Case Study: Engineering Safety for AI-Enabled Medical Applications

Presented by :

Attrayee (Atty) Chakraborty, MS, MSc., CQSP

Disclaimer

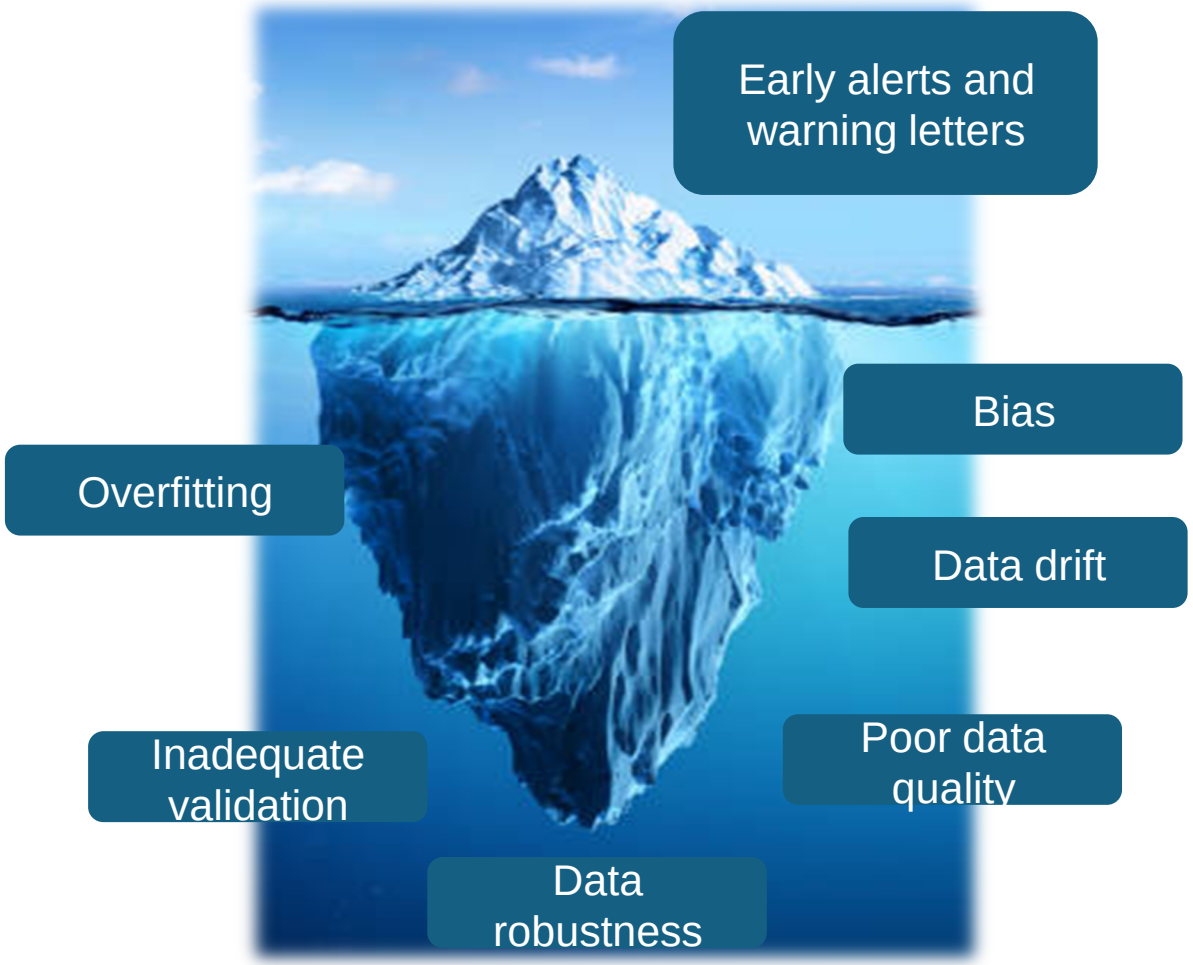
The views and opinions expressed in this presentation are solely those of the individual presenters and do not necessarily represent the views, positions, or policies of their respective employers, or any of their affiliates, directors, officers, or employees.

The content is provided for informational purposes only and should not be construed as official guidance, policy, or endorsement by any organization with which the presenters are affiliated.

