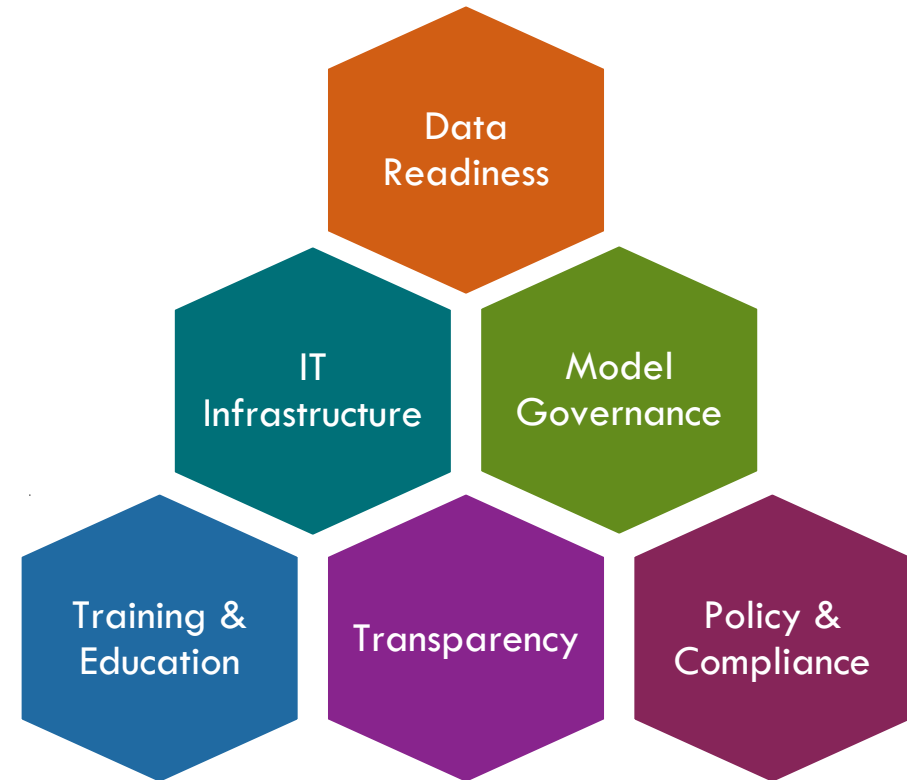
- Clinical
- Informatics
- Bioethics
- Health IT
- Safety/Quality
- Finance
- Government Affairs
- Education
- Engineering

# Clinically-led Responsible AI Program

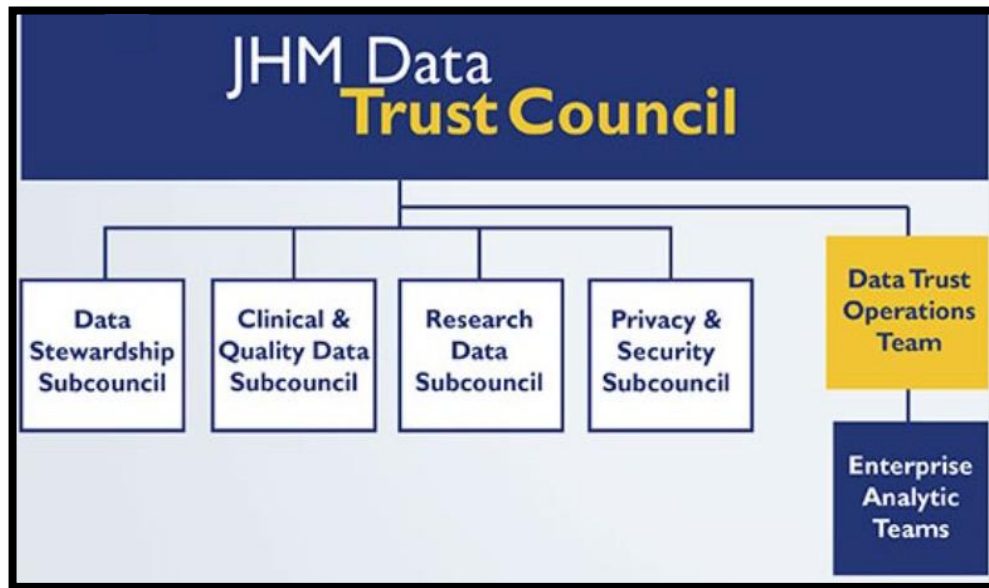
## 7 Core Principles For Use



## Implementation Plan



# Existing Governance Structures



The Data Trust governs the infrastructure, standards, policies and procedures for patient and member-related data.

