RADIOLOGY LEADERSHIP INSTITUTE
AMERICAN COLLEGE OF RADIOLOGY

RLI
Power Hour

Experience and Lessons Learned in AI Product Development, Validation, and Clinical Implementation

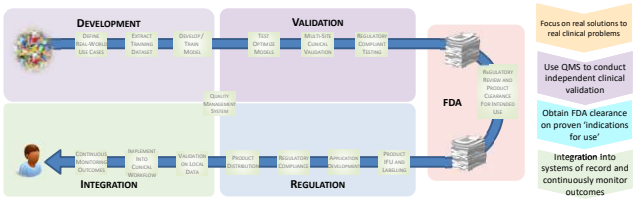
Bernardo C. Bizzo, MD, PhD

October 3, 2023

1

AI Product Journey

4 key phases and 15 process steps to impact patient care



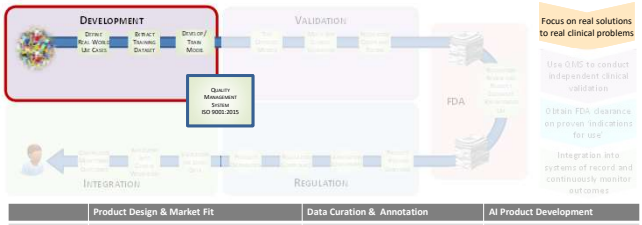
Through investment and support of industry partners, MGB has built an AI ecosystem that supports all aspects of the AI product lifecycle for our institution and collaborators

RLI Power Hour Webinar Series

2

Differentiated Product Development Capabilities

Emphasis on Product Design and Quality Considerations



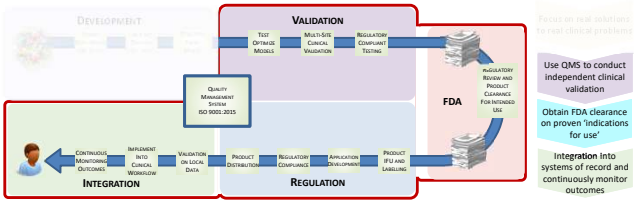
	Product Design & Market Fit	Data Curation & Annotation	AI Product Development
Core Assets & Capabilities	• Access to clinical ecosystem and sub-specialty expertise • Seasoned product managers • Quality Management System (ISO 9001:2015)	• Clinical data assets & infrastructure • Expertise in workflow management and tooling • Quality Management System	• ML model development & engineering • GPU computing infrastructure • Quality Management System

RLI Power Hour Webinar Series

3

CRO Activities Tailored to Meeting Demands of Evolving Market

Services to Support Pre- and Post-FDA Clearance Activities



	Pre FDA clearance		Post FDA clearance
Core Activities	**Regulatory & Validation Strategy** *Support planning and design of regulatory submissions, validation studies, and publications to maximize clinical relevance and adoption*	**Performance Assessment & Reader Study** *- Compare AI product outputs to ground truth to assess accuracy* *- Compare how clinicians perform when using or not using AI product to assess efficiency*	**Clinical Impact & Surveillance** *Implement AI products on 'real-world' data from MGB hospitals to assess failure modes, clinical impact, and post-market surveillance*

4

CRO Offerings Focused on Clinical Relevance

Validation Strategy → Product Performance Assessment → Deployment & Monitoring

	Pre FDA clearance		Post FDA clearance
Outcome	Establish connection with portfolio client, enable FDA submission	FDA clearance, enable clinical impact	Market adoption, enable model refinement
Core Activities	**Regulatory & Validation Strategy** *Support planning and design of regulatory submissions, validation studies, and publications to maximize clinical relevance and adoption*	**Performance Assessment & Reader Study** *- Compare AI product outputs to ground truth to assess accuracy* *- Compare how clinicians perform when using or not using AI product to assess efficiency*	**Clinical Impact & Surveillance** *Implement AI products on 'real-world' data from MGB hospitals to assess failure modes, clinical impact, and post-market surveillance*

