SIIM ACR Data Science Summit
VIRTUAL 2020

The AI Lifecycle

Evaluating and Monitoring AI Algorithms

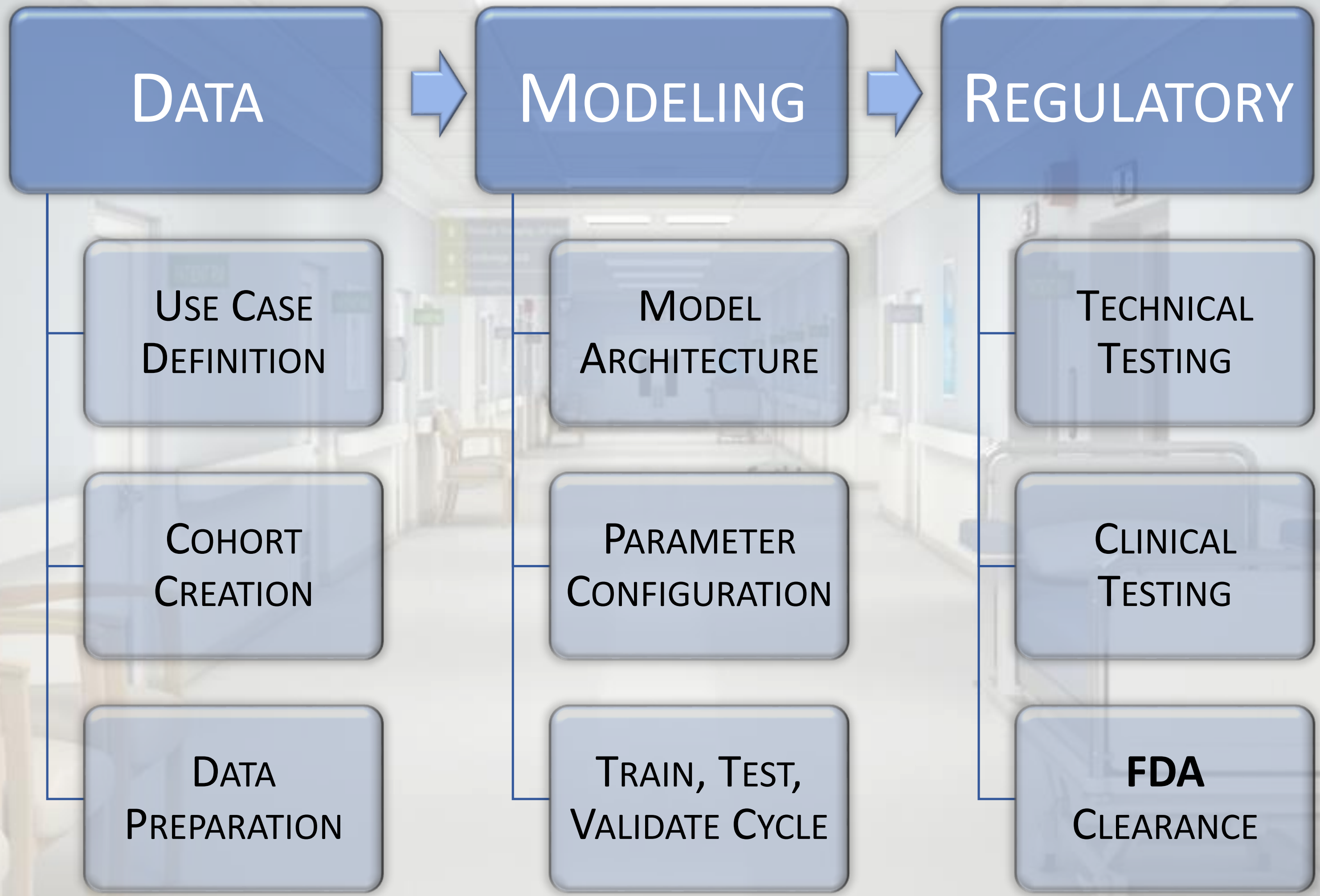
ACR DSI Summit

June, 2020

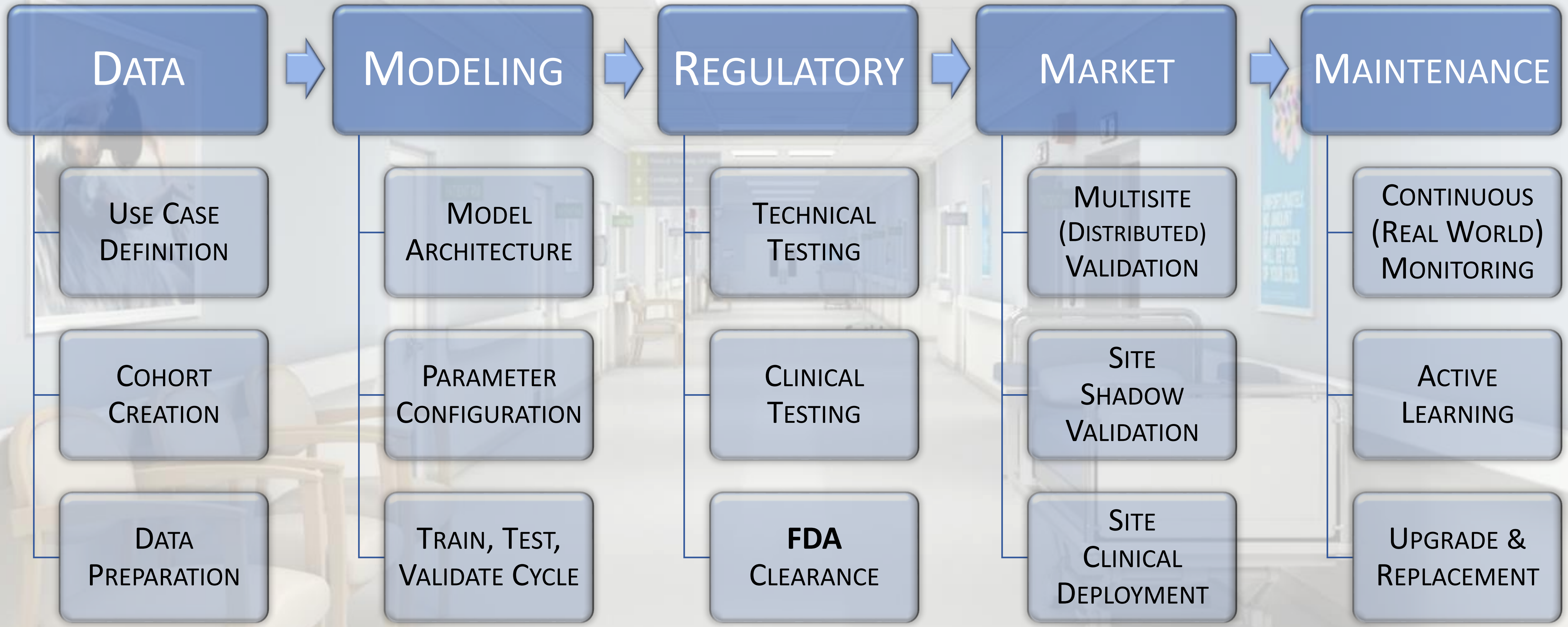
Keith J. Dreyer DO, PhD, FACR
Chief Science Officer, ACR DSI

DATA SCIENCE INSTITUTE™
AMERICAN COLLEGE OF RADIOLOGY

The AI Lifecycle



The AI Lifecycle



The AI Lifecycle

Data → **Modeling** → **Regulatory** → **Market** → **Maintenance**

Data
- Use Case Definition
- Cohort Creation
- Data Preparation

Modeling
- Model Architecture
- Parameter Configuration
- Train, Test, Validate Cycle

Regulatory
- Technical Testing
- Clinical Testing
- FDA Clearance

Market
- Multisite (Distributed) Validation
- Site Shadow Validation
- Site Clinical Deployment

Maintenance
- Continuous (Real World) Monitoring
- Active Learning
- Upgrade & Replacement

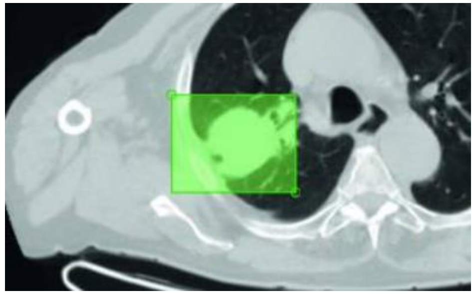
Define Use Case

Custom created:

P(Lung Cancer)=.45

Vs.

