

CORE TRAINING WORKSHOP

# Hands-On Training on Artificial Intelligence

Understanding AI Research Oversight in Healthcare

**Mark Lifson, PhD**

AI/ML Director

Core Training Workshop Series

2025

AI vs Traditional Research

Clinical Intent

Risk Assessment

Oversight Framework



# Mark Lifson, PhD

Director, AI/ML Engineering  
Mayo Clinic Center for Digital Health

### TECHNICAL EXPERTISE

Medical AI/ML

FDA/CE Regulatory

Diagnostic Biosensors

Clinical Translation

Health Equity

SaMD Governance

### KEY CONTRIBUTIONS

**Regulatory & Oversight**

- AI/ML Expert IRB Reviewer
- FDA/CE-Mark Approvals
- Medical & Scientific Affairs

**Technical Achievements**

- Hematology AI Systems
- Biosensor Platforms
- Point-of-Care Diagnostics

### IMPACT METRICS

EXPERIENCE **15+ Years**

CITATIONS **1300+**

PATENTS **Diagnostics, Regulatory & Mechanical Systems**

### CAREER IMPACT

Biosensor Development  
*Label-free detection platforms*

AI/ML in Clinical Practice  
*FDA-approved diagnostic algorithms*

Healthcare Governance  
*Frameworks for responsible AI deployment*

Regulatory Innovation  
*Automated oversight systems & patents*

Health Equity Research  
*Bias mitigation in diagnostic AI*

ENGINEERING

BIOMEDICAL

RESEARCH

HEALTHCARE

INDUSTRY

BIOSENSORS

RIT

UNIVERSITY of ROCHESTER

STANFORD

MAYO CLINIC

Abbott

A

## Traditional Research



VS

## AI Research



AI research **starts with human data** - there's no pre-human exploration phase



# Phases of AI Clinical Evaluation

## Exploratory/Discovery

*Pre-clinical*

- Identify study aims & hypotheses
- Literature searches
- Secondary data analysis
- Professional society guidelines
- Initial algorithm development

## Pilot/Validation

*Early feasibility*

- Preliminary safety assessment
- Performance evaluation
- Small study cohorts
- Analytical validation
- Clinical validation metrics

## Intervention/Treatment

*Clinical efficacy*

- Confirms clinical efficacy
- Safety & risk assessment
- Impacts patient care
- Pivotal trials
- Post-market surveillance