Offering Details	• Pre-submission (Q-Sub)
	• Pilot and pivotal study design
	• Scientific publication roadmap
	• Clinical implementation strategy
	• Usability protocol
	• PCCP definition
		• Standalone performance assessment
		  ○ Image processing/quantification (Verification & Validation (V&V))
		  ○ CADt (triage)
		  ○ CADe (detection)
		  ○ CADx (diagnosis)
		• Clinical reader study (CADe, CADx)
		• Ground truth / reads by sub-specialty physicians
			• Shadow-mode 'real-world' failure analysis
			• Clinical workflow impact assessment
			• Patient outcome and economic benefit
			• Monitoring & post-market surveillance
			• Monitoring+ with model refinement
			• Usability and human factors testing

Why us?	• 6+ years of end-to-end AI product development and regulatory experience • Access to multiple clinical end users	• Broad network of clinical sub-specialty • Vast multi-site data access • AI focused QMS processes	• Direct access to live clinical data • AI feedback in the clinical workflow • QMS enabled post-market surveillance

RLI Power Hour Webinar Series

5

Clinical Validation Study Types

Standalone Performance Assessment

Assess model accuracy by comparing AI outputs to a ground truth

Clinical Performance Assessment

Assess model impact on physicians by comparing how clinicians perform when using or not using AI product

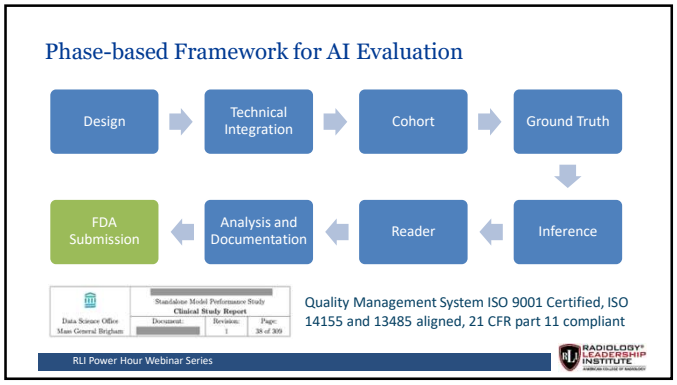
Regulatory pathways	MIMPS (Image Processing), CADt (triage) and CADe/x (detection / diagnostic)	CADe, CADx, CADe/x
Experts	3 sub-specialty MDs (typically 2+1 strategy)	~8-20 MD readers

These CAD device types are general not organ/disease specific

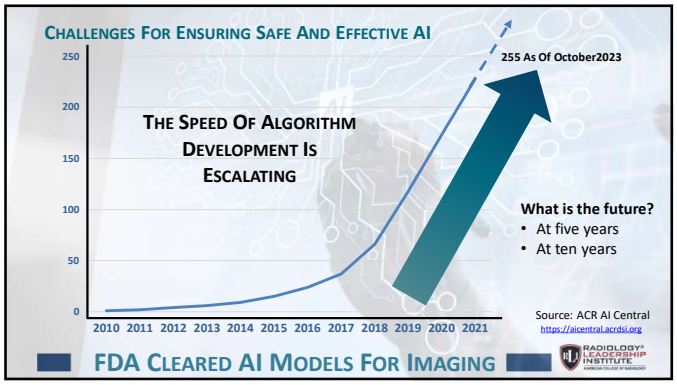
Source: Nicholas Petrick (FDA)

RLI Power Hour Webinar Series

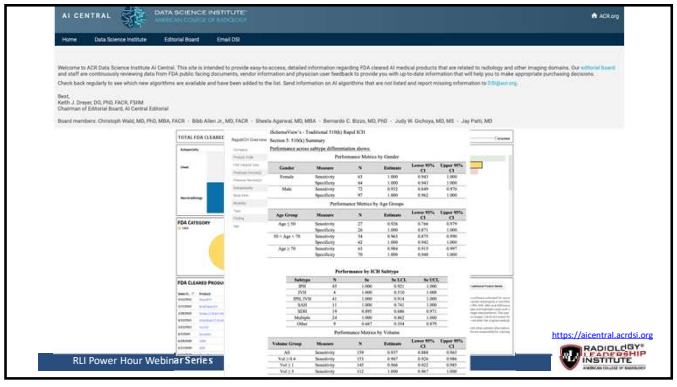
6


Phase-based Framework for AI Evaluation
Design
Technical Integration
Cohort
Ground Truth
Inference
Reader
Analysis and Documentation
FDA Submission
Standalone Model Performance Study
Clinical Study Report
Data Science Office
Mass General Brigham
Document
Revision:
1
Page:
38 of 309
Quality Management System ISO 9001 Certified, ISO 14155 and 13485 aligned, 21 CFR part 11 compliant
RLI Power Hour Webinar Series