Clinical Standard Use Case:



Incidental Pulmonary Nodules on Chest Radiograph

Purpose	Detection and characterization of incidental pulmonary nodules on chest radiographs (CXR). These are nodules that are detected on CXRs performed for other reasons than lung cancer screening.
Tag(s)	
Panel	Thoracic
Define-AI ID	08190004
Originator	Thoracic Panel
Panel Chair	Warren B. Gefter, MD; Eric J. Stern, MD
Panel Reviewers	Thoracic Panel
License	Creative Commons 4.0
Status	Public Comment
RadElement Set	RDES94

Technical Specifications

Inputs

DICOM Study

Procedure	XRAY, Chest
Views	CXR: PA, lateral, AP, apical lordotic, obliques CR, DR, dual-energy, and bone-suppression CXRs
Data Type	DICOM
Modality	XRAY
Body Region	Chest
Anatomic Focus	Lung
Pharmaceutical	N/A
Scenario	N/A

Primary Outputs

Detection of nodule

RadElement ID	RDE564
Definition	The definition of pulmonary nodule detection includes: 1) The center x and y coordinates of a candidate nodule bounding box with reference to the superior and right-most pixel in the bounded area (referencing the patient for sidedness, zero indexed); 2) The dimensions of a bounding box in pixels (x and y); and 3) The probability that the bounded CXR opacity represents a true lung nodule.
Data Type	Numeric

Considerations for Dataset Development

Procedures	CXR, CR, DR, dual-energy, and bone-suppression CXRs
Views	PA, lateral, AP, apical lordotic, obliques
Age	≥ 18 years old
Sex at birth	Male, Female
Nodule Validation	CT within 1 month of CXR. Corresponding nodule location on CXR confirmed by chest radiologists.
Nodule attenuation based on CT confirmation	solid, part-solid, groundglass, internal fat density, calcification, cavitary
Size (in mm)	[5,40]
Shape	round, oval, triangular, lobular, irregular
Margin	smooth, irregular, spiculated
Location	broad sampling of lung regions, apex to base, central to peripheral
Comorbidities	Smokers, non-smokers, COPD, travel/exposure history, other primary malignancy or history of primary malignancy, bronchitis, bronchiolitis, pneumonia, tuberculosis, fungal and other pulmonary infections, focal inflammatory lesions, usual interstitial pneumonia and other diffuse lung diseases, pleural effusion.
Other Considerations	Range of CXR technologies (CR, DR, dual-energy, bone

Clinical Implementation

Value Proposition

Lung cancer, the leading cause of cancer-related deaths in both women and men, frequently presents as a pulmonary nodule on chest radiographs (CXRs) or CT scans. While low-dose CT is utilized for lung cancer screening, chest radiography, being among the most highly utilized diagnostic imaging procedures worldwide, is the most common thoracic imaging study in which incidental lung cancers are discovered. Nonetheless, interpretation of chest radiographs is challenging and prone to many reading errors. Thus nodules are frequently missed on CXRs, with studies showing approximately 20-30% (even up to 90%) seen only in retrospect. The causes for these frequent errors are multifactorial, including: overlapping anatomic structures such as the ribs, clavicles, thoracic spine, pulmonary vessels, heart, mediastinum and diaphragms; errors in visual search, lesion recognition or decision-making; and suboptimal image quality. Small, ill-defined nodules with low attenuation and conspicuity are particularly susceptible to being overlooked. As early detection of lung cancer reduces mortality, missed or delayed diagnosis due to these CXR errors may negatively impact patient survival.

Furthermore, such errors carry significant medicolegal risks, being the second most common cause (after breast cancer) for malpractice litigation in radiology. Algorithms based upon machine learning therefore offer an important use case to assist radiologists in more accurate detection, characterization, and any communication and recommendation for further study of these nodules. This may be particularly true for less experienced radiologists or in places without access to radiology expertise. These algorithms show promise in improving upon traditional CAD (computer-assisted detection) systems.

Narrative(s)

A 45-year-old man with cough and fever has a CXR for evaluation of possible pneumonia. Algorithm evaluates the lungs and detects a non-calcified, irregular-shaped nodule at the right lung apex partially obscured by the anterior first rib. Lesion is highlighted on annotated image, so as not to be overlooked by the radiologist. Radiologist confirms that this is a new finding compared with older CXRs and recommends further evaluation with a chest CT scan. Appropriate communication with the referring clinician is made.

DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

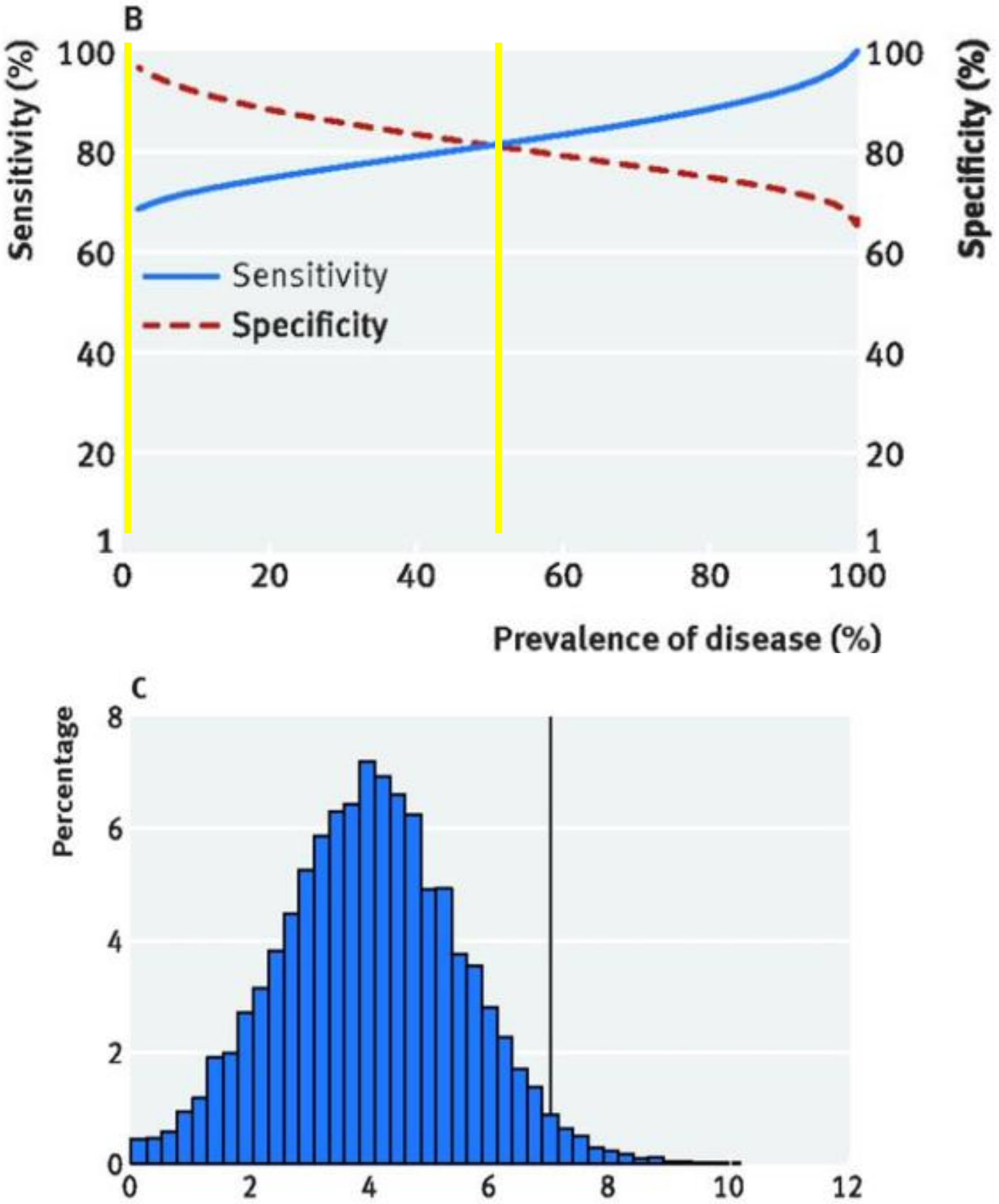
The AI Lifecycle

Create Cohort

- Representative of the target population
  - Disease Prevalence
  - Selection Bias or Spectrum Effect

Create Cohort

- Representative of the target population
  - **<u>Disease Prevalence</u>**
  - Selection Bias or Spectrum Effect



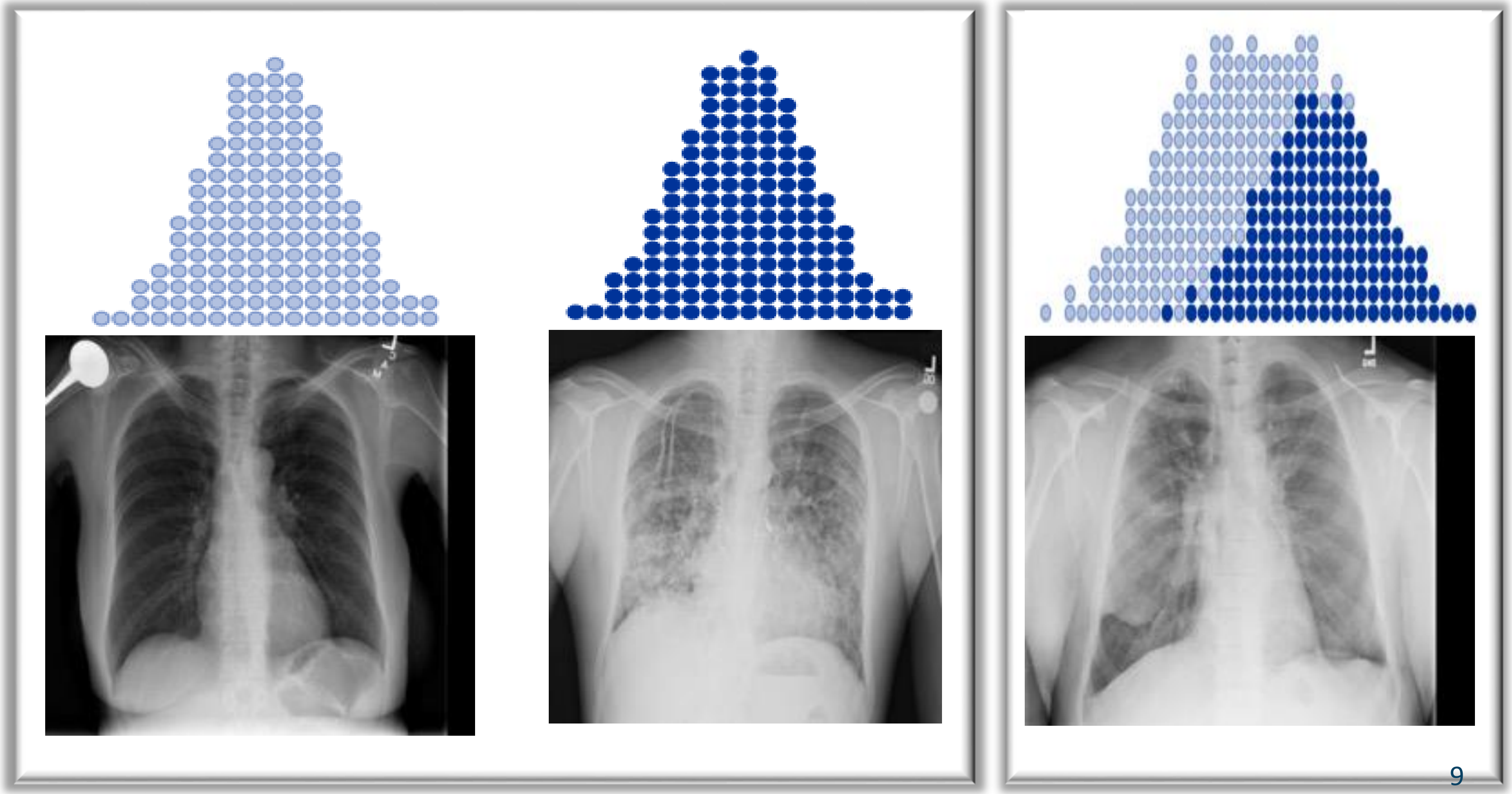
Disease prevalence = 50%
2%

DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

BMJ 2016;353:i3139

Create Cohort

- Representative of the target population
  - Disease Prevalence
  - **Selection Bias or Spectrum Effect**



DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

The AI Lifecycle

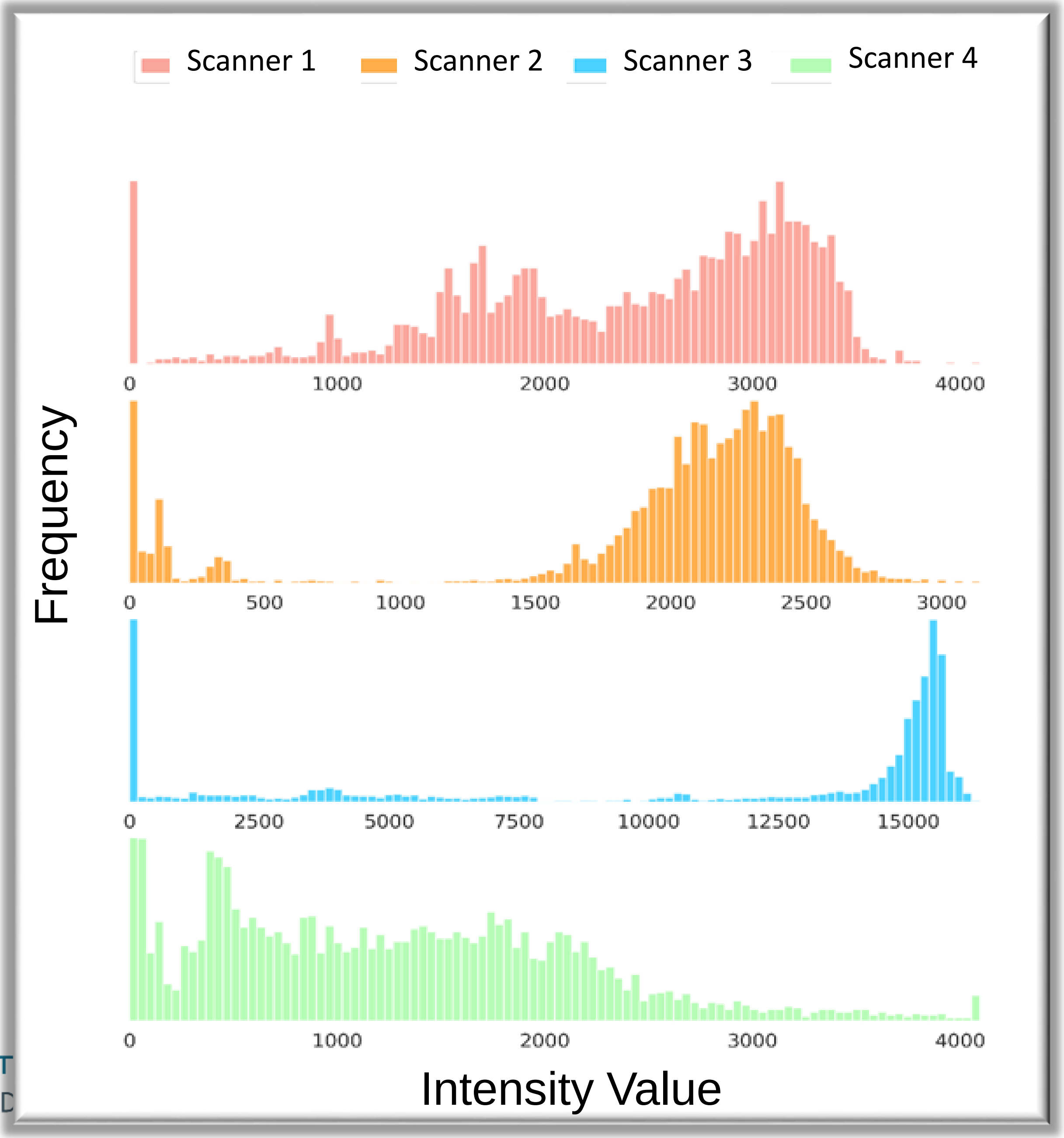
Data
- Use Case Definition
- Cohort Creation
- Data Preparation

Modeling
- Model Architecture
- Parameter Configuration
- Train, Test, Validate Cycle

Regulatory
- Technical Testing
- Clinical Testing
- **FDA** Clearance

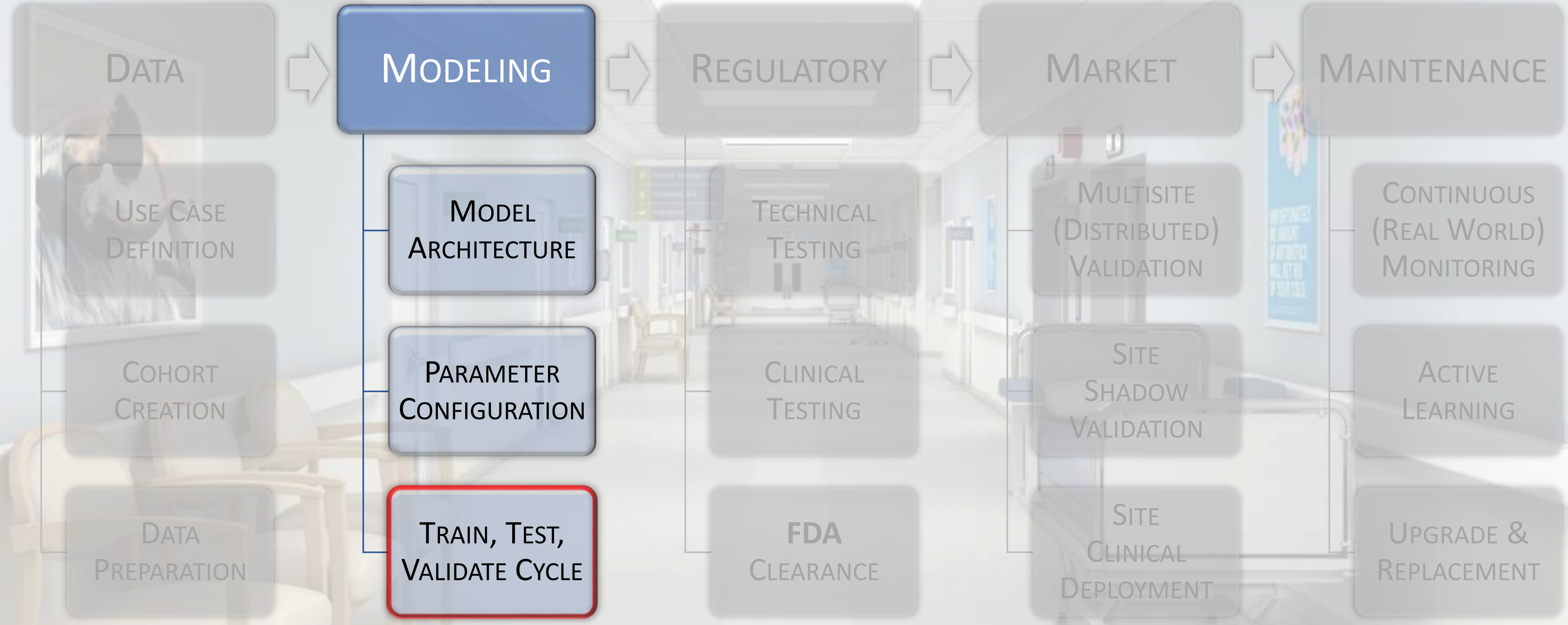
Market
- Multisite (Distributed) Validation
- Site Shadow Validation
- Site Clinical Deployment