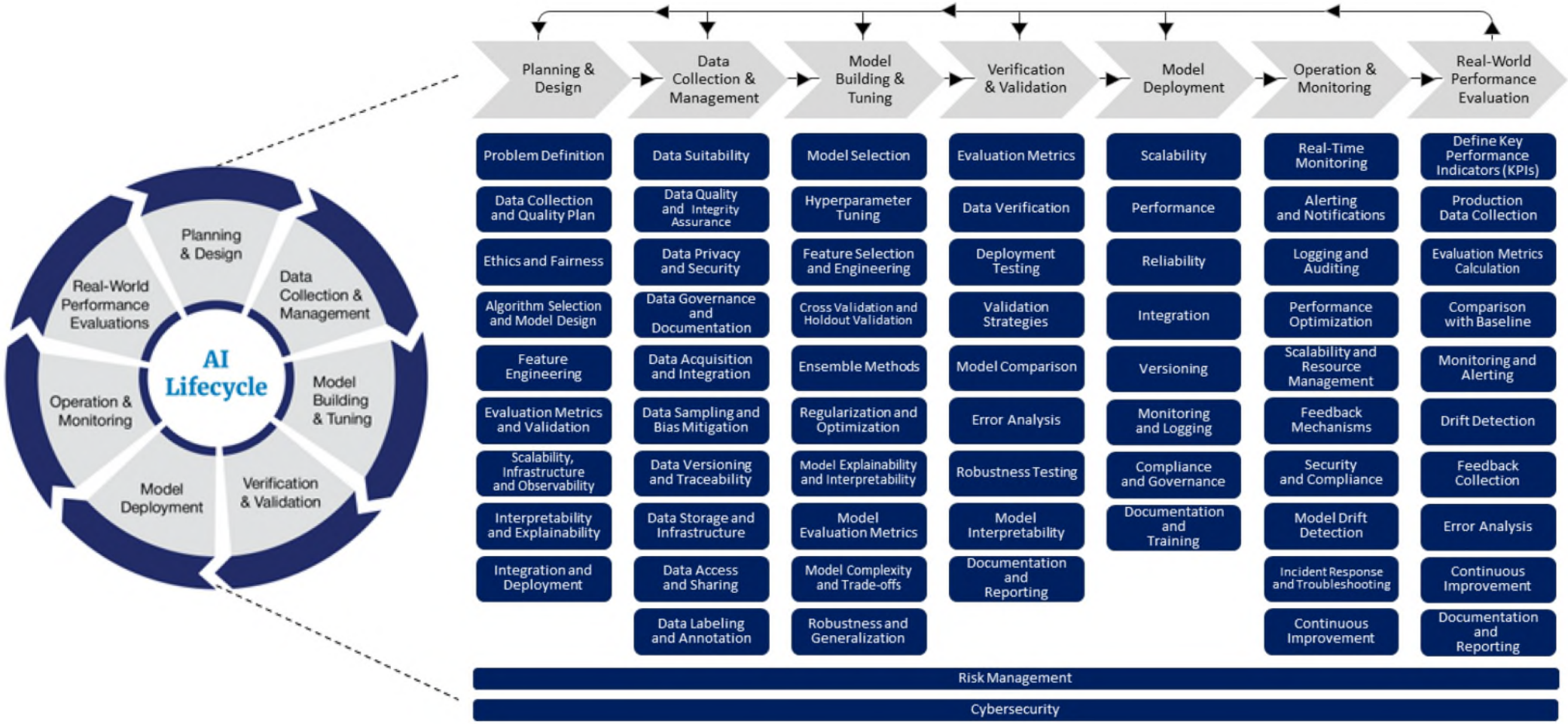
Agenda

1. Risk management frameworks for AI-enabled medical devices
2. Understanding fundamental differences in deterministic vs non-deterministic medical devices
3. Case-study: example of assessing risks for AI-enabled medical device

We just see the tip of the iceberg...



Risk management across AILC



[Blog: A Lifecycle Management Approach toward Delivering Safe, Effective AI-enabled Health Care | FDA]