### Analytical Validation

Accuracy, Reliability, Precision

### Clinical Validation

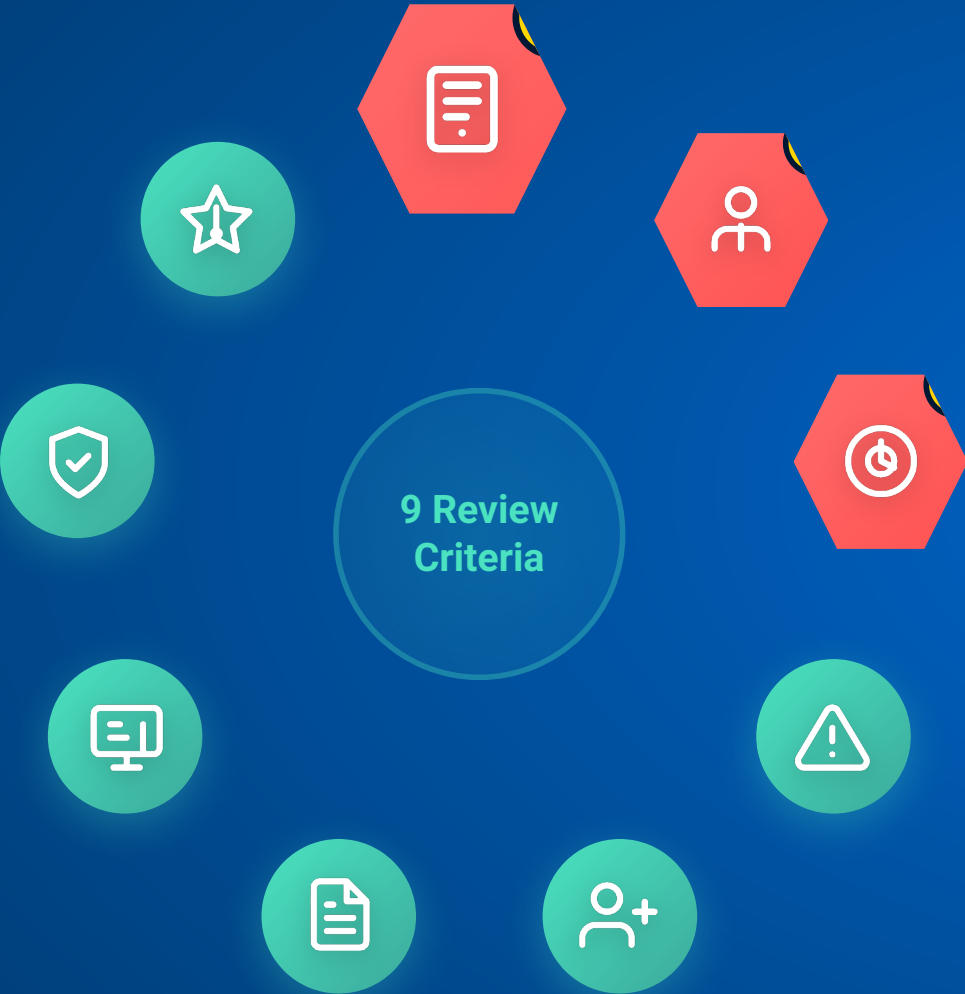
Sensitivity, Specificity, PPV/NPV

### Performance Monitoring

Long-term effectiveness & safety

AI requires **three distinct phases** - each with unique validation requirements

- All Criteria (9)
- Research Definition (3)
- Risk Management (2)
- Subject Protection (4)



FDA-Regulated Device?

- 21 CFR 812, 820
- Policy
- FDA Guidance
- ISO Standards

Software intended for diagnosis, treatment, prevention, or mitigation of disease with quality system requirements

IRB MUST DETERMINE

Compliance with FDA Quality System Regulation (21 CFR 820) design controls, including verification and validation activities per 21 CFR 820.30(g), software validation per 21 CFR 820.70(i), and determination of significant risk vs non-significant risk device classification under 21 CFR 812.

# What Both Sides Want

## IRB SPECIALISTS

### Speaking Regulatory

"Per 45 CFR 46.111(a)(2), risks must be reasonable in relation to anticipated benefits..."

Need: Complete documentation demonstrating compliance with Common Rule, FDA regulations, and institutional policies

## AI RESEARCHERS

### Speaking Technical

"Our CNN achieves 94% AUC with cross-validation on the hold-out test set..."

Want: Clear guidance on what's needed without having to learn regulatory frameworks

TRANSLATION GAP

**The Challenge:** How do we bridge this gap without requiring either side to become experts in the other's domain?



The same technology described differently triggers different regulatory pathways

### ⚠ Clinical Trigger Words

⚠ diagnose   ⚠ treat   ⚠ screen   ⚠ detect   ⚠ clinical decision  
⚠ deployed   ⚠ patient care

### ✓ Research Indicators

✓ investigate   ✓ analyze   ✓ explore   ✓ retrospective   ✓ patterns  
✓ correlations   ✓ hypotheses

#### STUDY OBJECTIVES

The primary objective is to develop a deep learning model that will ⚠ diagnose pneumonia from chest X-rays with accuracy comparable to board-certified radiologists. The system will be ⚠ deployed in emergency departments to ⚠ expedite treatment decisions for patients with respiratory symptoms.

⚠ CLINICAL INTENT - FDA PATHWAY LIKELY

#### RESEARCH AIMS

This study will ✓ investigate novel biomarkers in retinal images by ✓ analyzing patterns across large retrospective datasets. We aim to ✓ explore correlations between retinal features and cardiovascular risk factors to ✓ generate hypotheses for future prospective studies.

✓ EXPLORATORY RESEARCH - IRB FOCUS ON DATA PRIVACY

#### METHODOLOGY

Our AI platform will ✓ analyze patient data to ⚠ predict 30-day readmission risk. Results will be ? provided to clinicians for ✓ evaluation of potential ⚠ treatment modifications. The system will ? support clinical teams in ? decision-making processes.