## The Software Intake Process (SIP): Requesting a New Software-Based Solution

[Submit a Request](#)

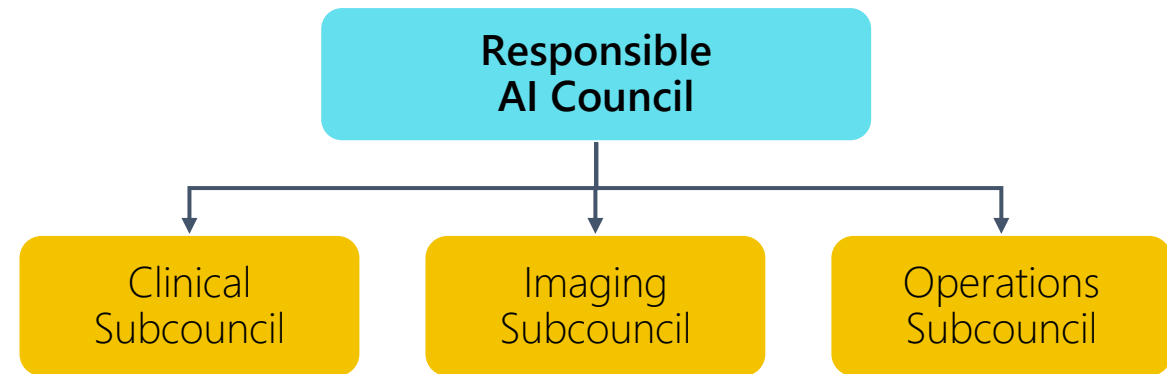
Are you looking to purchase or implement a software-based solution for a Johns Hopkins Medicine entity or department – including The Johns Hopkins Health System Corporation, affiliated hospitals, joint ventures, and the Johns Hopkins University School of Medicine? Will your software be implemented enterprise-wide? If so, review the types of requests that should go & should not go through SIP below. If you have questions about whether or not the technology you would like to bring onboard is appropriate for this process, please contact [sip@lists.johnshopkins.edu](mailto:sip@lists.johnshopkins.edu).

The Software Intake Process is a streamlined, single-entry point to collect and govern new technology requests.

