Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Key Considerations Before AI Deployment in Healthcare

Tint Lwin Win

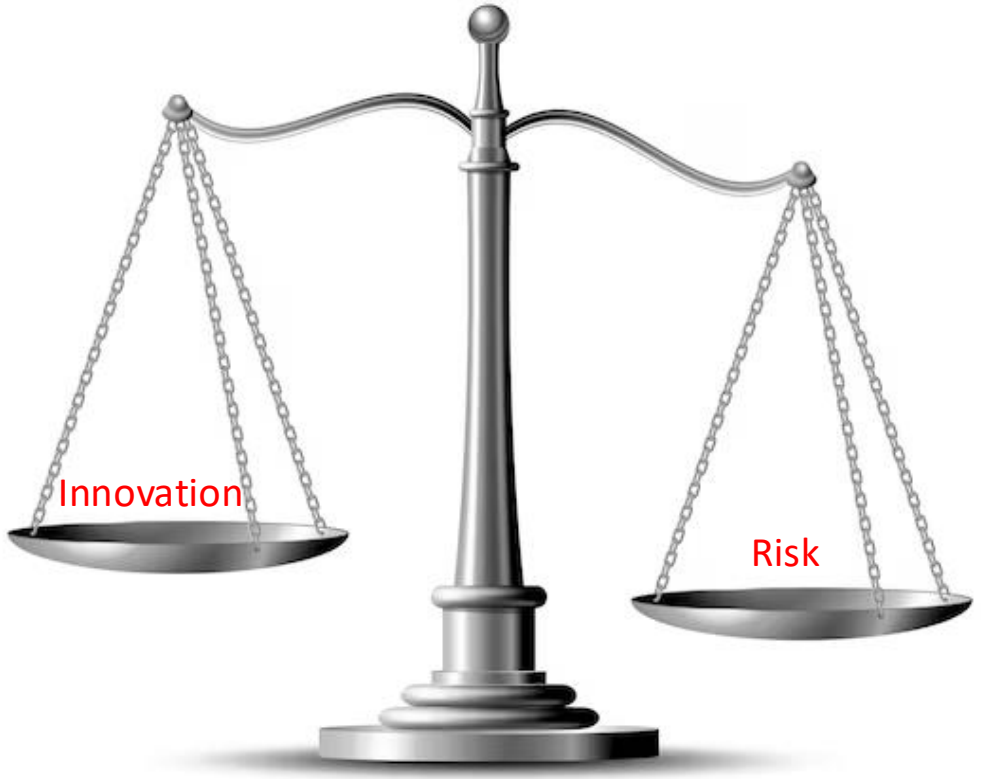
Student | Data Science in Healthcare and Clinical Informatics

23 January 2026

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Balancing Innovation and Risk in Clinical AI

- Rising adoption of AI in clinical practice.
  - Potential Benefits:
    - Improved outcomes, efficiency, reduce cost.
  - Potential Risks:
    - Health inequities, patient harm.
- Need for **governance and risk controls**.



- To ensure AI is beneficial, safe, and equitable.

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

AI-in-Healthcare Frameworks

2020

SPIRIT-AI & CONSORT-AI[1]

First international reporting standards for clinical trials of AI interventions

2021

WHO Ethics & Governance of AI for Health

Global ethical principles for safe AI use.

Good Machine Learning Practice (GMLP)[2]

for Medical Device Development – Guiding Principles (FDA, Health Canada, MHRA

2022

NIST AI Risk Management Framework[3]

National Institute of Standards and Technology Systematic framework for AI risk governance.

2023

FUTURE-AI[4]

Trustworthy AI standards for medical imaging

2024

WHO Generative AI & LMM Guidance[5]

Governance for large multimodal models in health

2025

FAIR-AI[6]

A Practical Framework for Appropriate Implementation & Review of AI in Healthcare

STARD-AI[7]

Standards for Reporting Diagnostic Accuracy (STARD) for AI tools.