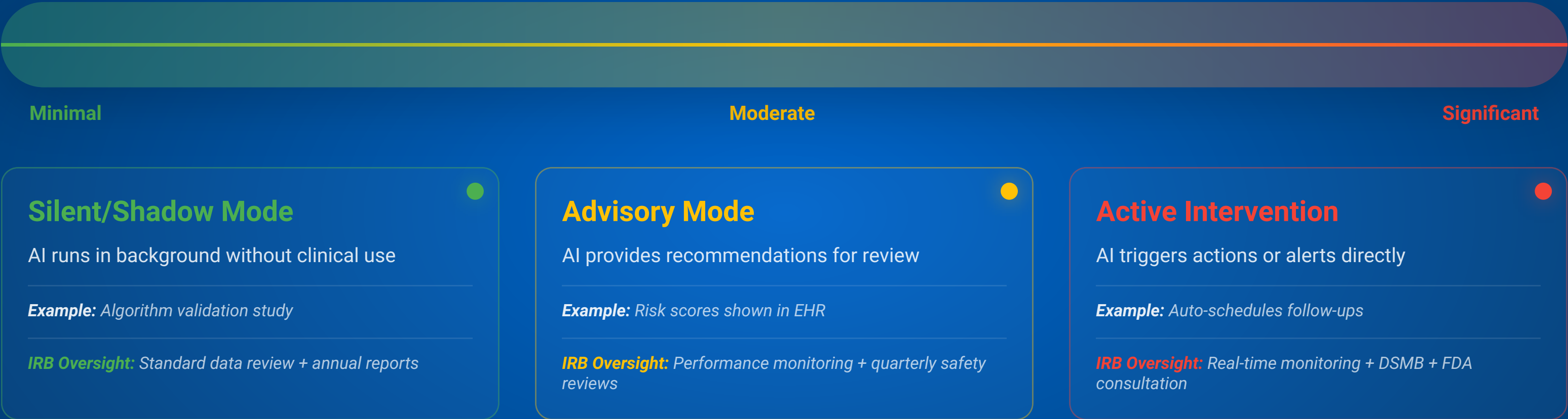
? MIXED SIGNALS - CONTAINS BOTH CLINICAL AND RESEARCH LANGUAGE

#### DATA ANALYSIS PLAN

The AI algorithm will ⚠ screen mammography images to ⚠ identify suspicious lesions requiring immediate radiologist review. Cases flagged as high-risk will ⚠ trigger automatic scheduling for follow-up imaging, with the AI assessment ⚠ documented in patient records.

⚠ CLINICAL INTENT - DIRECT PATIENT IMPACT

Patient impact level determines IRB oversight requirements



Oversight Intensity: Proportional to the degree of patient care impact



RISK CONTROL PRIORITIZATION MATRIX

Q2: Patient Impact	Direct Care	<div>Enhanced Review</div> <div>Medium-High Risk</div> <div>Exploratory with direct impact</div>	<div>Full IRB</div> <div>High Risk</div> <div>Translational affecting patients</div>	<div>Full IRB + Monitoring</div> <div>Critical Risk</div> <div>Clinical deployment to patients</div>
	Indirect	<div>Standard Review</div> <div>Low-Medium Risk</div> <div>Exploratory with indirect effects</div>	<div>Enhanced Review</div> <div>Medium Risk</div> <div>Translational validation phase</div>	<div>Full IRB</div> <div>Medium-High Risk</div> <div>Clinical decision support</div>
	No Impact	<div>Minimal</div> <div>Low Risk</div> <div>Pure research, no patient involvement</div>	<div>Standard Review</div> <div>Low-Medium Risk</div> <div>Algorithm development phase</div>	<div>Enhanced Review</div> <div>Medium Risk</div> <div>Clinical readiness testing</div>
		Exploratory	Translational	Clinical

Q1: Clinical Intent

Q3 Technology Characteristics

Select technology features to determine appropriate controls

ML/Predictive Models

☐ Autonomous operation  
Operates without human input

☐ Clinical decision influence  
Outputs used in care decisions

☐ Continuous learning  
Updates from new data

☐ Age extremes included  
Users <10 or >80 years

☐ Patient data usage  
Processes medical records

Complexity Score 0 / 17

Based on ISO 14971 Risk Management