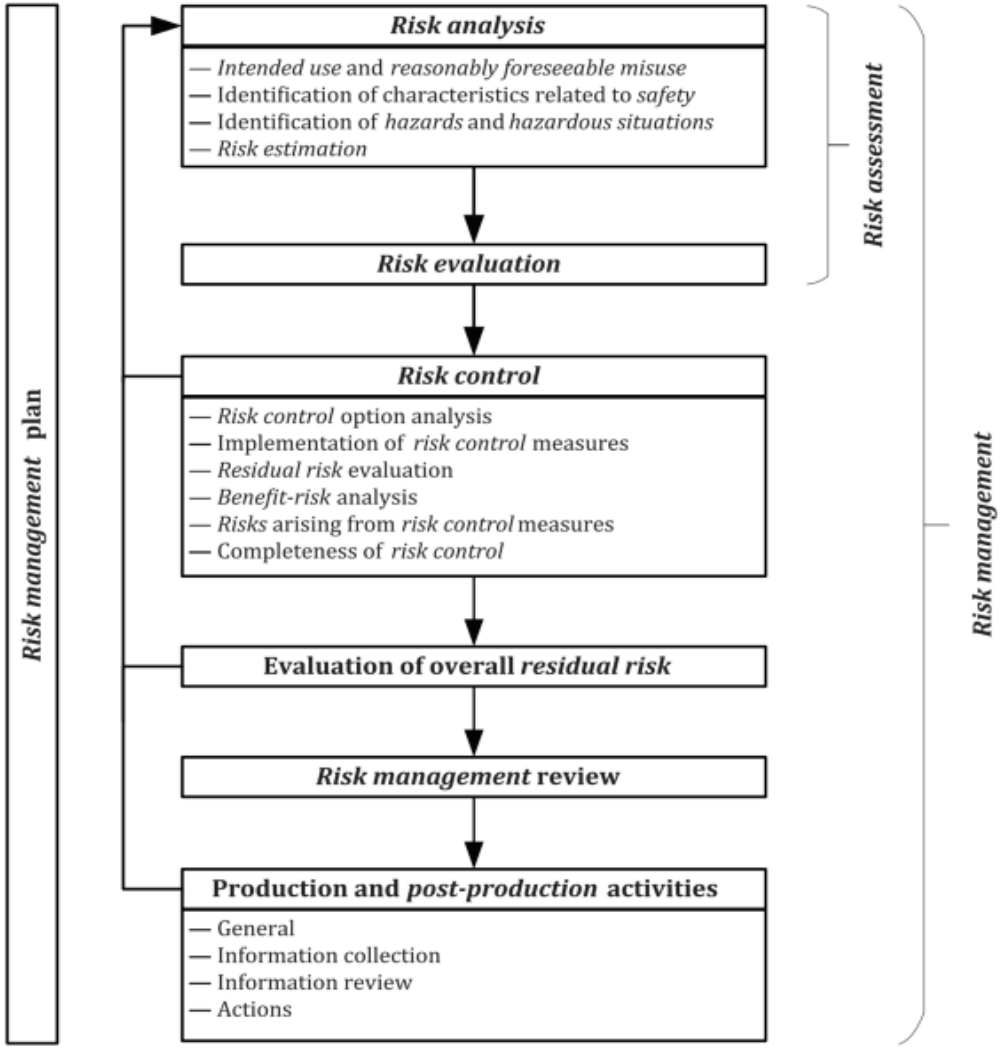
Risk management frameworks for AI-enabled medical devices

AAMI 34971

AAMI 34971 vs ISO 14971

Risk Management Frameworks for AI-Enabled Medical Device

ISO 14971:2019



BS/AAMI 34971:2023

User training

Explainability

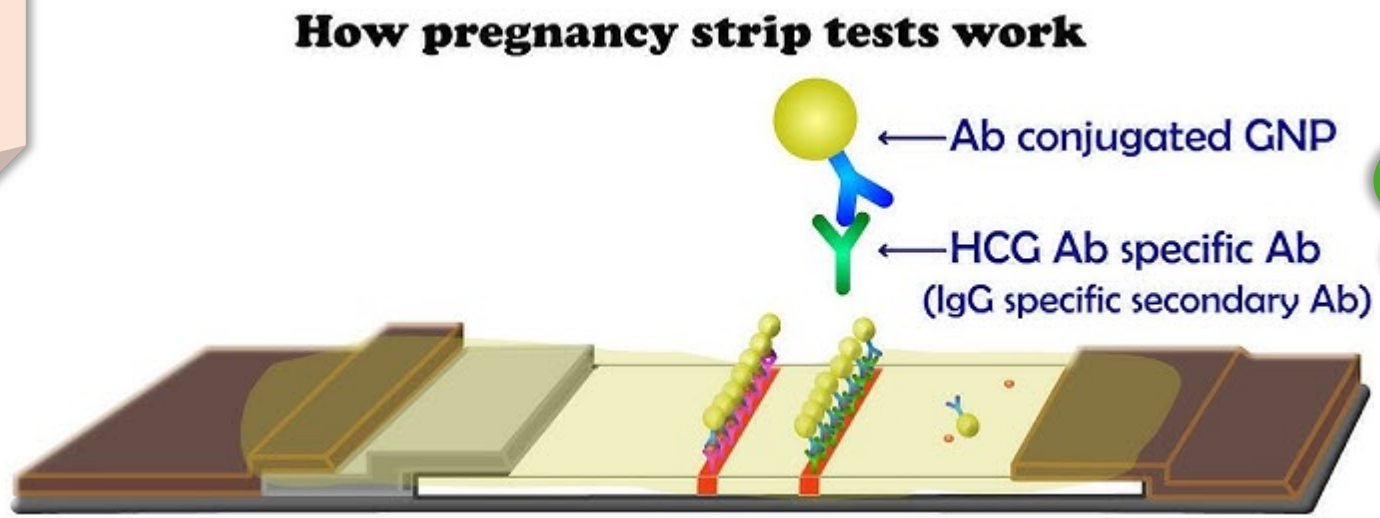
Data Drift

FMEAs for AI enabled medical devices: what's different?