[1] Ibrahim, H., Liu, X., Rivera, S. C., Moher, D., Chan, A. W., Sydes, M. R., Calvert, M. J., & Denniston, A. K. (2021). Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. In *Trials* (Vol. 22, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13063-020-04951-6
[2] Diamond, & Matthew. (2021). *Good Machine Learning Practice for Medical Device Development: Guiding Principles*.
[3] Tabassi, E. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. https://doi.org/10.6028/NIST.AI.100-1
[4] Lekadir, K., Frangi, et. al. (2025). FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ*. https://doi.org/10.1136/bmj-2024-081554
[5] Ethics and governance of artificial intelligence for health. Guidance on large multi-modal models. Geneva: World Health Organization; 2024. Licence: CC BY-NC-SA 3.0 IGO.
[6] Wells, B. J., et. al. (2025). A practical framework for appropriate implementation and review of artificial intelligence (FAIR-AI) in healthcare. *Npj Digital Medicine*, *8*(1). https://doi.org/10.1038/s41746-025-01900-y
[7] Sounderajah V, Ashrafian H, Golub RM, Shetty S, De Fauw J, Hooft L, Moons K, Collins G, Moher D, Bossuyt PM, Darzi A. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ open. 2021 Jun 1;11(6):e047709.

The Problem: Missing practical guidance

- Current frameworks are often:
  - Non-binding or too theoretical
  - Lacking practical guidance for real-world implementation
  - Inflexible regarding 'high-risk' definitions in healthcare contexts
- **Goal**: To create a standardized, repeatable, and transparent framework.

A framework for appropriate implementation and review of artificial (**FAIR-AI**)[1]

- It was proposed in 2025
- System-level framework for safe, ethical, and appropriate AI deployment.
- Focuses on readiness, risk assessment, and structured review.
- Provides practical guidance for real-world implementation.

Methodology: How FAIR-AI was built

Review

Narrative Review: Synthesized best practices (Validity, Equity, Transparency).

Interviews

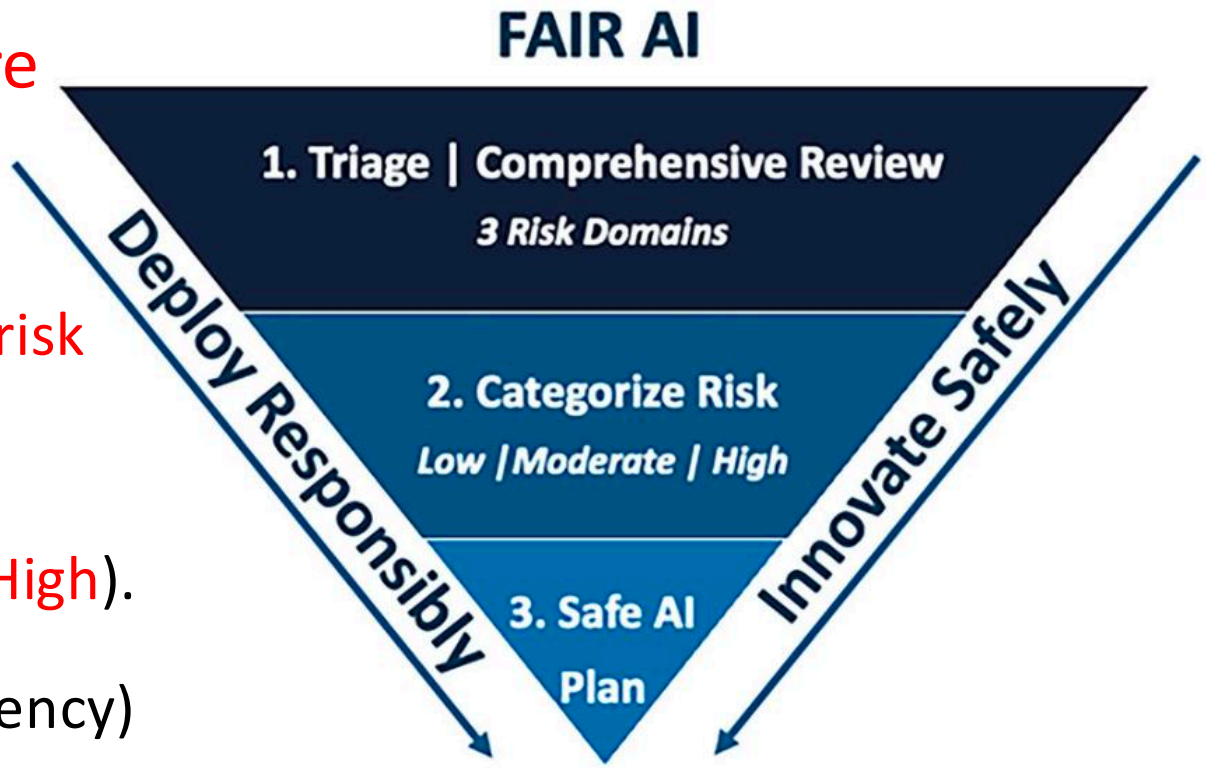
Stakeholder Interviews: Leaders, developers, healthcare providers, and patients.