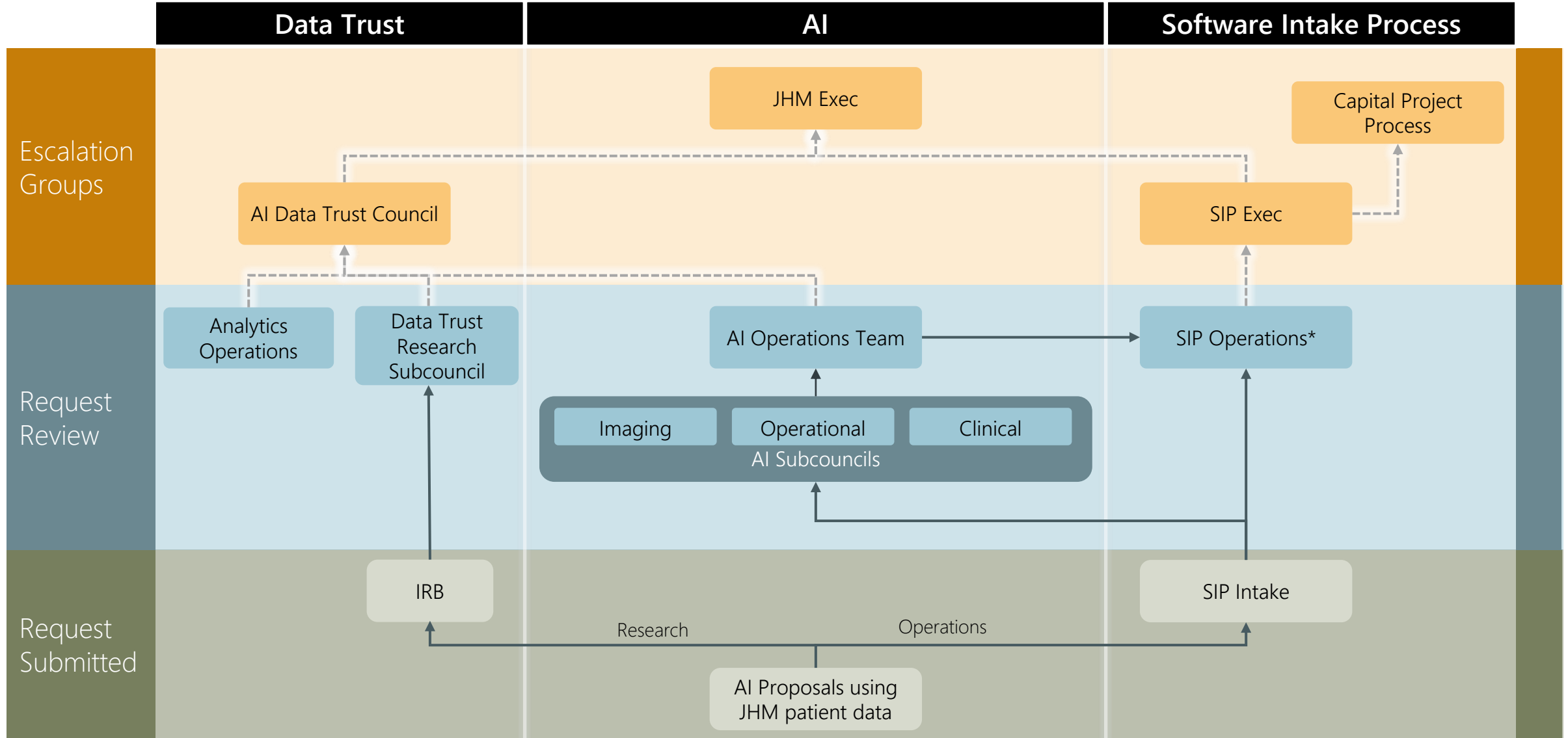
# Recommendations: Governance Structure

## JHM Responsible AI Council

- internal & external experts
- provide oversight for AI tool use, ongoing evaluation, and cessation
- provide input on AI model development, implementation, risk mitigation, and strategy