DETERMINISTIC MEDICAL DEVICES

Input	Process	Output
hCG levels	Detect hCG concentration Clinician is aware of mechanism of action	Positive: 20–25 mIU/mL

Explicit, deterministic, and reproducible



Hardware failures
Manufacturing defects
Expired reagents

Image Source: https://www.youtube.com/watch?v=Tv9B4m7o364

FMEAs for AI enabled medical devices: what's different?

NON-DETERMINISTIC MEDICAL DEVICES

Opaque, probabilistic, and non-reproducible

Input	Process	Output
Dermoscopic or smartphone image of a skin lesion	**Clinicians cannot see *which specific pixel patterns or anatomical features* led to that decision** • The network may attend to subtle texture, color gradients, or vessel-patterns that are **not visible or interpretable** to dermatologists, even with saliency maps. • The same image resized, color-corrected, or captured under different lighting can yield different outputs, despite the input being "essentially" the same.	Classification (e.g., "benign," "indeterminate," "malignant") with confidence.

Out-of-Distribution input
Small/overfitted datasets
Data bias

FMEA

Failure mode and effects analysis for tablet packaging

Process step/input	Potential failure mode	Potential failure effects	Severity	Potential causes	Occurrence	Current controls	Detection	RPN	Actions recommended	Responsibility	Actions taken	Severity	Occurence	Detection	RPN
What is the process step/input under investigation?	In what ways does the key input go wrong?	What is the impact on the key output variables (customer requirements) or internal requirements?	How severe is the effect to the customer?	What causes the key input to go wrong?	How often does cause of failure mode occur?	What are the existing controls and procedures (inspection and test) that prevent the cause of the failure mode? **Should include an SOP number.**	How well can you detect cause or failure mode?		What are the actions for reducing the occurrences of the Cause or improving detection?	Who's responsible for the recommended action?	What are the completed actions taken with the recalculated RPN? **Be sure to include completion month/year.**				
Sealing/ temperature	Temp. too high	Burned blister pack	10	Wrong setting	6	Verification of batch record	7	420	Provide infrared temperature device to operator	M. Peña	Temperature device implemented (8/11)	10	4	2	80
Sealing/ temperature	Temp. too high	Blister pack not sealed completely	9	Wrong setting	6	Verification of batch record	7	378	Provide infrared temperature device to operator	M. Peña	Temperature device implemented (8/11)	9	4	2	72
Sealing/ press time	Too much time	Burned blister pack	10	Machine not set properly	5	Verification of batch record	7	350	Provide a visual display to see time elapsed	J. Rodriguez	Visual display implemented (10/11)	10	3	2	60
Sealing/ press time	Not enough time	Blister pack not sealed completely	9	Machine not set properly	5	Verification of batch record	7	315	Provide a visual display to see time elapsed	J. Rodriguez	Visual display implemented	9	3	2	54

RPN = risk priority number
SOP = standard operating procedure

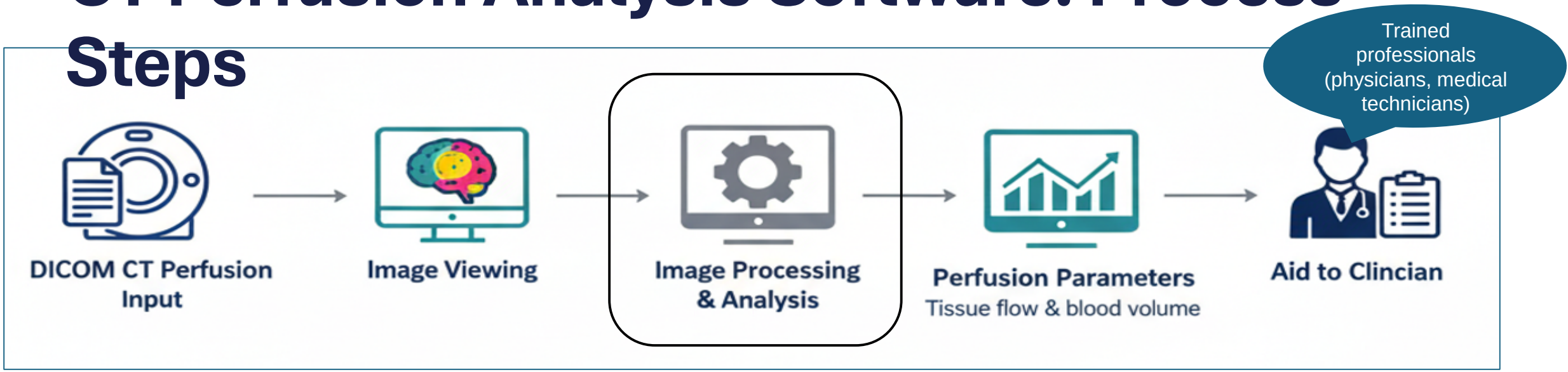
What is FMEA? Failure Mode & Effects Analysis | ASQ