Workshop

Design Workshop: Consensus from 33 multidisciplinary experts (individuals with expertise in legal affairs, regulatory compliance, cyber security, ethics, clinical care, clinical informatics, data science, and research)

Overview: the triangle of responsibility

- The framework is built on three core components
    1. Triage & Comprehensive Review (3 risk domains).
    2. Categorize Risk (Low | Moderate | High).
    3. Safe AI Plan (monitoring & transparency)



Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Foundational Elements

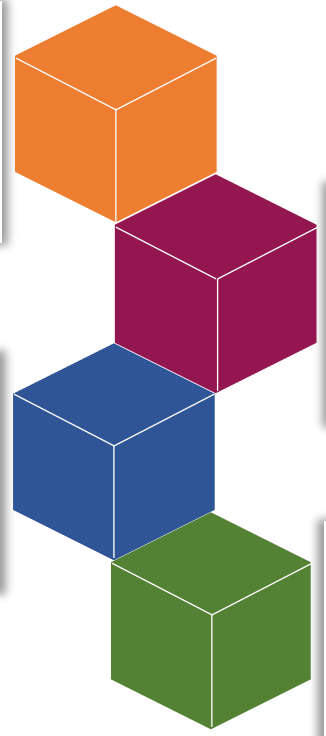
Required before implementation of AI tools in healthcare

Artifacts

Ethics statement & guiding principles.