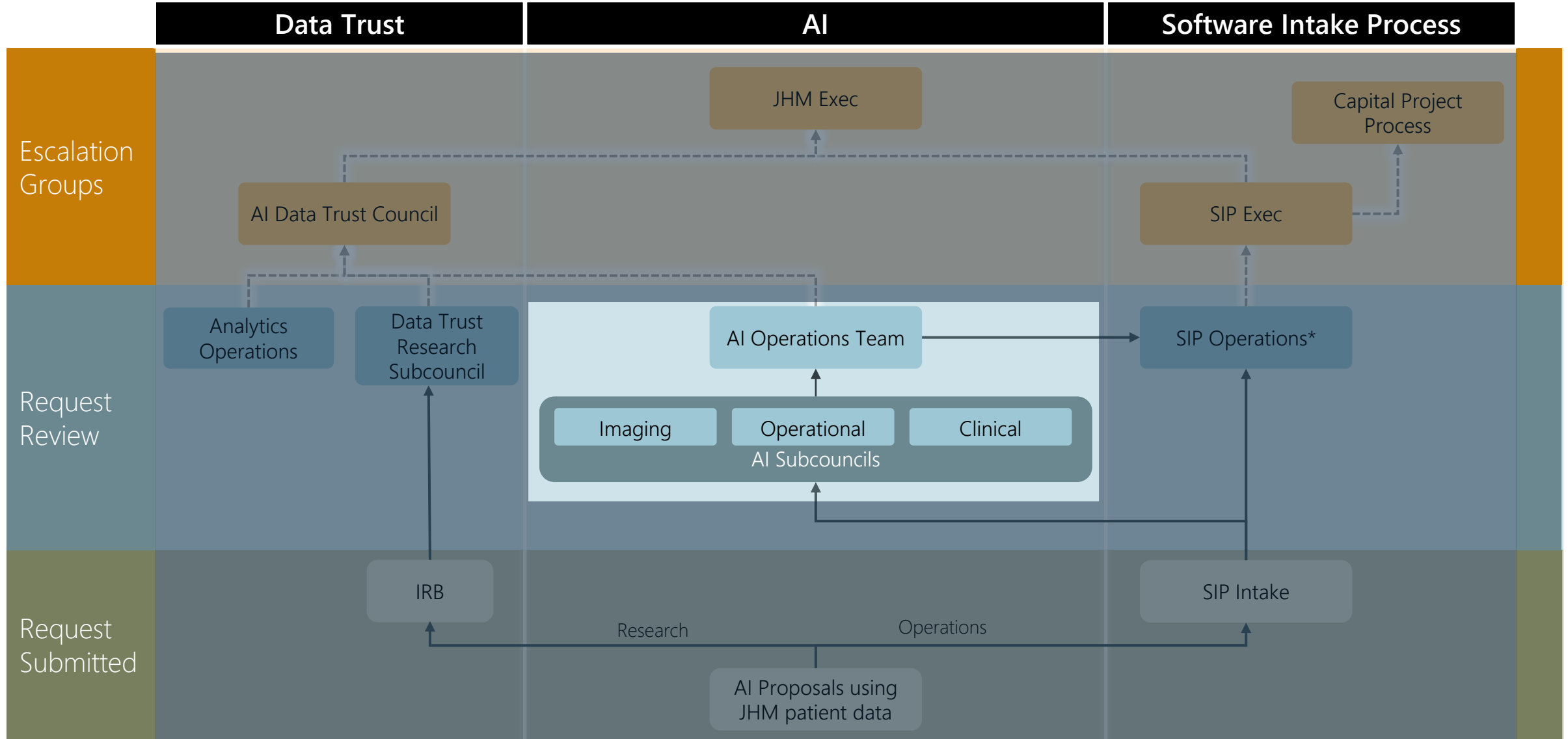


# Integrated Governance Structure



SIP should be initiated when there is software integration with a clinical system, when the software is intended for enterprise-wide use, or upon request of the study team to review the software