Case study:

CT Perfusion image analysis software for stroke assessment

PROCESS FLOW

CT Perfusion Analysis Software: Process Steps



Intended Use: Visualization, processing, analysis of CT perfusion data for clinical decision support for ischemic stroke detection

Regulatory classification: Class II (510(k))

Functionality:

- Viewing, processing, analysis of CT perfusion images
- Visualization of contrast over time
- Calculation of perfusion and blood volume parameters

STEP 1: What is the process step/input under investigation?

CRITICALITY OF THE DEVICE

How critical are AI outputs on patient safety?

Level of human control on output

Table D.2 — *Possible adaptation of SaMD Categories to autonomous systems*

State of healthcare situation or condition	Significance of information provided by SaMD to healthcare decision				
	Treat or diagnose with no intervention possible	Treat or diagnose with override	Treat or diagnose with approval	Drive clinical management	Inform clinical management
Critical	TBD	TBD	IV	III	II
Serious	TBD	IV	III	II	I
Non-serious	IV	III	II	I	I

[SOURCE: Building Explainability and Trust for AI in Healthcare, Xavier Health [18]]

Severity Level	Key Characteristics	Disease/Condition Type	Target Population	User Type
Critical	Accurate/timely action vital to avoid death, long-term disability, or serious health deterioration; may impact public health.	Life-threate (incl. inc requir the in s	Fragi pedi	Specia ained
Serious	Accurate diagnosis/treatment vital to avoid unnecessary interventions or mitigate long-term irreversible consequences. Note: Lay users without professional support → treat as Critical.	Moderate progression (often curable); no major interventions; not time-critical.	Not fragile.	Specialized or lay users.
Non-Serious	Accurate diagnosis/treatment important but not critical for avoiding irreversible consequences.	Slow/predictable progression (minor/chronic); manageable; minor/noninvasive interventions.	Individuals (may not be patients).	Specialized or lay users.