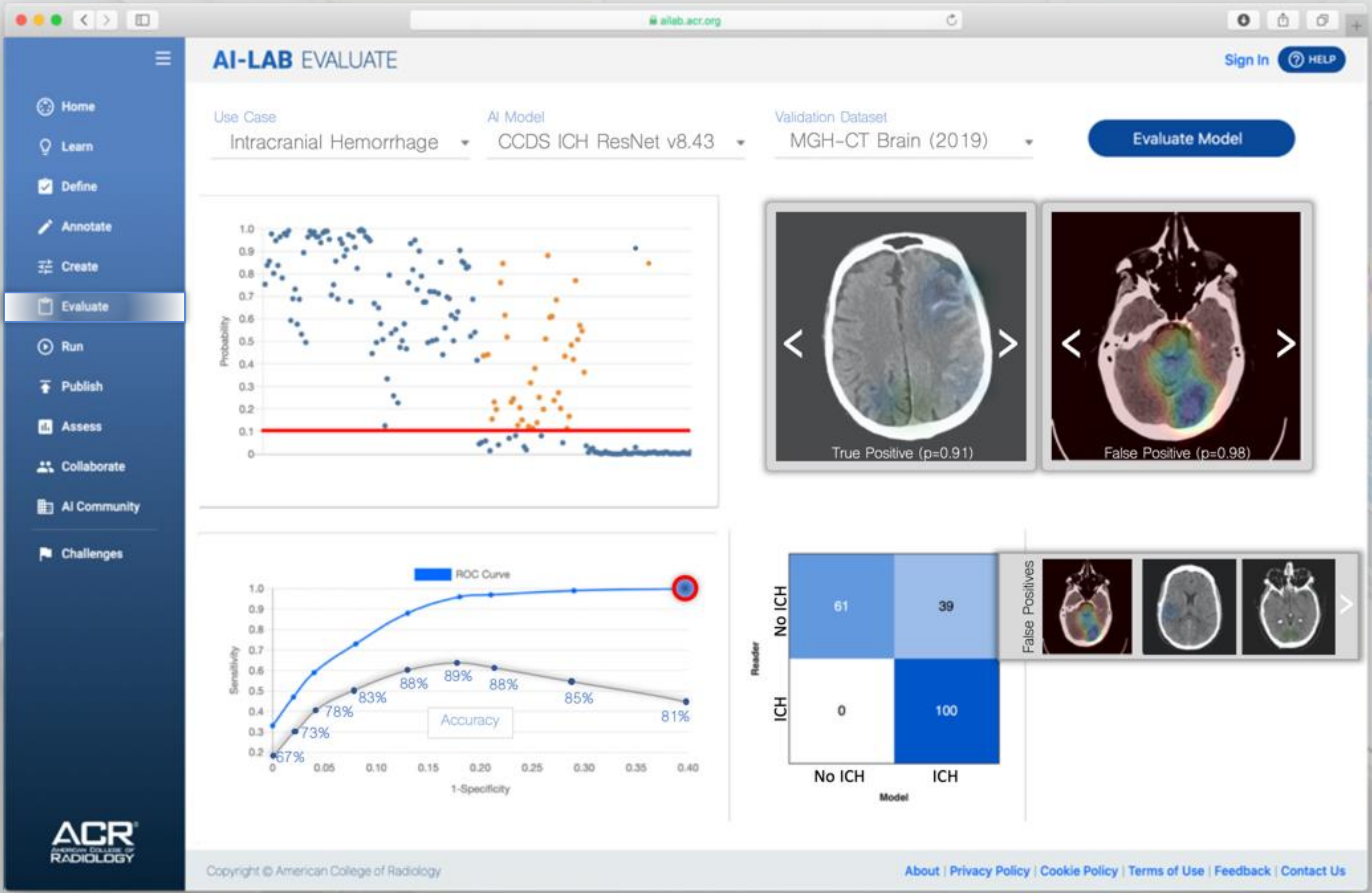
Maintenance
- Continuous (Real World) Monitoring
- Active Learning
- Upgrade & Replacement



DATA SCIENCE INST
AMERICAN COLLEGE OF RAD

THE AI LIFECYCLE





Optimization by Application

↑ Accuracy

Most broadly applicable algorithm

↑ Specificity

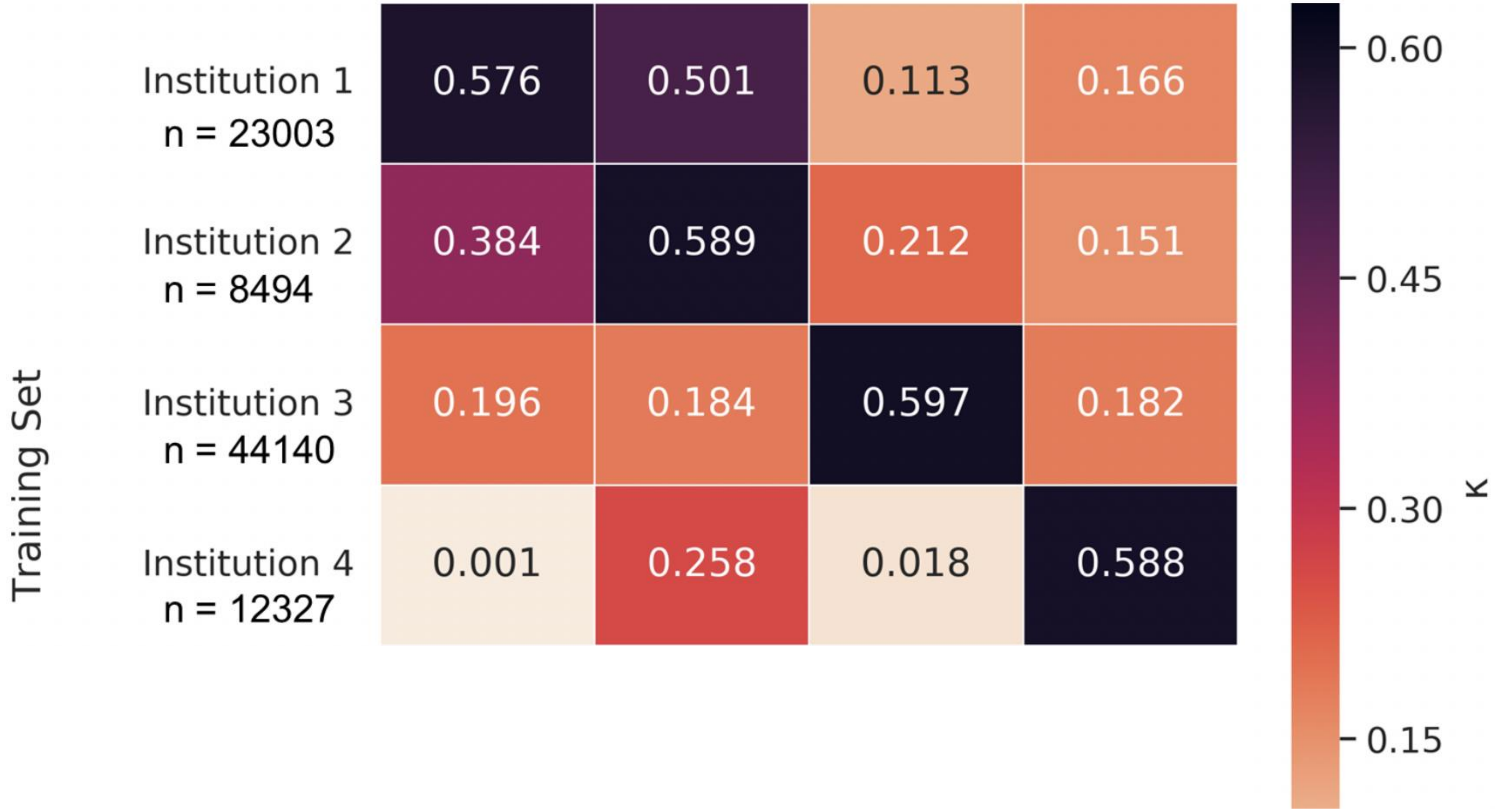
Optimized to promote to an acute state

↑ Sensitivity

Optimized to demote from an acute state

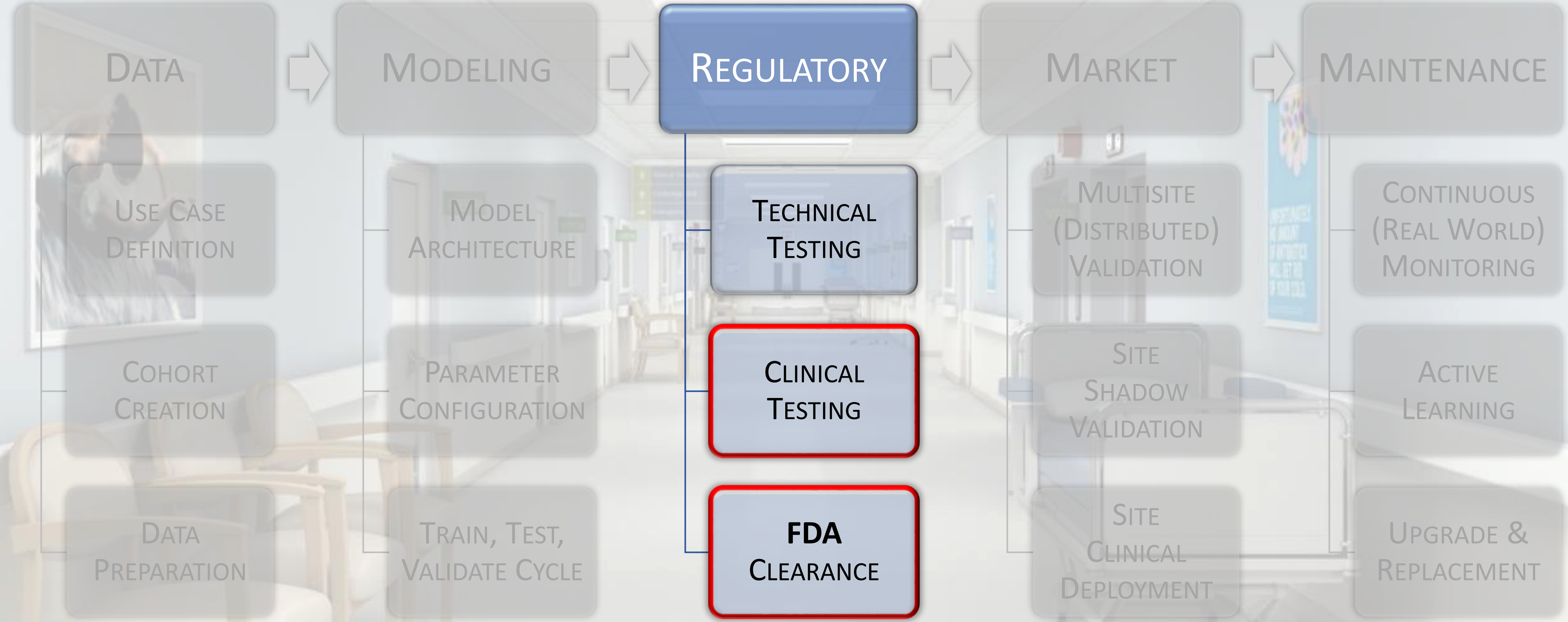
The AI Lifecycle

- Data
  - Use Case Definition
  - Cohort Creation
  - Data Preparation
- Modeling
  - Model Architecture
  - Parameter Configuration
  - Train, Test, Validate Cycle
- Regulatory
  - Technical Testing
  - Clinical Testing
  - **FDA** Clearance
- Market
  - Multisite (Distributed) Validation
  - Site Shadow Validation
  - Site Clinical Deployment
- Maintenance
  - Continuous (Real World) Monitoring
  - Active Learning
  - Upgrade & Replacement