Escalation Body

Multidisciplinary Governance Committee



Personnel

Accountable data science team

Inventory Tool

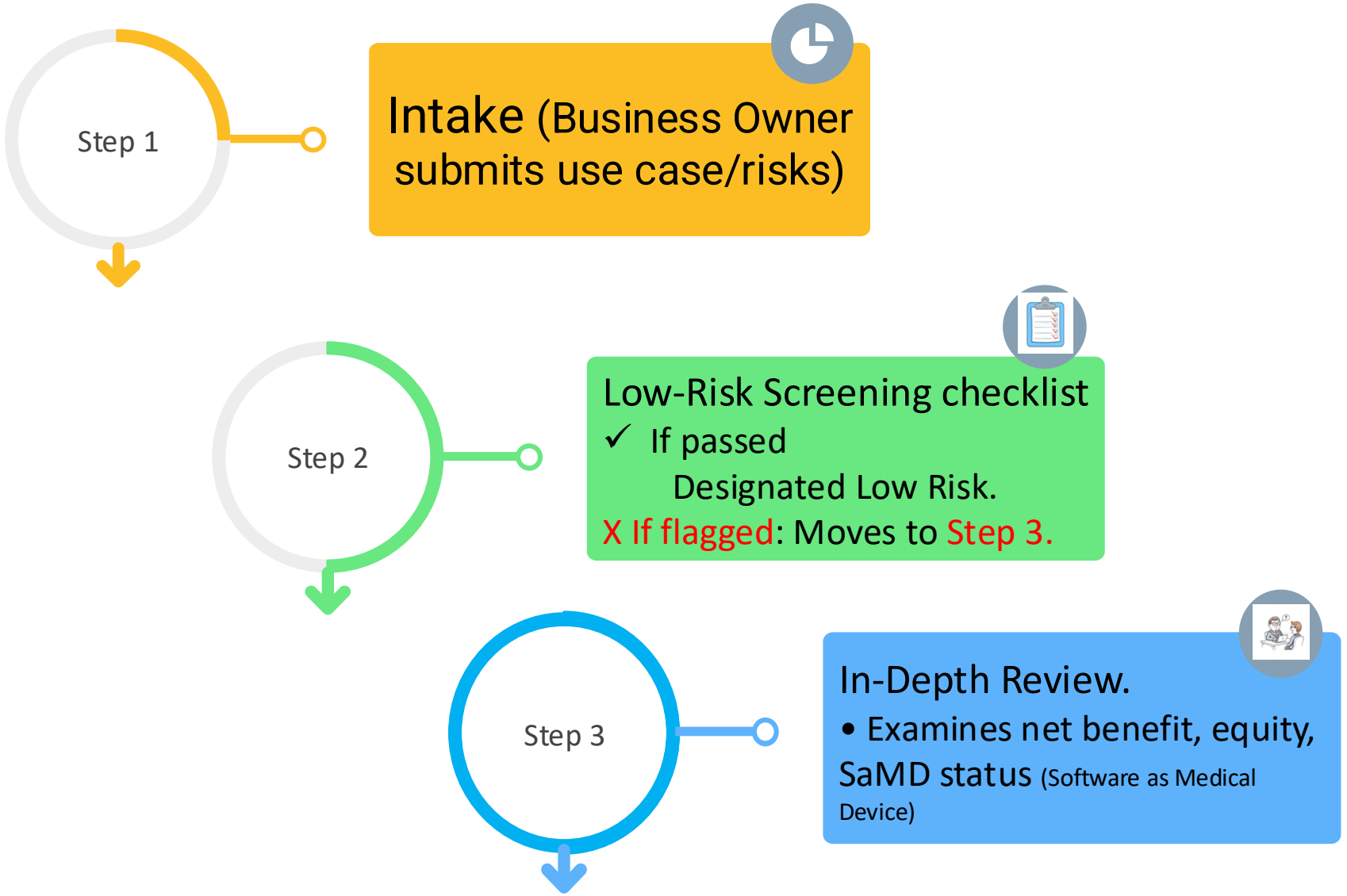
'Single source of truth' catalog.

Scope: What Goes Through FAIR-AI?

- **Inclusion**
  - Broad definition of AI: any computer system capable of activities normally associated with human cognitive effort
- **Exclusions**
  - Simple scoring systems & rules-based tools
  - Physical medical devices (e.g. FDA regulated devices)
  - IRB-approved research protocols

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

The FAIR-AI review process flow



Risk categories as determined by FAIR-AI evaluation and escalation to AI Governance

Risk Category	Data Science Team	AI Governance Committee
	Evaluation	Escalation Determination
Low	All low risk criteria passed	n/a
Medium	No high risk criteria on in-depth review	n/a
High	Any high risk criteria, and/or supporting evidence is lacking	Proceed under high risk conditions Proceed as high risk - pilot or research required Do not proceed - intolerable risk

Transparency: The AI Label

High-risk solutions require an 'AI Label' for end-users.

AI Label

AI Solution Name
- Inventory ID
- Developer
- Business owner
- Model type

FAIR AI Summary

Risk category & initial review date
- Low, Medium, High, Uncertain

- Last review date
- Intended use
- Directions for interpretation of AI solution output
- Performance
- Known limitations
- Potential risks
- FDA approval, if applicable (include type)

Pre-Implementation: The In-Depth Review

1. **Technical Evaluation**
   - Validation evidence & FDA status.
2. **Ethical & Equity Analysis**
   - Bias screening & data representativeness.
   - Explainability of the "Black Box".
3. **Workflow & Safety**
   - **Net Benefit:** Confirming utility exceeds risk.
   - **Human Oversight:** Ensuring a "human-in-the-loop" for critical decisions.