7


CHALLENGES FOR ENSURING SAFE AND EFFECTIVE AI
250
200
150
100
50
0
2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
255 As Of October2023
THE SPEED OF ALGORITHM DEVELOPMENT IS ESCALATING
What is the future?
• At five years
• At ten years
Source: ACR AI Central
https://aicentral.acrdsi.org
FDA CLEARED AI MODELS FOR IMAGING

8


AI CENTRAL
DATA SCIENCE INSTITUTE
Home
Data Science Institute
Editorial Board
Email DSI
https://aicentral.acrdsi.org
RLI Power Hour Webinar Series

9

10

MGB AI Imaging Governance Committee

Overview

Mission / Objective

The **MGB Imaging AI Governance Committee** oversees clinical AI deployment, ensuring safe and effective use, with the overarching purpose of improving clinical workflows, improving outcomes and empowering further research

Governance Structure & Cadence

- Committee **led by the Radiology departments**
  - Presence of other clinical departments
- **Representation from various teams across institution** that enable and facilitate the deployment of AI
- Meetings with **monthly cadence for the first 6 months**; followed by a shift to quarterly schedule

Discussion Topics

- **Setting policy** and protocol
- Prioritize and approve **new algorithm deployment**
- **Monitoring algorithm performance** and potentially revoke use based on clinical risk assessment
- Potential opportunities for **scientific impact**
- Potential opportunities for **collaboration with commercial entities**

RLI Power Hour Webinar Series

11

Deployment Intake Process

Deployment request

Deployment request → Request assessment → Governance Committee Assessment → Deployment Initiation

- **Type of deployment**
  - MGB homegrown
  - Commercial product
- **Program type**
  - Research ("shadow-mode")
  - **Clinical deployment**
  - **Clinical/research deployment**

Mass General Brigham

Guidance for Deploying Medical Image-based Software – including Algorithms – in the Clinical Setting

- **Deployment method**
  - On premise
  - Cloud-based
- **Data routing requirements**
  - Full imaging studies
  - Series/images
  - EHR data
- **Results endpoints**
  - AI Repository
  - Text
    - Reporting tool (CARDS / ACR Assist)
    - Notification, worklist prioritization, EHR
  - Overlay/image/series
    - Research or Clinical Imaging viewer

RLI Power Hour Webinar Series

12

Deployment Intake Process

Request assessment

Deployment request → Request assessment and scoring → Governance Committee Assessment → Deployment Initiation

Technical requirements

- Interview with AI vendor or lab for technical deep dive
- Technical Approach Plan to assess implementation complexity and resources
- Scope long-term maintenance and support needs

Clinical safety and benefits

- Clinical sponsor to justify patient safety and clinical workflow improvements
- Comparison with other software currently available within MGB or in the marketplace (RFP)
- Clinical validation plan with pre-defined performance thresholds

Business aspects

- Number of scans, users, departments, and sites impacted
- Alignment with MGB Radiology mission and overall system priorities
- Cost savings, reimbursement, and/or opportunities to offset deployment costs, license fee, and maintenance costs

RLI Power Hour Webinar Series

RADIOLOGY LEADERSHIP INSTITUTE

13

AI Deployment Types

Research "Shadow-Mode" Deployment: without impacting workflow

Phase 1: Algorithm deployment in non-clinical setting (Months 1-3) → Phase 2: Review (Month 3+)

- **Algorithm integration** in 'shadow-mode'
- **Performance assessment** on retrospective cohort, or prospective study