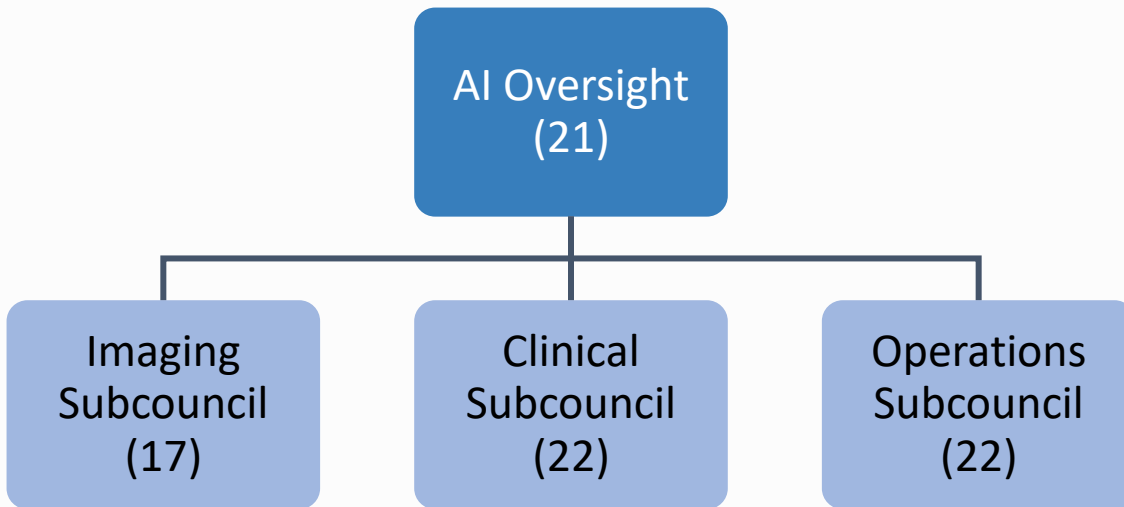
# Integrated Governance Structure



SIP should be initiated when there is software integration with a clinical system, when the software is intended for enterprise-wide use, or upon request of the study team to review the software

# AI Governance: Operational View

1. Should we pilot?
2. Should we widely deploy?
3. Regular review of AI tools in production
4. Oversee the model cards creation & inventory



## AI Oversight Council

- Co-Chairs:
  - Dr. Kim Peairs, DOM
  - Josh Newman, Health IT
- Subcouncil Co-Chairs:
  - CMIO All Children's
  - Rad Diagnostic Imaging
  - Peds Otolaryngology
  - Internal Medicine
  - Ophthalmology
  - Revenue Cycle
- CMIO Bayview
- JHM CMIO
- JHHS CNIO
- Internal Medicine
- JH Imaging
- Community Physicians
- Controllers Office
- Berman Inst of Bioethics
- Pediatrics
- Legal - IT
- Patient Safety
- Research Governance
- Office of the Data Trust

# Core Principles

Subcouncils ensure proposals sufficiently meet our core principles.



ADVANCE



DISAPPROVE

- Can table for corrections, further info
- Often request pilot phase and re-eval



# Questionnaire Questions

1. What clinical or operational problem are you trying to solve?
2. How will AI help solve that problem?
3. Who are the users of the AI system (Patient, Clinician, etc.), the users of the output, and what actions will they be taking? What is the impact on these users?
4. What are the metrics for effectiveness, and how will you measure them? How frequently with the AI system be evaluated?
5. What steps have been taken to ensure that the AI system is fair and will be used ethically?
6. What new or material risks does the AI system introduce for the user, the patient, organization, or others impacted by the output? How will those be mitigated?

# AI Evaluation Rubric

	1 - Inadequate	2 - Poor	3 - Fair	4 - Good	5- Excellent
1. What clinical or operational problem are you trying to solve?					
2. How will the AI system help solve that problem?					
3. Who are the users of the AI system (patient, clinician, etc.) and the users of the output? What actions will they take? What is the impact on these users? <sup>1</sup>					

4. What are the metrics for effectiveness, and how will you measure them? What is the ROI/value of the AI system? How frequently will the model be evaluated?
5. What steps have been taken to ensure that the AI system is fair and will be used ethically?
6. What new or material risks <sup>2</sup> does the AI system introduce for the use of the patient, the organization, and others impacted by the output? How will those risks be mitigated?

Questions with Rating Criteria:

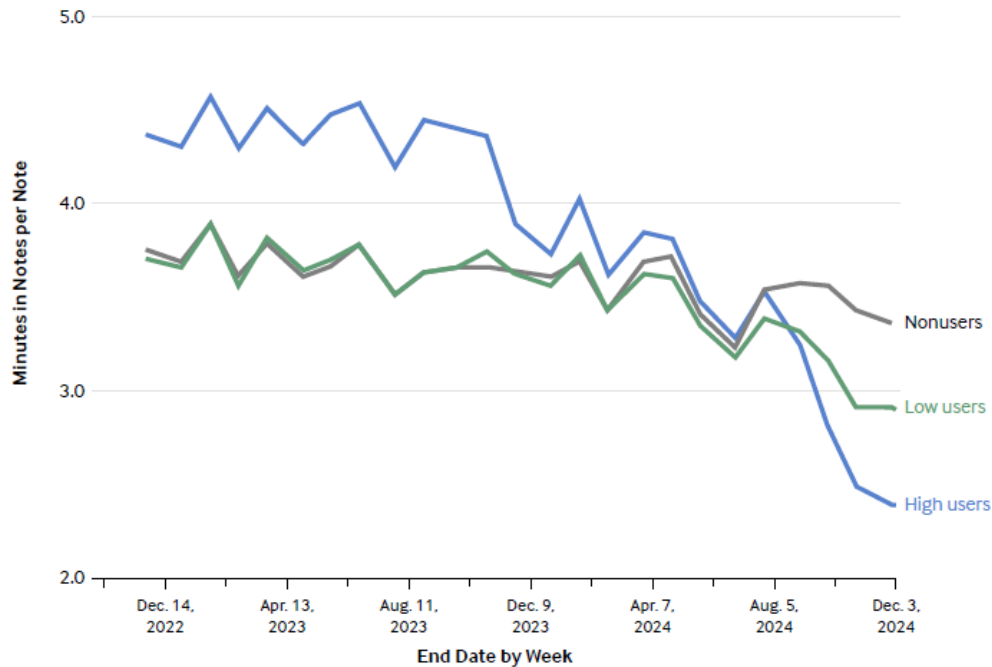
	1 – Inadequate	2 – Poor	3 – Fair	4 – Good	5- Excellent
1. Importance of clinical, imaging or operational problem to be addressed with an AI system	The issue presented is not a genuine clinical, imaging or operational problem.	The issue is vaguely defined and lacks enough context to determine its importance. It is unclear whether the problem is worth solving.	The problem is adequately defined, with some detail or context provided. It appears to be a genuine issue. However, the significance of solving this problem may not be clear.	The problem is well-defined, with sufficient detail and context provided. The issue is clearly a genuine clinical or operational problem, and the proposed AI solution is relevant and appears to have the potential to effectively address the problem.	The problem is exceptionally well-defined, with comprehensive detail and context provided. The issue is a significant clinical or operational problem that is worth solving.
2. Degree to which the AI system could address the identified problem	The proposed AI solution lacks a clear connection to the problem, or the explanation of how the solution will address the issue is absent or	The proposed AI solution has a weak or unclear connection to the problem. The explanation of how the solution will address the issue is vague,	The proposed AI solution has an identifiable connection to the problem. The explanation of how the solution will address the issue is moderately clear, but it	The proposed AI solution has a strong connection to the problem. The explanation of how the solution will address the issue is clear and	The proposed AI solution has an exceptional connection to the problem. The explanation of how the solution will address the issue is

Using an evaluation rubric to help guide sub-councils.

