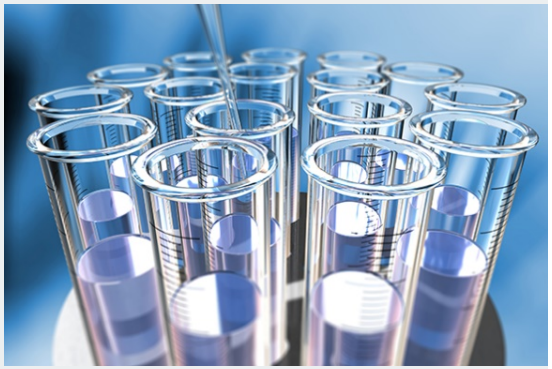
HEALTH ADVANCES

Strategy Consultants for the Healthcare Industry



Artificial Intelligence in Diagnostic Development

Kristine C. Mechem

Health Advances LLC

BOSTON | SAN FRANCISCO | ZUG | HONG KONG

www.healthadvances.com

Presented at

AAPM Annual Meeting 2021

Jan 7-8th, Hosted Virtually

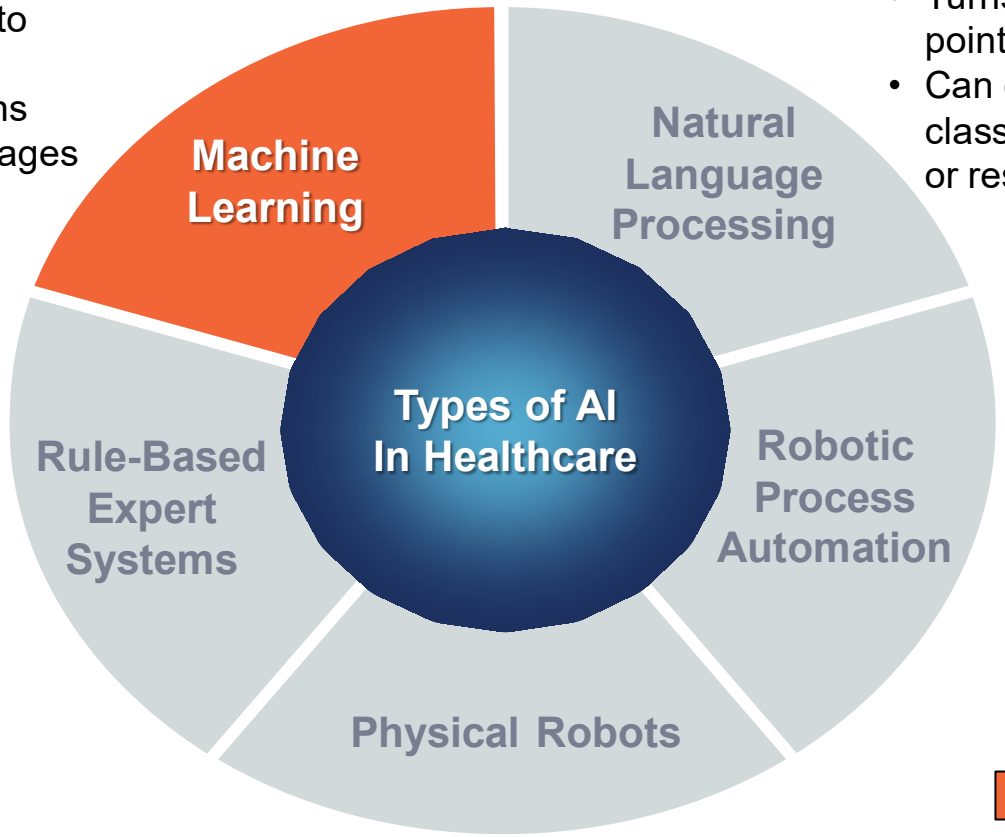
January 7, 2020

Introduction to Artificial Intelligence

Artificial intelligence (AI) refers to a wide umbrella of techniques that employ computer software to perform 'human-like' skills and can be employed for various uses in healthcare.