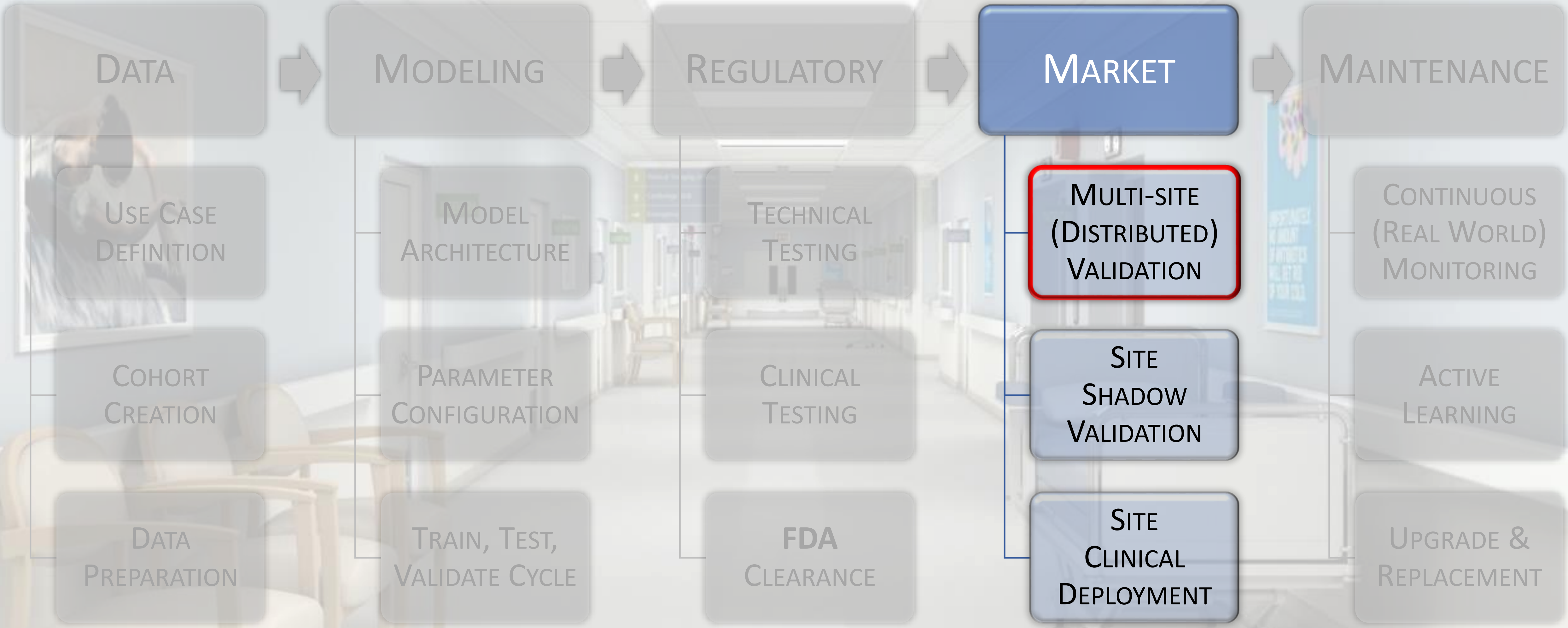


*Centralized model building, federated model shows higher accuracy

The AI Lifecycle

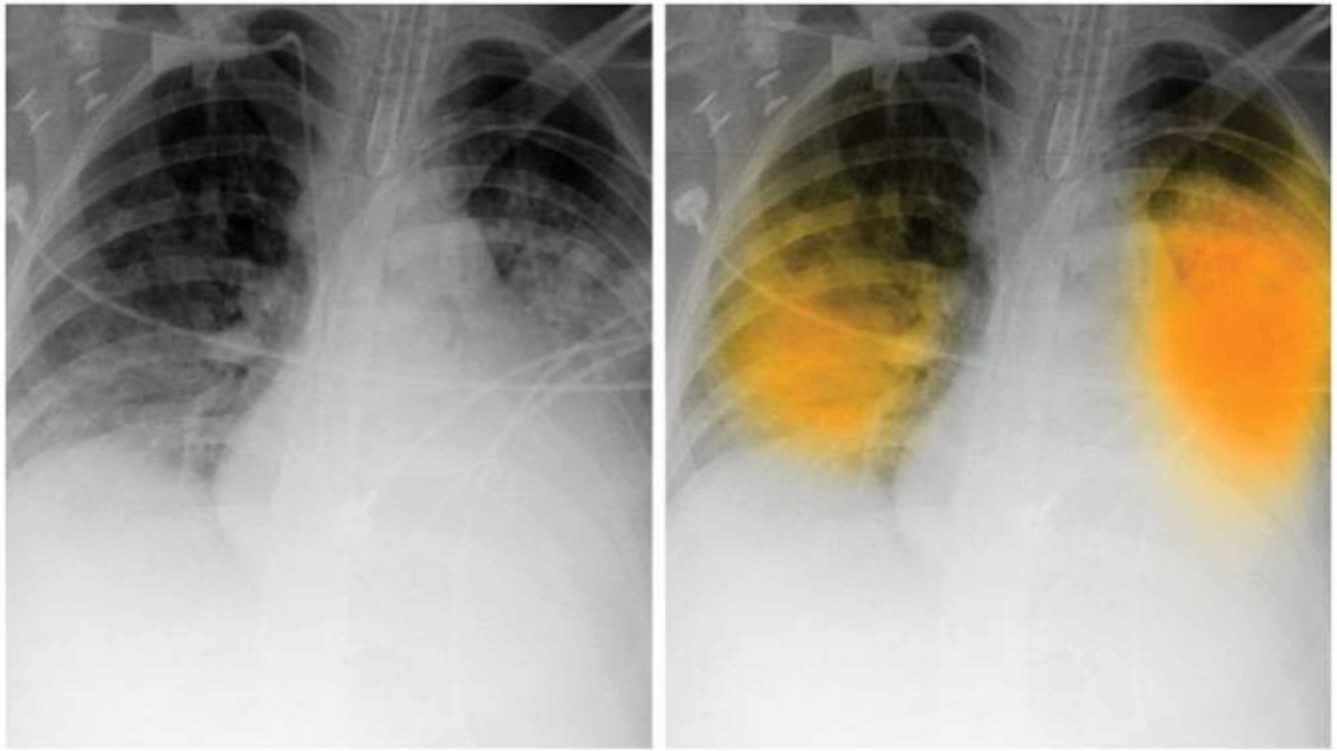


The AI Lifecycle



Algorithms are brittle

- Demographic variables
- Comorbidities
- Technology
- Disease severity



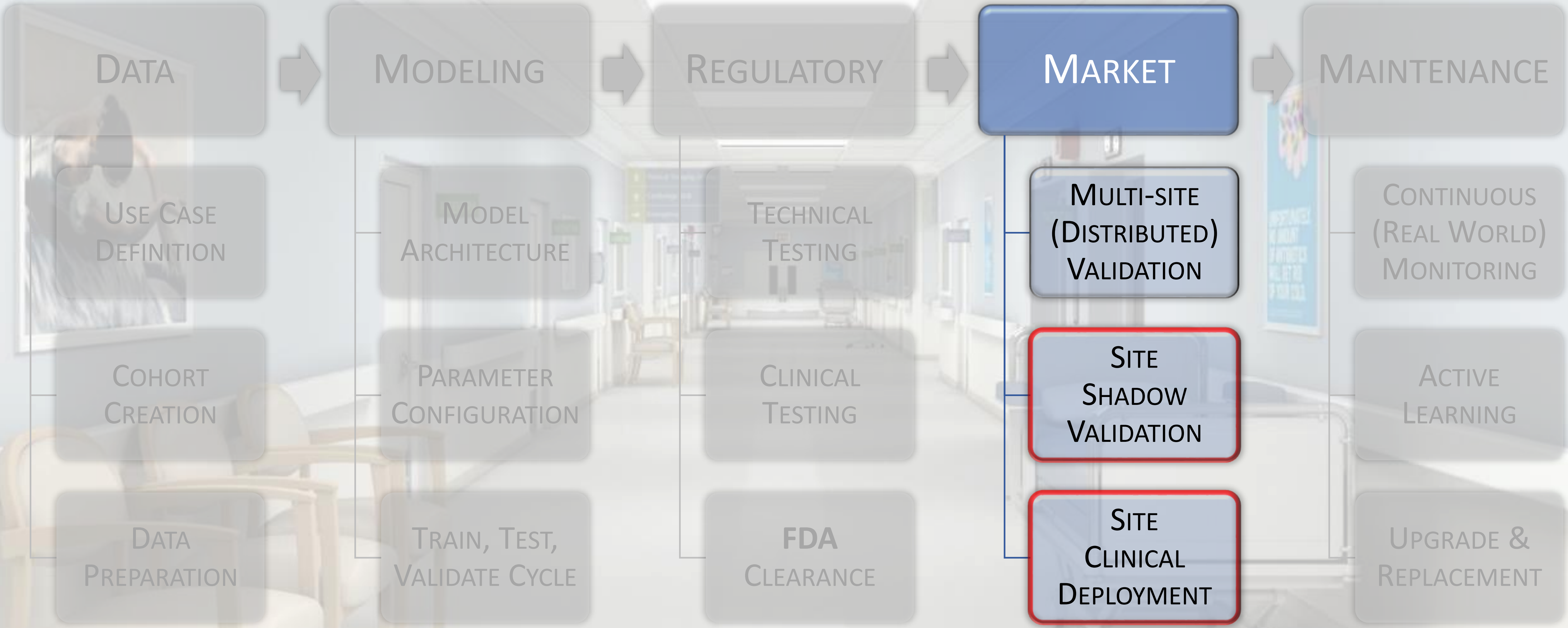
Scientists are developing a multitude of artificial intelligence algorithms to help radiologists, like this one that lights up likely pneumonia in the lungs. ALBERT HSIAO AND BRIAN HURT/UC SAN DIEGO AIDA LABORATORY