Safe AI Plan: Post-Deployment Continuous Monitoring

- AI models drift and workflows change over time.
- **Requirement**: Periodic attestation by business owner.
  - **Confirm** alignment with intended use case.
  - **Confirm** data/workflows haven't changed.
  - **Confirm** expected benefits are realized.

Limitations of FAIR-AI

- Not a regulatory framework.
- **Resource Intensity**: Requires significant time, personnel, and leadership support, training required for evaluators and end-users.
- **Evaluator Dependency**: Framework relies on the diligence and expertise of reviewers
- **Effectiveness depends** on local implementation.

Real-world AI deployment (The ASSURE Study)

- **The Workflow:** A "Multistage" approach utilizing **Human-in-the-loop**.
  - ***Step 1****:* Radiologist reads with AI support.
  - ***Step 2 (SafeGuard)****:* If the Radiologist says "Normal" but AI says "High Risk," the case is routed to a **second human expert**.
- **Scale:** Comparison of 208,891 AI-assisted exams vs. 370,692 Standard of Care exams across 109 US sites.
- **Key Results:**
  - **Cancer Detection Rate (CDR):** Increased by **21.6%**
  - **Precision (PPV1):** Increased by **15.0%** (More cancers found without excessive false alarms)
  - **Equity:** Improvements were consistent across Black, Hispanic subpopulations

Evaluating the ASSURE Study through the FAIR-AI Framework

- Based on the **FAIR-AI framework**, the **DeepHealth Breast AI deployment** in this healthcare scenario would be categorized as **Moderate Risk,** because
  - This AI tool is not autonomous and is used as a **clinical decision support** system as it requires **radiologist approval** and is subject to **human oversight**.
  - The **radiologist's oversight** and the **SafeGuard Review process** mitigate the risk, but **false negatives** potentially delay the cancer diagnosis.

Evaluating the ASSURE Study through the FAIR-AI Framework

- Good Practices
  - **Safety Net:** The "SafeGuard Review" directly addresses FAIR-AI's "Human Oversight" requirement. The AI cannot unilaterally reject a cancer diagnosis; it triggers a human safety check.
  - **Equity Validation:** They explicitly validated against vulnerable subgroups (race/age) to ensure the model didn't widen disparities, satisfying the FAIR-AI Equity screen.
  - **Net Benefit:** Demonstrated clear clinical utility (increases in CDR/PPV) outweighing the modest increase in recall rate.

Evaluating the ASSURE Study through the FAIR-AI Framework

- Areas for Improvement
  - The study operated under a "waiver of consent". It is unclear if patients were notified that AI was assessing their cancer risk.
  - The paper lacks detailed reporting on governance structure and oversight committees.
  - No explicit categorization of the AI tool's risk level (low, moderate, high).
  - No detailed documentation of ethical review.
  - The study does not mention long-term continuous performance monitoring, and does not have sufficient follow-up data to report sensitivity, specificity, false-negative rates, interval cancers or cancer stage at diagnosis.

Reference

- Wells BJ, Nguyen HM, McWilliams A, Pallini M, Bovi A, Kuzma A, Kramer J, Chou SH, Hetherington T, Corn P, Taylor YJ. A practical framework for appropriate implementation and review of artificial intelligence (FAIR-AI) in healthcare. NPJ digital medicine. 2025 Aug 11;8(1):514.
- Louis, L. D., Wakelin, E. A., McCabe, M. P., Ng, A. Y., Kim, J. G., Lee, C. I., Buist, D. S. M., Gregory Sorensen, A., & Haslam, B. (2025). Equitable impact of an AI-driven breast cancer screening workflow in real-world US-wide deployment. *Nature Health*. https://doi.org/10.1038/s44360-025-00001-0


Thank You