- Learns patterns to classify data
- Infers conclusions about data or images

- Turns language into data points
- Can create, understand, and classify clinical documentation or research reports



- Sets up 'if-then' rules to guide prompts
- Used in electronic medical records for clinical decision support

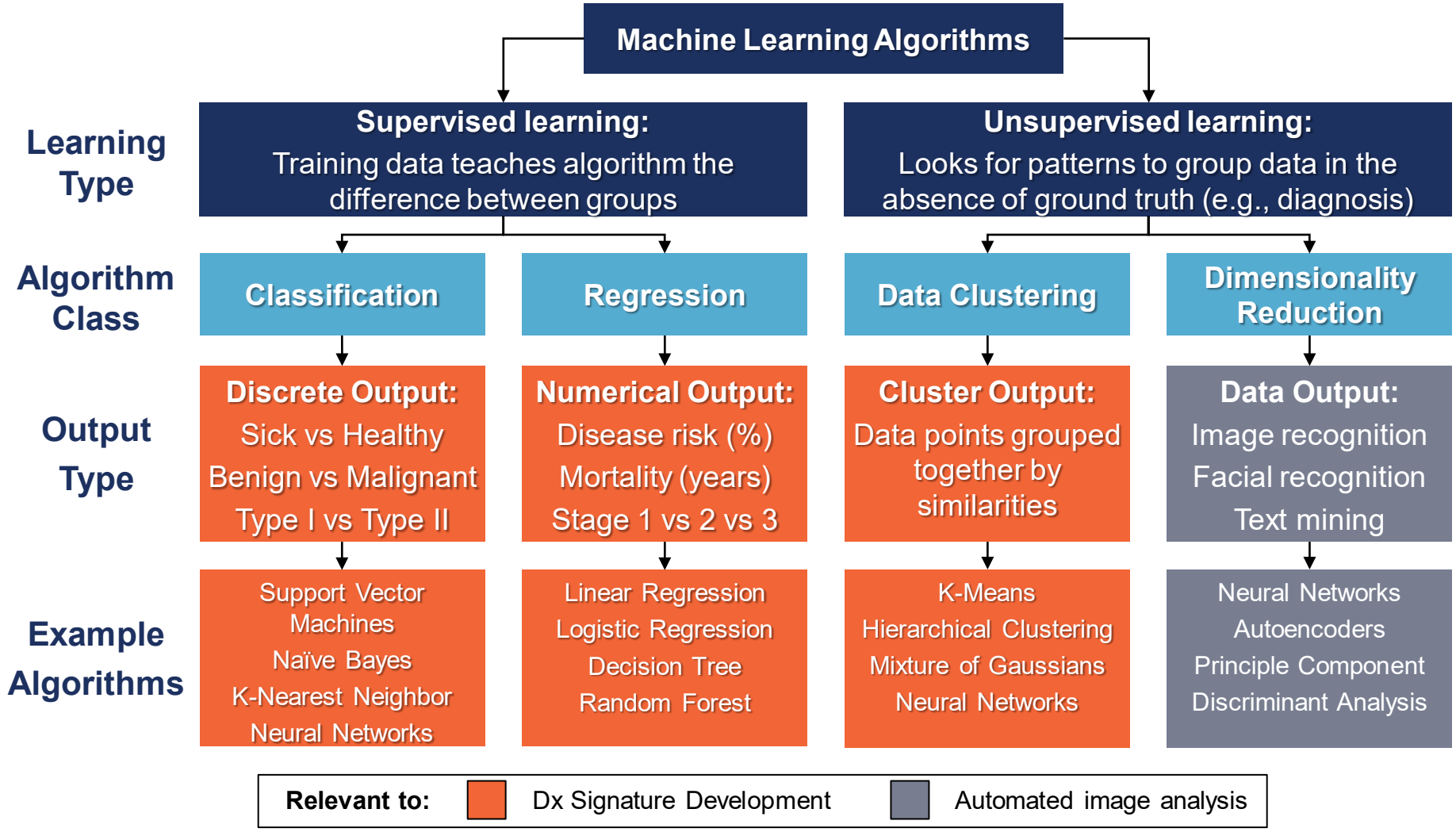
- Performs structured digital automations
- Used in routine tasks such as hospital billing and records

- Drives robotic surgical systems
- Used in many types of surgical procedures

Source: Health Advances analysis, Davenport 2019 Future Healthc J.

Overview of Machine Learning Algorithms

Machine learning (ML) is a form of artificial intelligence that generates algorithms to detect useful patterns from large, unstructured data sets.



Source: Health Advances analysis, Medium, Uddin 2019 BMC Med Inform Decis Mak.