Artificial intelligence could revolutionize medical care. But don't trust it to read your x-ray just yet

By Jennifer Couzin-Frankel | Jun. 17, 2019 , 12:45 PM

DATA SCIENCE INSTITUTE®
AMERICAN COLLEGE OF RADIOLOGY

The AI Lifecycle

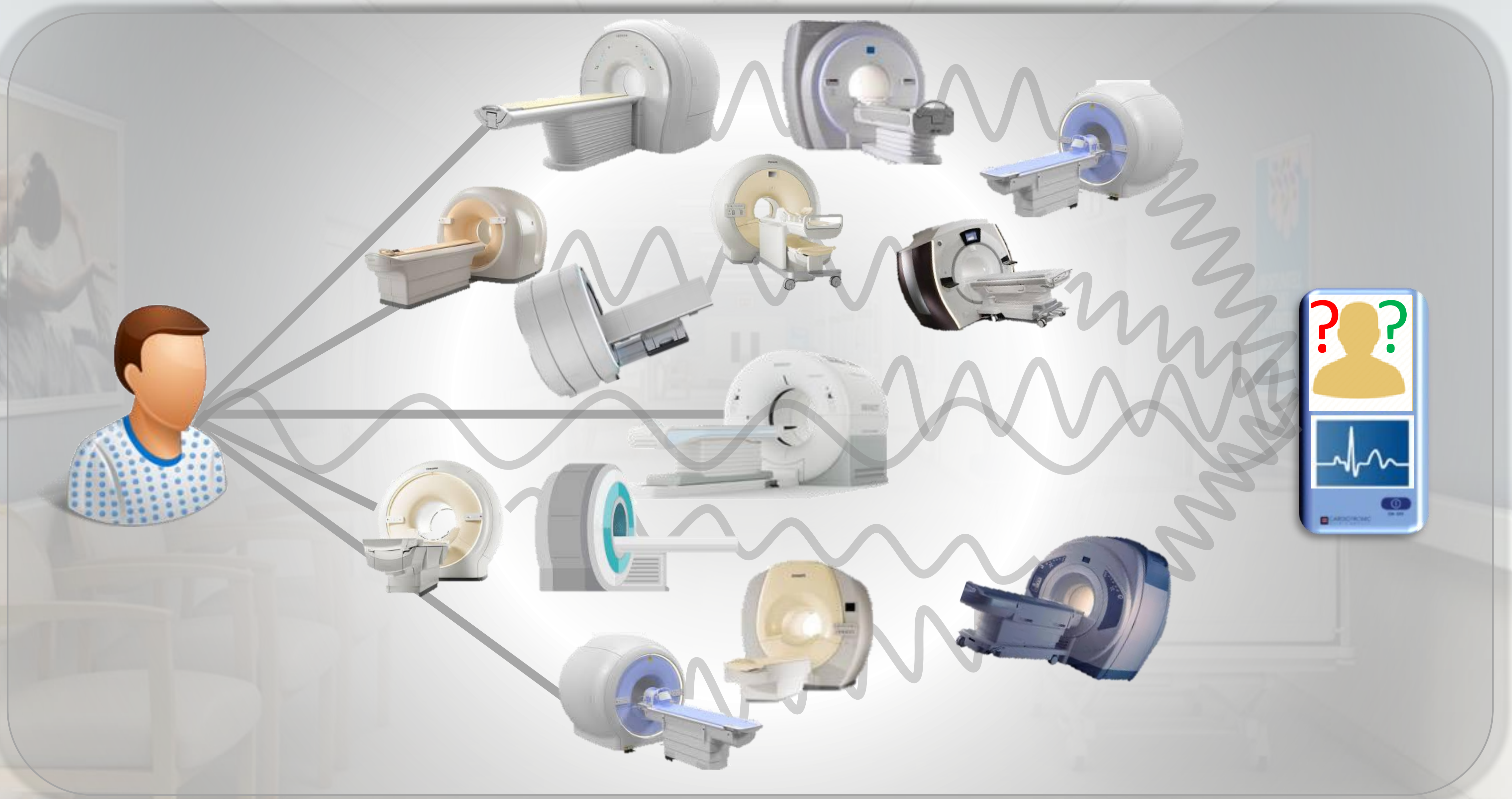


The AI Lifecycle

- **Data**
  - Use Case Definition
  - Cohort Creation
  - Data Preparation
- **Modeling**
  - Model Architecture
  - Parameter Configuration
  - Train, Test, Validate Cycle
- **Regulatory**
  - Technical Testing
  - Clinical Testing
  - **FDA** Clearance
- **Market**
  - Multisite (Distributed) Validation
  - Site Shadow Validation
  - Site Clinical Deployment
- **Maintenance**
  - Continuous (Real World) Monitoring
  - Active Learning
  - Upgrade & Replacement

