Core Training Workshop

Hands-On Training on Artificial Intelligence

Understanding AI Research Oversight in Healthcare

Mark Lifson, PhD
AI Research & Ethics Specialist

Core Training Workshop Series

2025

AI vs Traditional Research

Clinical Intent

Risk Assessment

Oversight Framework

Speaker Introduction
ML
Impact Metrics
Experience
15+ Years
Citations
1300+

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

Career Impact
Biosensor Development
Label-free detection platforms
AI/ML in Clinical Practice
FDA-approved diagnostic algorithms
Healthcare Governance
Frameworks for responsible AI
Regulatory Innovation
Automated oversight systems

What Makes AI Different? Data
Traditional Research
1
Lab bench
2
Animals
3
IRB starts
4
Humans
5
Trials
VS
AI Research
1
Human data
2
Training
3
Validation
4
Testing
?
IRB?
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
Mark Lifson, PhD
← → Navigate
2025 - Core Training Workshop

9 Criteria, 1 Review

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

The Translation Gap

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

Translation Gap

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

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

Determining Clinical Intent

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 toexpedite 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

Minimal
Moderate
Significant

Silent/Shadow Mode

AI runs in background without clinical use
Example: Algorithm validation study

IRB Oversight: Standard data review + annual reports

Advisory Mode

AI provides recommendations for review
Example: Risk scores shown in EHR

IRB Oversight: Performance monitoring + quarterly safety reviews

Active Intervention

AI triggers actions or alerts directly
Example: Auto-schedules follow-ups

IRB Oversight: Real-time monitoring + DSMB + FDA consultation
Oversight Intensity: Proportional to the degree of patient care impact

Risk Control Prioritization Matrix

Direct Care	Enhanced Review Medium-High Risk Exploratory with direct impact Potential Risk Controls Full IRB
Indirect	High Risk Translational affecting patients Potential Risk Controls Full IRB + Monitoring
	Critical Risk Clinical deployment to patients Potential Risk Controls Standard Review
No Impact	Low-Medium Risk Exploratory with indirect effects Potential Risk Controls Enhanced Review
	Medium Risk Translational validation phase Potential Risk Controls Full IRB
	Medium-High Risk Clinical decision support Potential Risk Controls Minimal
	Low Risk Pure research, no patient involvement Potential Risk Controls Standard Review

Based on ISO 14971 Risk Management