Ischemic stroke

Elderly, acute patients

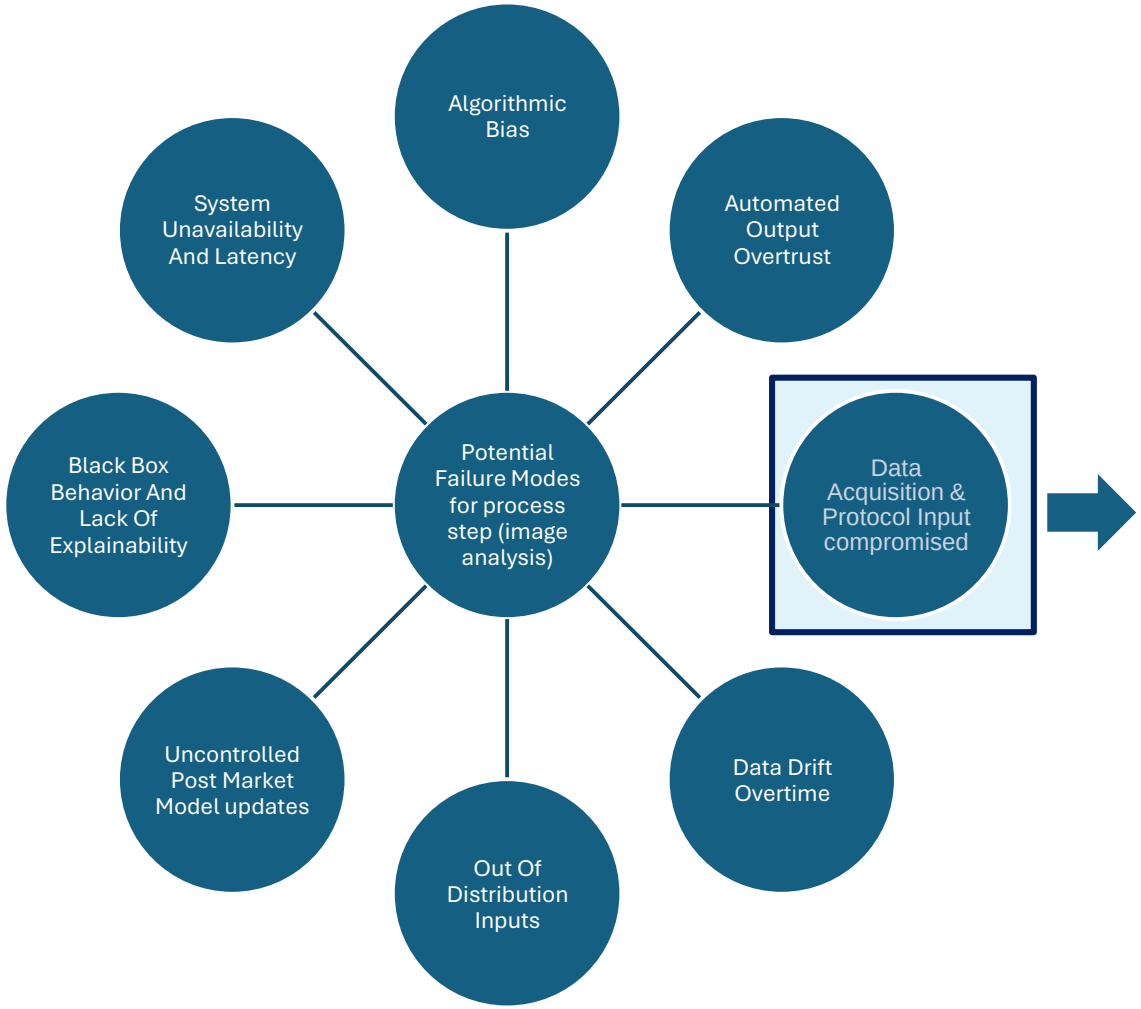
Radiologists

"Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations

Relate back to intended use

POTENTIAL FAILURE MODE

Identify Potential Failure Modes for process step (image analysis)



Protocol input is not equivalent to the software's intended operating conditions.

Most CT perfusion software expects a **standardized protocol** and specific **DICOM metadata** (slice thickness, reconstruction kernel, contrast timing, table position, acquisition timing).

STEP 2: In what ways does the key input go wrong?

POTENTIAL FAILURE EFFECTS AND SEVERITY

Identify Potential Failure Effects and Severity

Potential effects of failure: OOD parameters have clinical effects

Software may fail to:

- Decode or pre-process DICOM incorrectly (compromise reliability)
- Produce biased or unstable perfusion maps
- Greater risk of false positive or false negative results
- Silent failures

5 (Critical/Catastrophic)

STEP 3 and 4: What is the impact on the key output variables (patient safety)? How severe is the effect to the customer (patient)?

POTENTIAL CAUSES

Identify Potential Failure Causes

Potential Causes of Failure

- Incomplete/corrupted DICOM data from non-standard scanner protocol
- AI model receives out-of-distribution imaging parameters (unknown image characteristics)

STEP 5: What causes the key input to go wrong?