The AI Lifecycle

Inconsistencies between acquisition devices

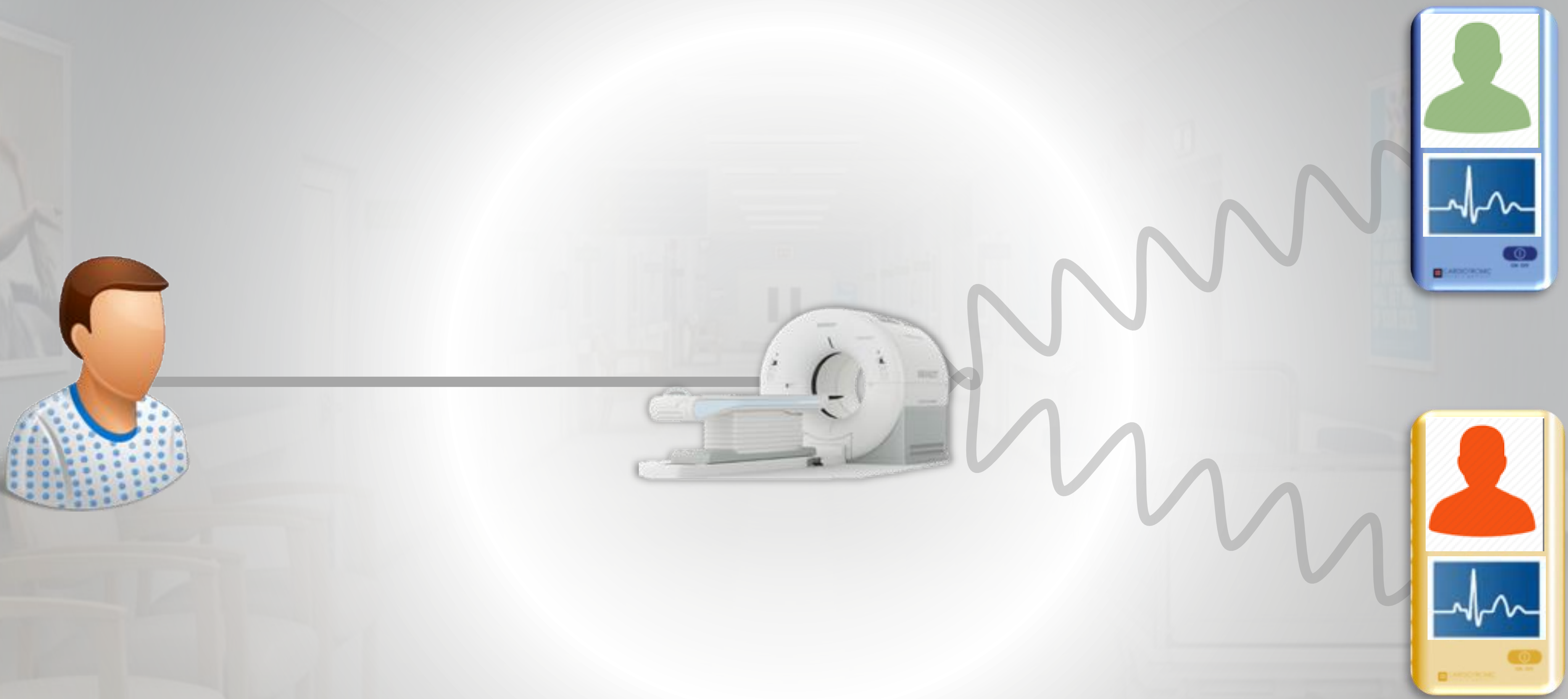


The AI Lifecycle

Inconsistencies between AI algorithms

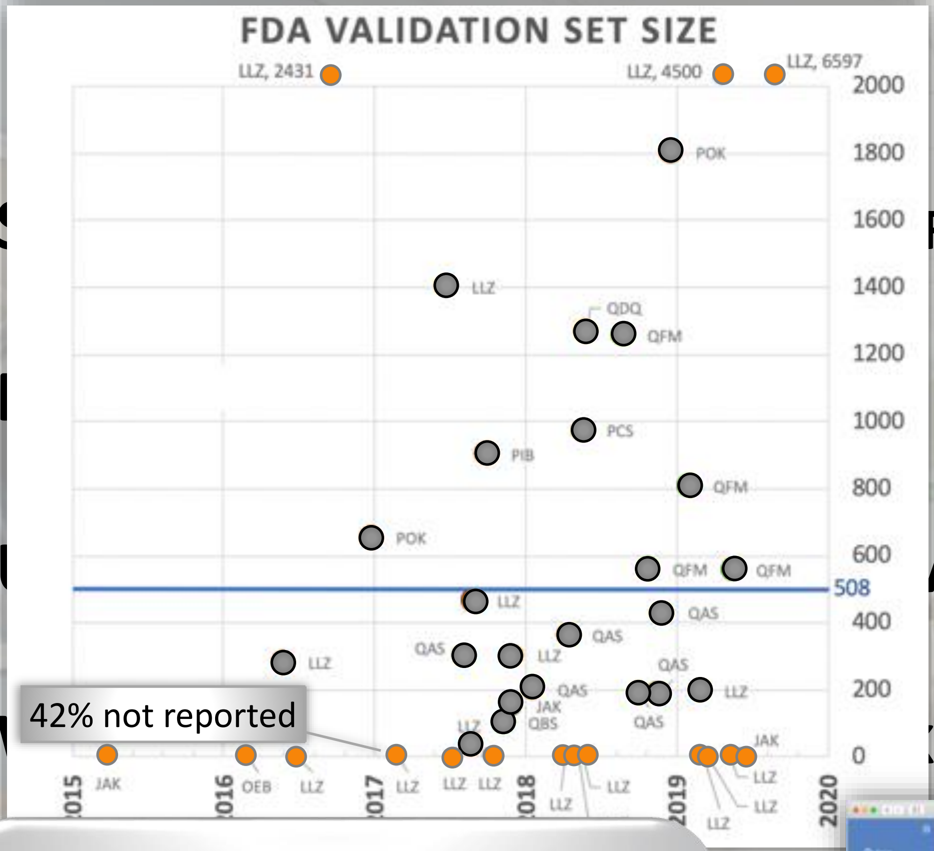
The AI Lifecycle

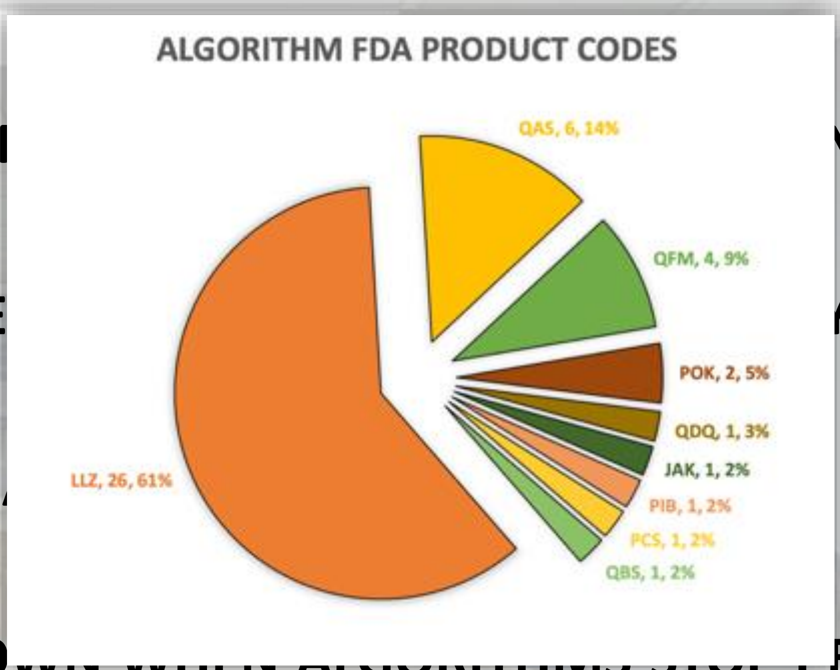
Inconsistencies between AI algorithms




ailab.acr.org
AI-LAB WELCOME
Sign In
HELP
Home
Learn
Define
Annotate
Create
Evaluate
Run
Publish
Assess
Collaborate
AI Community
Challenges
ACR AMERICAN COLLEGE OF RADIOLOGY
Welcome to ACR AI-LAB™
The ACR Data Science Institute has developed the ACR AI-LAB™, a data science toolkit designed to democratize AI by empowering radiologists to develop algorithms at their own institutions, using their own patient data, to meet their own clinical needs.
Learn
Learn how AI applies to imaging through a series of detailed videos.
Start Learning
Define Use Cases
Explore existing use cases for AI in medical imaging, or propose your own idea for a use case.
Learn More
Upcoming Events
Tips and Demos at RSNA
Drop by DSI's RSNA booth #11122, North Hall, to chat with our Resident Mammography Challenge winner David Qian on Tuesday, December 3rd from 12 - 1pm and hear his tips for developing a top-ranked model with AI-LAB. Congrats to David and all of our finalists! Hands-on AI-LAB demos led by ACR staff will be running all week.
See Challenge Rankings
For investigational use only. The performance characteristics of this product have not been established.

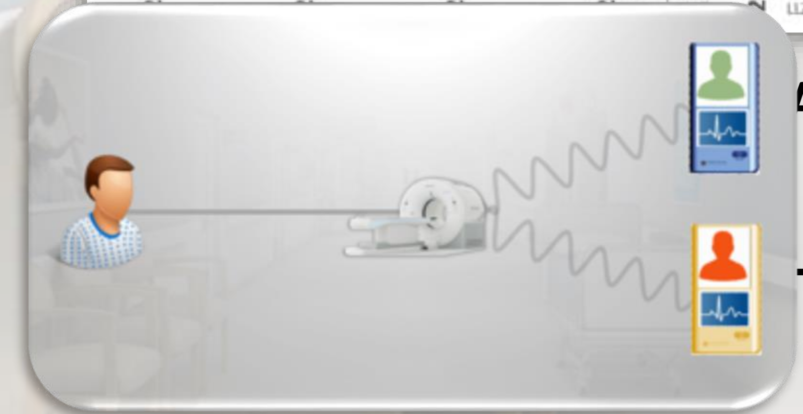
AI Life Cycle Summary

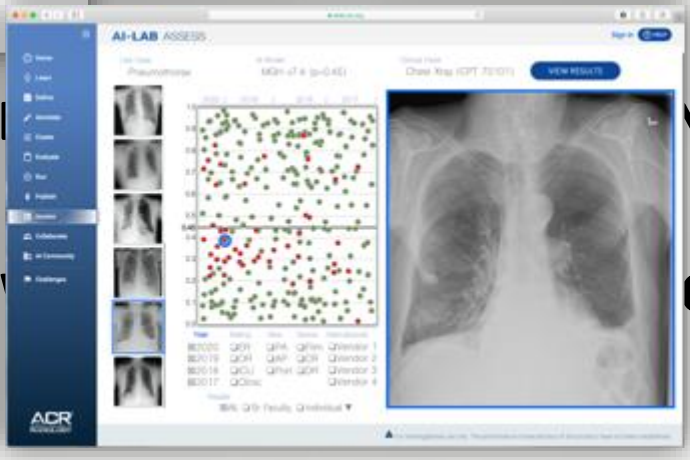


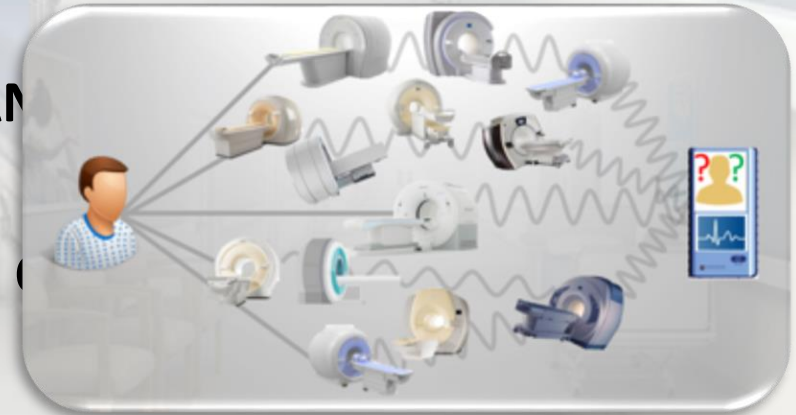


FDA Product Codes

LLZ: PACS
QAS: CAD for Triage
QFM: CAD for Prioritization
QDQ: CAD for Cancer
POK: CAD for Cancer
QBS: CAD for Fractures
LNH: MRI Device
JAK: CT Xray System
OEB: CAD for Lung CT
PCS: Liver Iron Imaging
PIB: Biophotometer