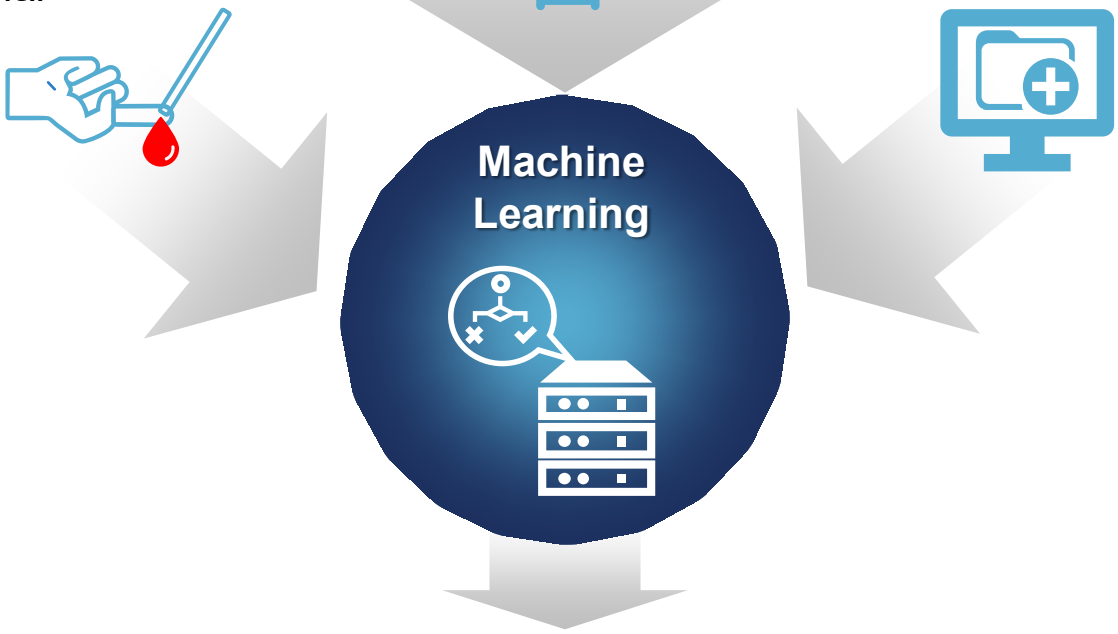
Application of ML in Diagnostic Panel Development

ML inputs large amounts of clinical and sample data to identify a biomarker signature that is characteristic of the disease state, which can then be used as a diagnostic test.

Nucleic Acid Sequencing
Genomic or cfDNA sequencing
RNA transcriptome analysis

Biological Markers
Protein markers
Metabolic analytes
Biopsy material

Clinical Covariates
Medical history
General factors
Tumor morphology



Biomarker Signature

Source: Health Advances analysis, Uddin 2019 BMC Med Inform Decis Mak.

Drivers of ML Utilization in Diagnostic Development

ML can facilitate new diagnostics for diseases that were previously undiagnosable or poorly diagnosed and improve treatment selection by bettering understanding disease progression

Drivers of ML Utilization

To diagnose a disease for which no single factor can provide a diagnosis

- For when no single gene mutation or antigen characteristically defines a given disease
  - Example disease: Lupus erythematosus is not diagnosable with any single-biomarker test

→ **Train ML on healthy vs disease to find biomarker signature**

To provide a more in-depth diagnosis

- For stratifying a disease into various subgroups and stages, or potential to respond to a therapy
  - Example: OncotypeDx for predicting need for adjuvant chemotherapy in breast cancer

→ **Train ML on patient sub-groups to find differentiating factors**

To improve existing diagnostic tests

- New panels could be easier to use, faster, or more cost-effective than existing tests, thus driving wider adoption
  - Example: Cologuard replaces invasive colonoscopy procedure with stool test and improves specificity over FIT

→ **Train ML to improve test metrics**

Source: Health Advances interviews and analysis.

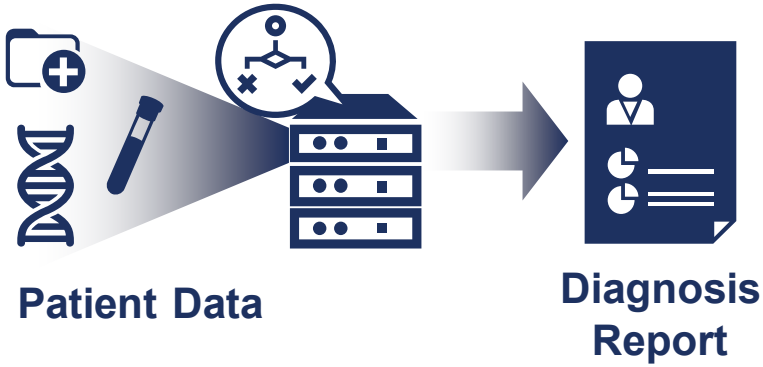
ML Use Cases in Diagnostic Panel Development

ML utilization in biomarker signature discovery and diagnostic development falls into two major use cases: panel-based tests and sequencing-based tests.