# Ambient "AI Scribe" Evidence from Kaiser

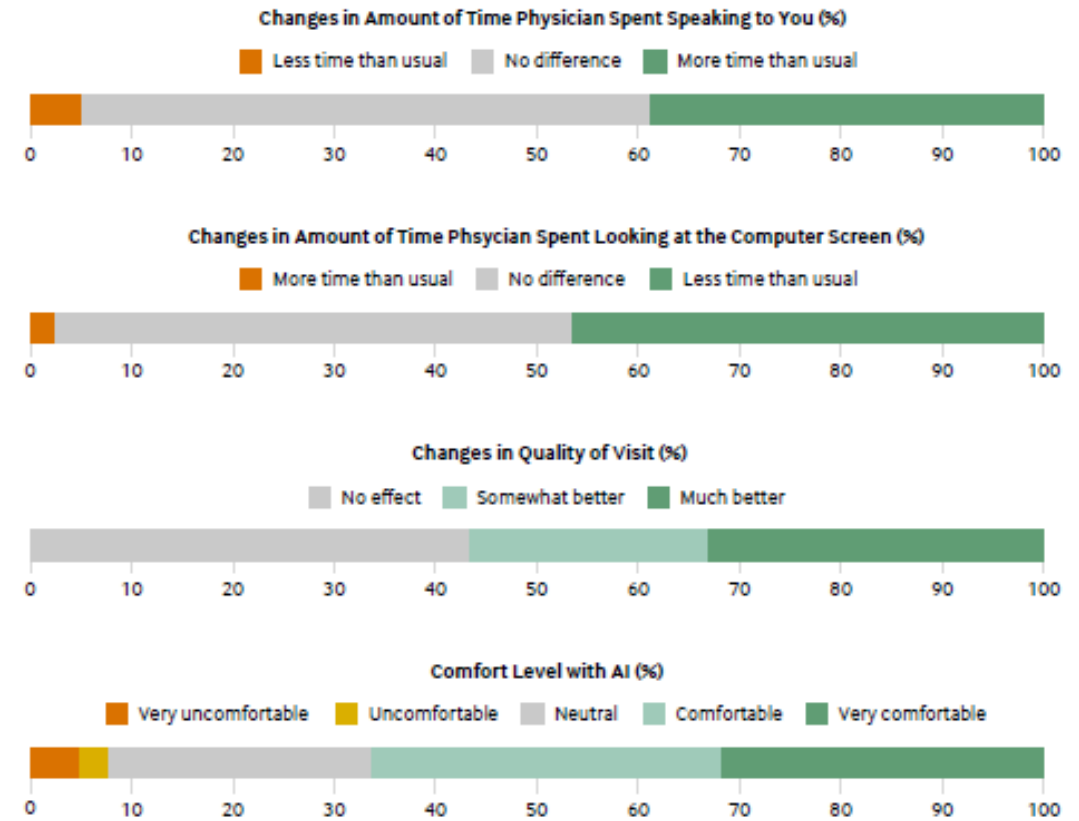
## Time in Notes per Note for all Physicians by AI Scribe Usage Level

This chart shows the time spent on AI scribe notes, in minutes per note, for nonusers, low users, and high users from December 2022 to December 2024.



“...reduction of **0.7 minutes per note for high users** compared with **0.15 for low users**, relative to nonusers... In aggregate, the time savings on time spent in notes for users, compared with nonusers, was estimated to be **15,791 hours or roughly 1,794 8-hour working days** in the year following implementation”

## Patient Perspective



Source: NEJM Catalyst

NEJM Catalyst (catalyst.nejm.org) © Massachusetts Medical Society

Source: “Ambient Artificial Intelligence Scribes: Learnings after 1 Year and over 2.5 Million Uses”, Tierney et al, NEJM Catalyst, May 2025