OCCURRENCE

Estimate Occurrence

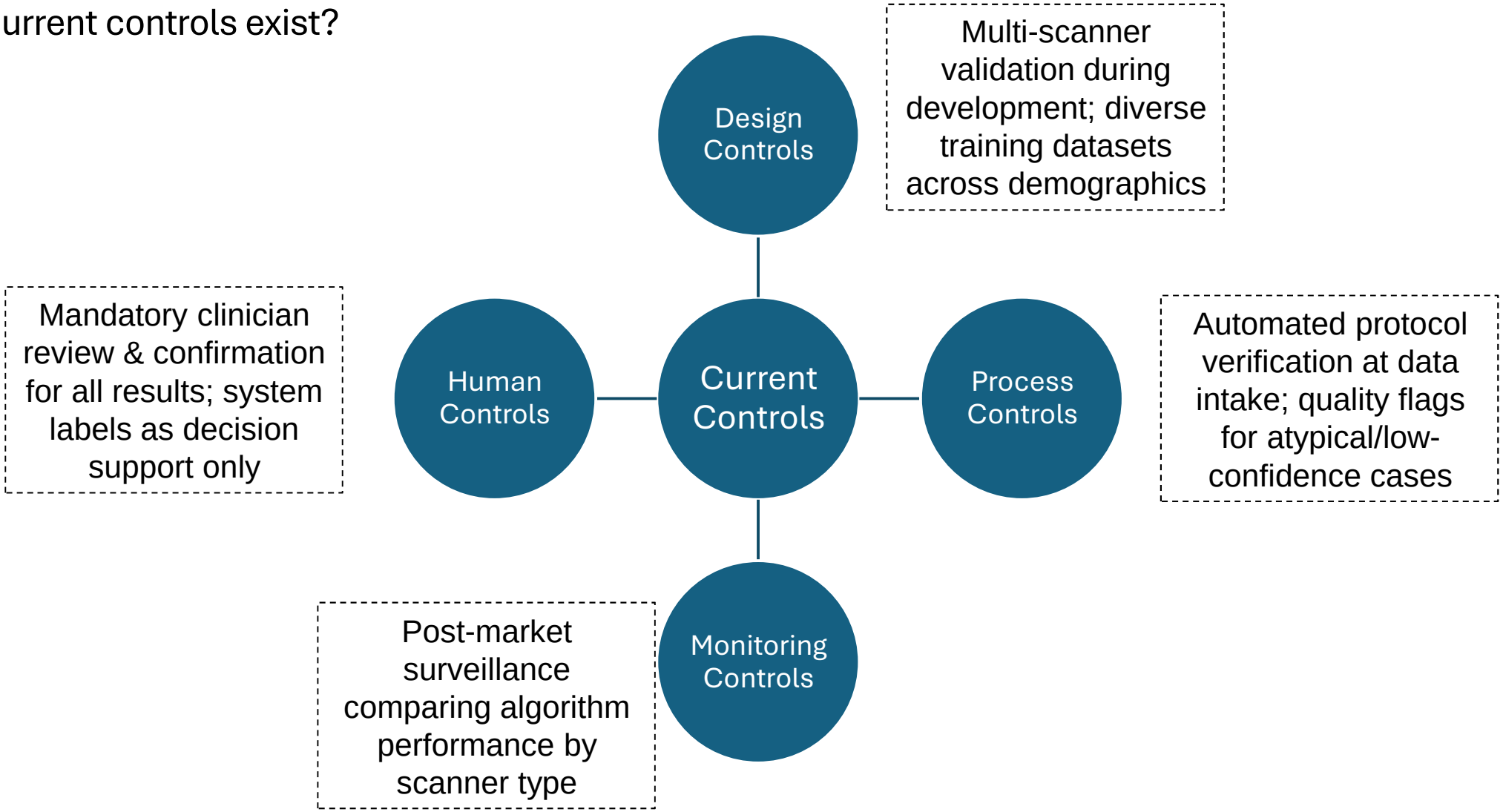

PMS updates RM

Failure Mode	Hazard	Hazardous Situation	Probability of Hazardous Situation Occurring (P1)	Probability of hazardous situation leading to harm (P2)	Probability of Occurrence of Harm (P = P1 X P2)	Key Evidence	Occurrence Rating
Data Acquisition & Protocol Input compromised	Corrupted/missing input data to AI perfusion analysis software.	Software fails to process/decode DICOM (e.g., missing tags, frame misalignment), producing invalid perfusion maps (CBF/Tmax).	0.1	0.1	0.01	<1 event per ~1k–10k transfers. PACS data/clinical reports confirm "remote" but not rare.	5
	Input outside trained distribution (e.g., unusual kVp/slice thickness).	Model outputs unreliable perfusion metrics (biased CBF/CBV), potentially misleading clinician interpretation.	0.01	0.01	10^-5	Trained radiologists cross-check with CTA/NCT; mismatch errors caught ~99%)	4

STEP 6: How often does cause of failure mode occur?

CURRENT CONTROLS

What current controls exist?



STEP 7: What are existing controls?

DETECTION

How to estimate detection?

Failure Cause	Detection Rating	Rationale
Incomplete/Corrupted DICOM data from non-standard scanner protocol	2 (High chance of detection)	Often caught automatically (80–95% via header validation); e.g., CT software flags non-standard syntax. Rare misses in rushed workflows.
AI model receives OOD imaging parameters (unknown image characteristics)	3 (Moderate chance of detection)	Effective with AI QC (90%+ AUROC in studies); e.g., iStroke/RAPID flags poor mismatch. Lower if subtle shifts.

STEP 8: How well can you detect cause of failure mode?

Putting it all together!