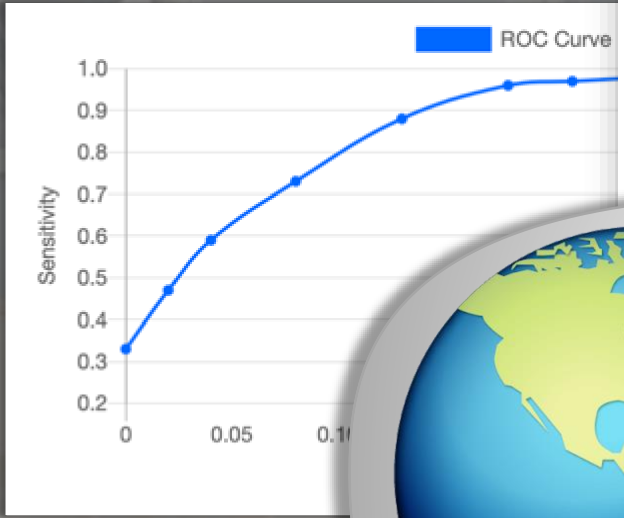
‘Rule-out’ Algorithm Considerations

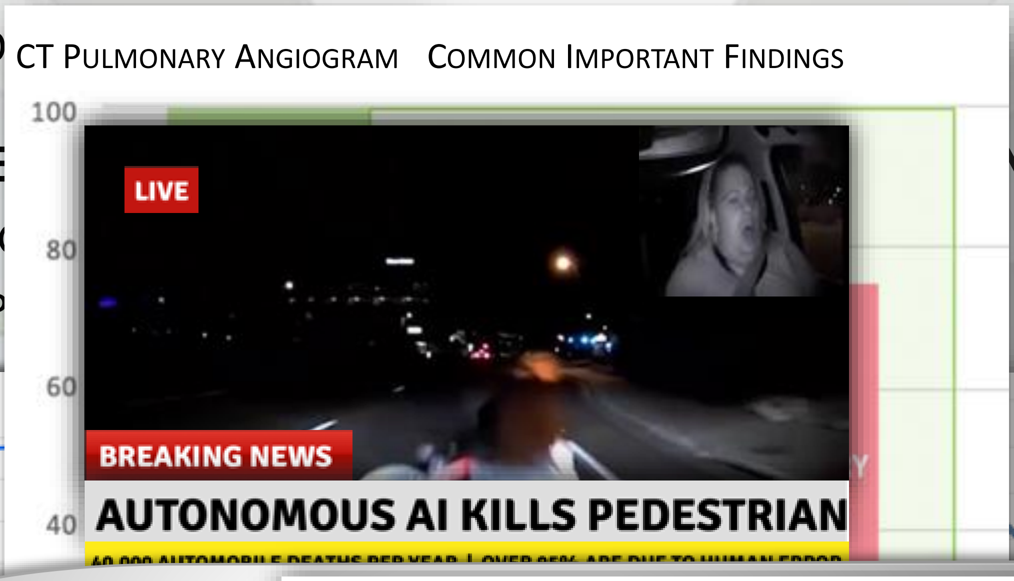
Rule-outs require an extremely high negative predictive value

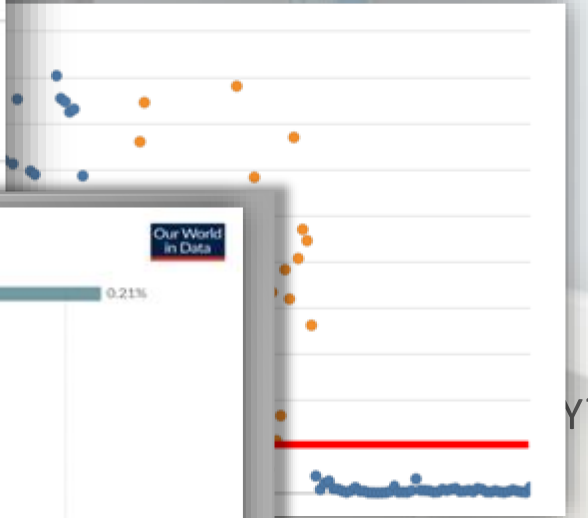
- Special Control: So

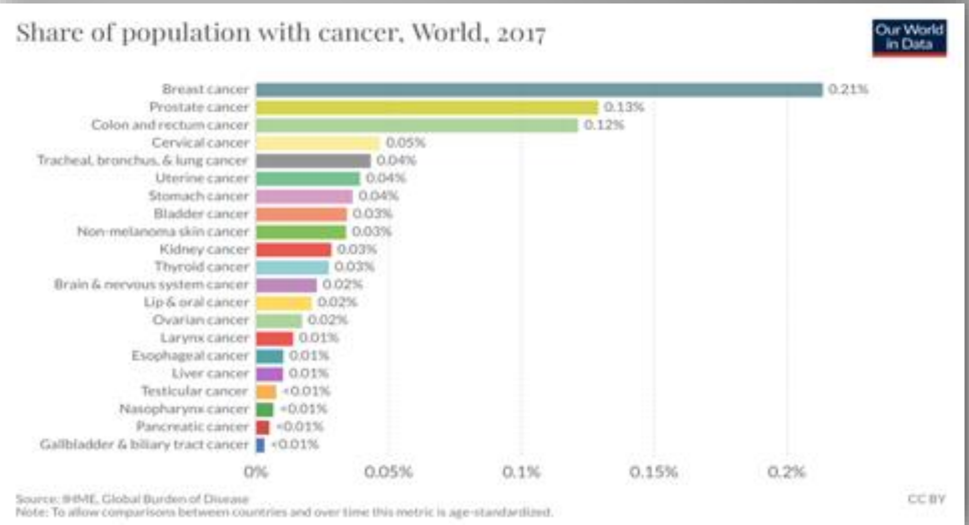
Tests for safety and e… other CAD

- Safe: Each False Ne
- Effective: Positive P









…y?)

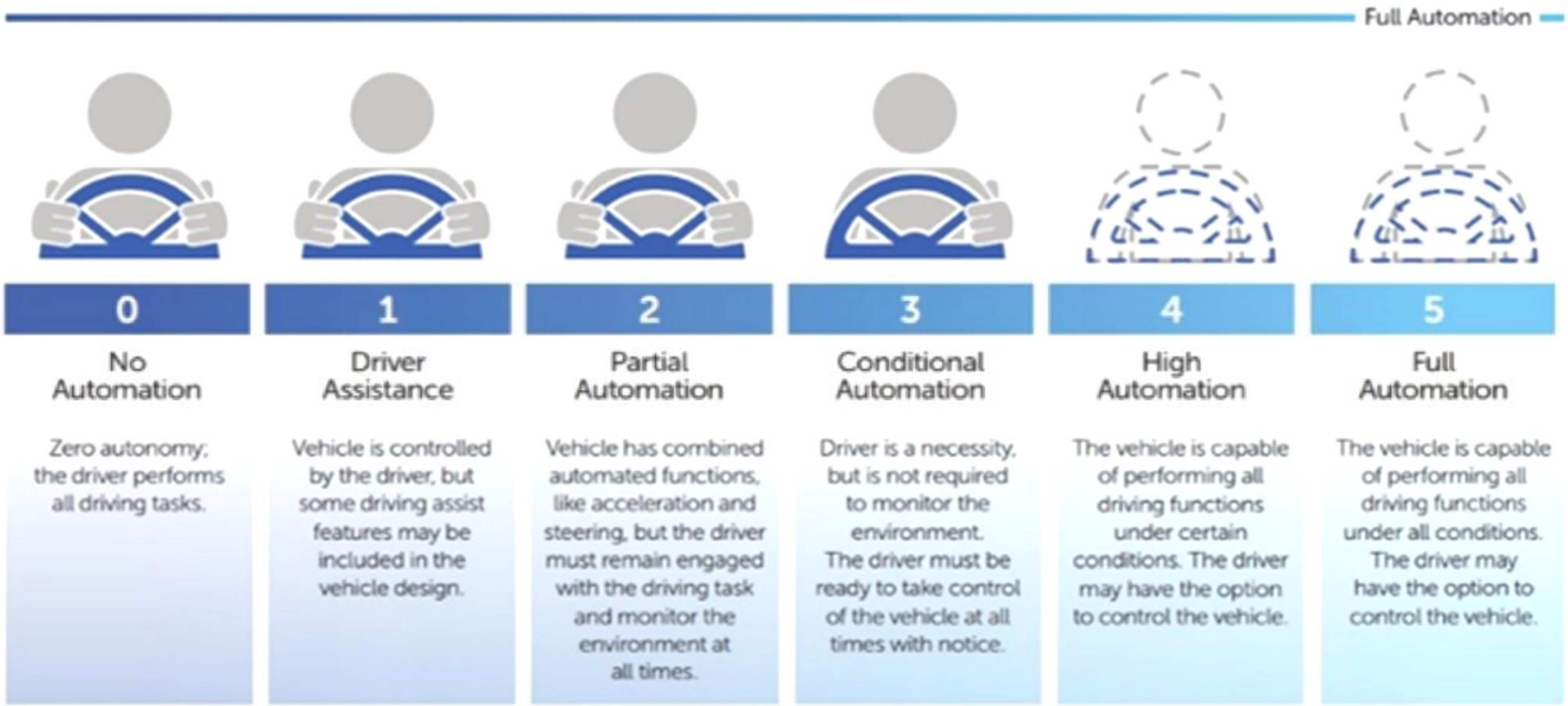
Radiology E

- When AI … her findings?
- Poor followi

The Road to Autonomous AI

Six Levels of Automation (SAE J3016)

U.S. National Highway Traffic Safety Administration



The Road to Autonomous AI

Clinical Functionality	0 No Automation	1 Driver Assistance	2 Partial Automation	3 Conditional Automation	4 High Automation	5 Full Automation
Determine Image Fitness						
Improve Image Quality (PostP)						
Determine Priority (CADt)						
Detect Finding (CADe)						
Quantify Measurements (PostP)						
Single-Shot Diagnoses (CADx)						

SIIM ACR Data Science Summit VIRTUAL 2020

The AI Lifecycle

Evaluating and Monitoring AI Algorithms

ACR DSI Summit

June, 2020

Keith J. Dreyer DO, PhD, FACR
Chief Science Officer, ACR DSI

DATA SCIENCE INSTITUTE™
AMERICAN COLLEGE OF RADIOLOGY