AI Gov Committee to **review summary**

Clinical Deployment: integration into production with a phased approach

Phase 1: Project kick-off after approval (Months 1-3)	Phase 2: Limited clinical deployment (Months 3-6)	Phase 3: Full clinical deployment (Month 6+)
• **Algorithm** in 'shadow-mode' • **Define validation plan** with pre-defined performance metrics • **Pilot performance assessment** on retrospective cohort	AI Gov Committee to **review** and allow to proceed	• **Limited** integration **into production**, user training • **Track** outputs and **clinician feedback**	AI Gov Committee to **review** and allow to proceed	• **Larger scale** model integration **into production,** additional sites • **Dashboard** for ongoing performance tracking	AI Gov Committee to **monitor**

Legend: Key Activity Types — Algorithm Implementation | Governance

RLI Power Hour Webinar Series

RADIOLOGY LEADERSHIP INSTITUTE

14

AI Deployment Types

ARTICLE IN PRESS

ORIGINAL ARTICLE

REPORT

Addressing the Challenges of Implementing Artificial Intelligence Tools in Clinical Practice: Principles From Experience

Bernardo Bizzo, MD, PhD[a,b], Giridhar Dasegowda, MBBS[b,c], Christopher Bridge, PhD[b,c], Ben Miller[b,c], James Hillis, MBBS, DPhil[d], Mannudeep K. Kalra, MD[b,c,e], Kimberly Durniak, PhD[a], Markus Stout[b,c,f], Tom Schultz[b,c,g], Tarik Alkasab, MD, PhD[b,c,h,i], Keith J. Dreyer, DO, PhD[b,c,j,k]

PROOF

Abstract

The multitude of artificial intelligence (AI)-based solutions, vendors, and platforms poses a challenging proposition to an already complex clinical radiology practice. Apart from assessing and ensuring acceptable local performance and workflow fit to improve imaging services, AI tools require multiple stakeholders, including clinical, technical, and financial, who collaborate to move potential deployable applications to full clinical deployment in a structured and efficient manner. Postdeployment monitoring and surveillance of such tools require an infrastructure that ensures proper and safe use. Herein, the authors describe their experience and framework for implementing and supporting the use of AI applications in radiology workflow.

Key Words: Artificial intelligence, machine learning, radiology, implementation, deployment

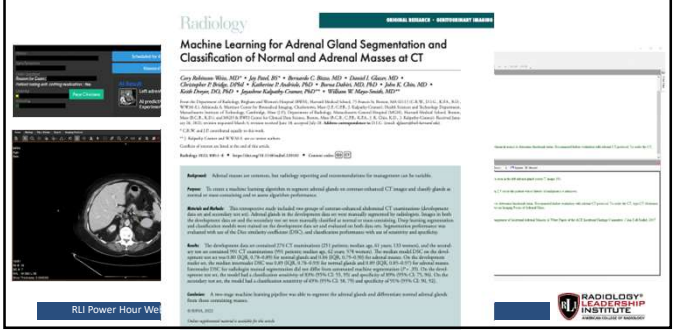
J Am Coll Radiol 2023 ■ ■ ■ Copyright © 2023 American College of Radiology

RLI Power Hour

RADIOLOGY LEADERSHIP INSTITUTE

15

"Limited Clinical Deployment" Evaluation



16

AI Monitoring

Monitoring Framework

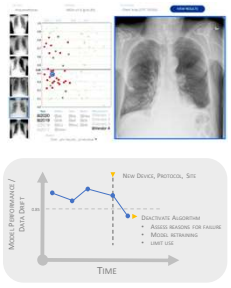
- Identify data drifts based on changes on input data
- Subgroup analyses (demographics, patient information, metadata)
- Alerts for model performance drops below pre-defined metric on specific population subgroups, protocols, devices, and/or sites
- Gatekeeping for multiple AI results without burdening radiologists

Post-market surveillance

- Clear clinical needs given data drift and model brittleness

Healthcare economics and patient outcome impact

- Beyond efficiency/accuracy improvements



RLI Power Hour Webinar Series

17

Thank you!

Questions?
bbizzo@mgh.harvard.edu
http://mgbai.com

RLI Power Hour Webinar Series

18