RPN and downstream activities

RPN and Downstream Activities

Before Mitigation

After Mitigation

Process input	Potential Failure Mode	Potential Failure Effects	Severity	Potential Causes	Occurrence	Detection	RPN	Addl. Risk Controls	Severity	Occurrence	Detection	RPN
Image analysis	Data Acquisition & Protocol Input compromised	Software may fail to: ▪ Decode or pre-process DICOM correctly ▪ Produce biased or unstable perfusion maps ▪ Reduce risk of false positive or false negative results	Critical/ catastrophic (5)	Incomplete/corrupted DICOM data from non-standard scanner protocol	5	2	50	Input data validation, standardized protocols	Critical/ catastrophic (5)	3	2	30
				AI model receives out-of-distribution imaging parameters (unknown image characteristics)	4	3	60	Alarm system to flag incompatible input		4	1	20

Additional risk control controls input

Additional risk control improves detection

Refer to BS/AAMI 34971 for between hazards, foreseeable sequences of events, hazardous situations, harm and potential risk control measures

Conclusion

Integrating thoughts into the AI Lifecycle (AILC)

Key Takeaways

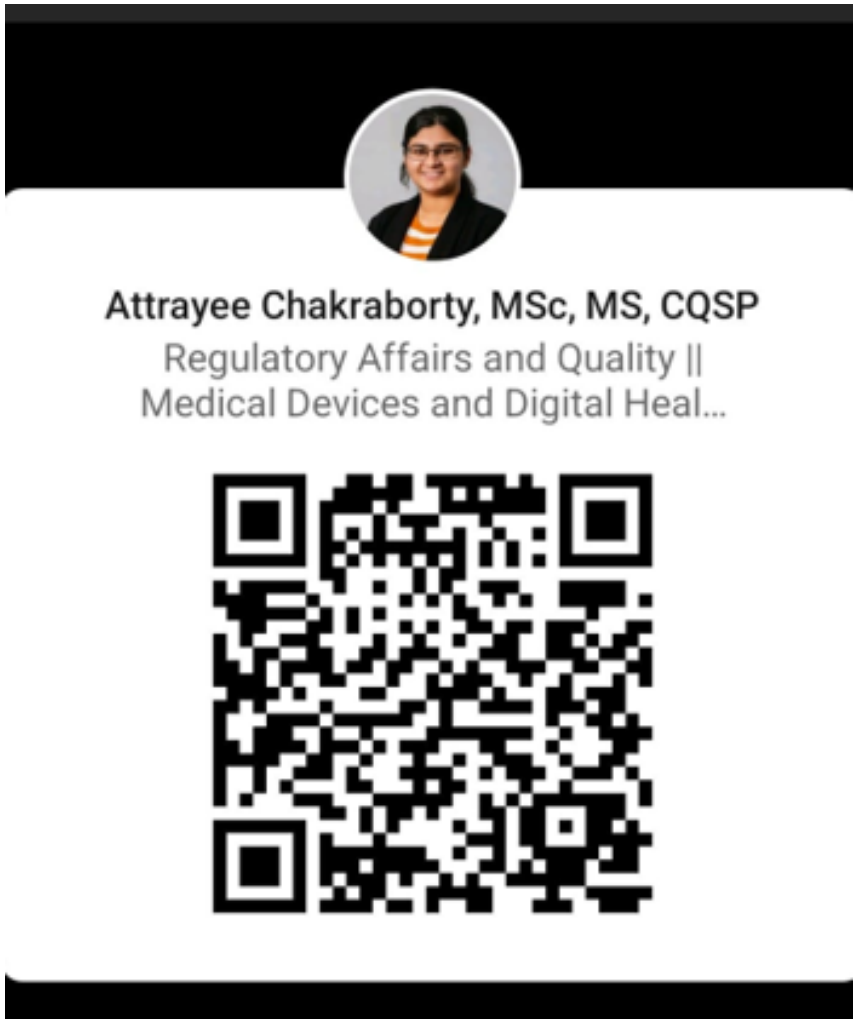
- AI risks hide below the surface — continuous vigilance required.
- Risk management spans the entire AI lifecycle.
- AI behaves non-deterministically — design accordingly.
- FMEA must account for AI-specific failure modes.
- Regulators expect lifecycle readiness, not point-in-time compliance

Key Takeaways

- AI risks hide below the surface — continuous vigilance required.
- Risk management spans the entire AI lifecycle.
- AI behaves non-deterministically — design accordingly.
- FMEA must account for AI-specific failure modes.
- Regulators expect lifecycle readiness, not point-in-time compliance

Thank You!

Any Questions?



Attrayee Chakraborty, MSc, MS, CQSP

Regulatory Affairs and Quality ||
Medical Devices and Digital Heal...