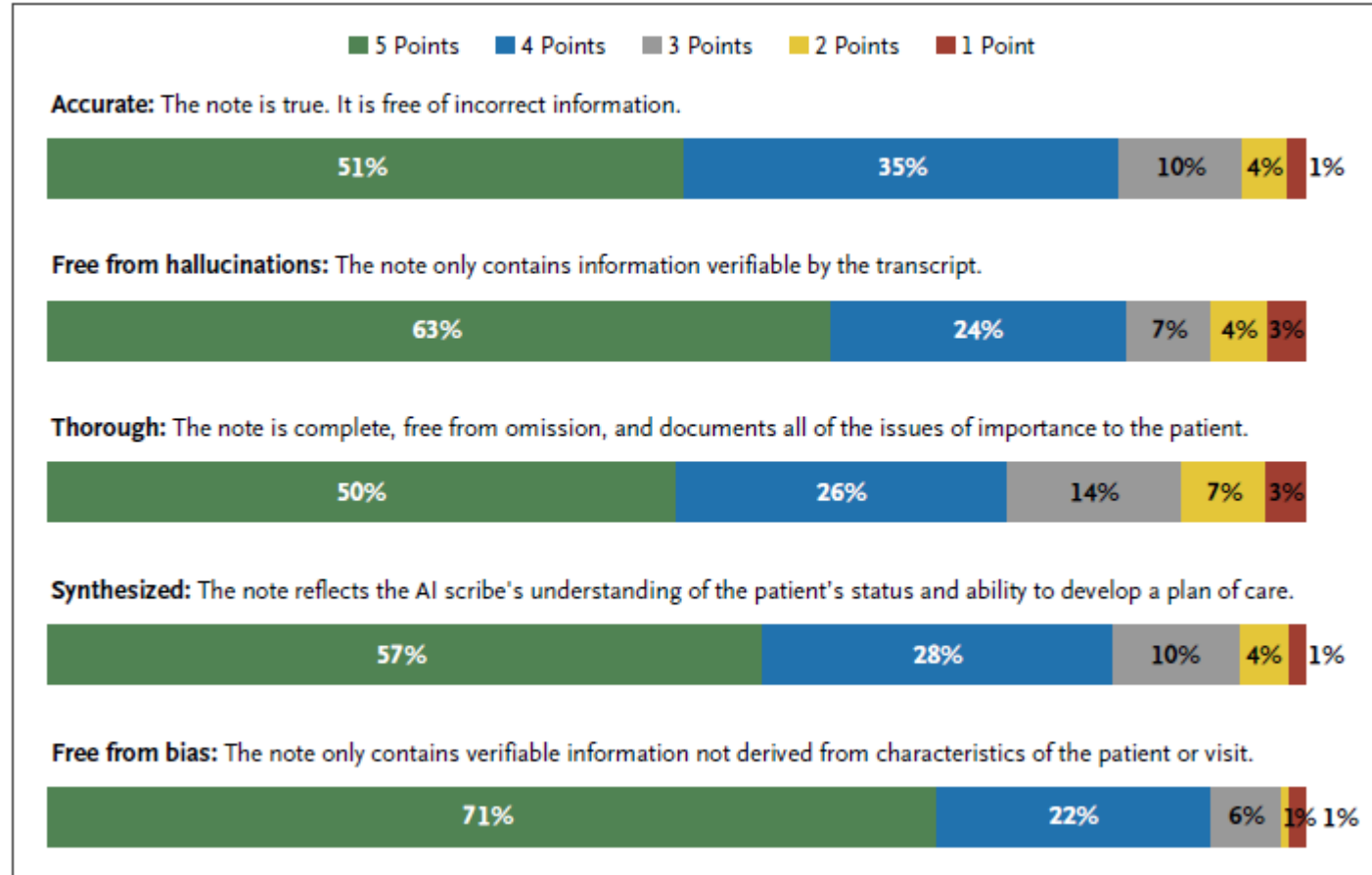
# Quality Assurance Paradigm for Rapid Implementation of AI “Scribe”

Executive Sponsors	Quality Assurance Team	Geographic Implementation Teams	Technology Users
<ul style="list-style-type: none"><li>• Allocate resources</li><li>• Review QA plan and its links to deployment decisions</li><li>• Review QA results</li><li>• Assess risk and interface with in organization</li><li>• Make deployment decisions</li></ul>	<ul style="list-style-type: none"><li>• Design QA plan</li><li>• Champion QA process</li><li>• Adapt data collection instruments</li><li>• Compile QA data</li><li>• Conduct data analysis</li><li>• Report results</li><li>• Provide feedback to stakeholders</li><li>• Provide feedback to vendor</li></ul>	<ul style="list-style-type: none"><li>• Select and train users</li><li>• Give input on QA plan</li><li>• Champion QA data collection</li><li>• Conduct initial analysis of user feedback</li><li>• Respond to user issues</li><li>• Determine specialty deployment readiness in partnership with clinical leaders</li></ul>	<ul style="list-style-type: none"><li>• Assess the technology for their discipline and personal practice style</li><li>• Provide feedback on technology performance (qualitative and quantitative)</li><li>• Contribute to requests for model improvements</li></ul>

Source: “Quality Assurance during the Rapid Implementation of an AI-Assisted Clinical Documentation Support Tool”, Cain et al, NEJM AI, March 2025

# Ambient "AI Scribe" Evidence from Kaiser

## Clinician Ambient Note QA Review



**Total 4 or 5 out of 5 Rating:**

Accurate: 86%

Free from hallucination: 87%

Thorough: 76%

Synthesized: 85%

Free from bias: 93%

**Source:** "Quality Assurance during the Rapid Implementation of an AI-Assisted Clinical Documentation Support Tool", Cain et al, NEJM AI, March 2025



THANK YOU