1 Panel-Based Tests

ML generates a signature in the form of a handful of biomarkers that can be integrated into a multi-marker diagnostic panel. The panel is locked and cannot be evolved over time.

2 Sequencing-Based Tests



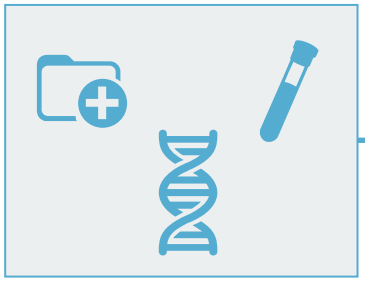
ML identifies a signature based on deep sequencing data. Full or extensive sequencing data is provided for each new diagnosis case allowing the algorithm ti evolve over time as it finds new predictive genes in the dataset.

Note: Sequencing-based tests include unbiased proteomic profiling.
Source: Health Advances analysis, company websites.

Use Case: Panel-Based Tests

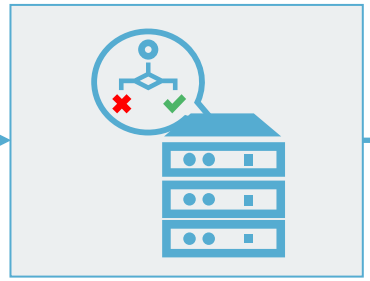
For panel-based tests, the use of ML is limited to the biomarker discovery step. Once the signature has been identified, the development process is similar to traditional diagnostics.

Data Collection



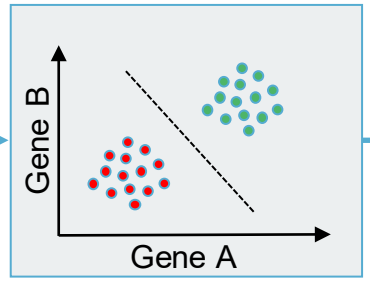
All available patient data is standardized and inputted

Supervised ML



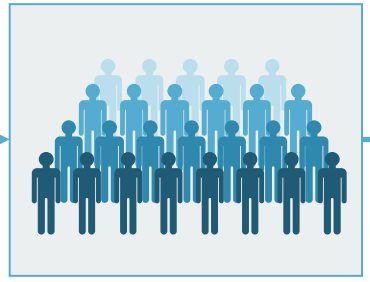
ML trains on known healthy and disease patients samples and looks for unique features

Signature Identification



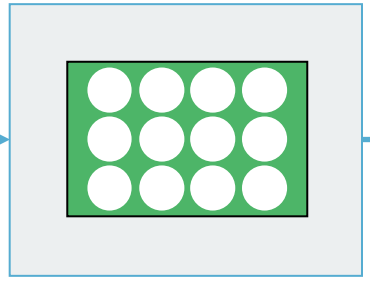
ML identifies a biomarker signature that is a limited number of genes or proteins that is unique to each group

Marker Validation



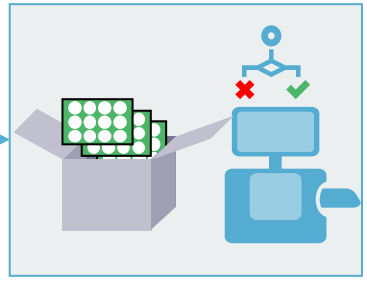
Set number of identified gene/protein markers validated in independent population (100s-1000s)

Panel Design



Markers validated as predictive of disease incorporated into diagnostic panel platform (e.g., ELISA, qPCR)

Final Product



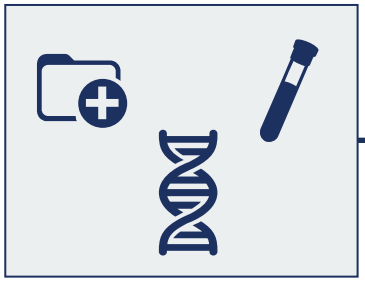
Test is a set panel; the output is put into the algorithm to calculate a score

Source: Health Advances analysis, company websites.

Use Case: Sequencing-Based Tests

For ML-algorithms, the algorithm retains full sequencing data for each new case, can take many data points into account, and can evolve over time as training is continuous.

Data Collection

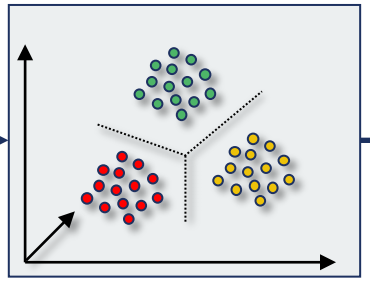


All available patient data standardized and inputted with diagnosis groups

Unsupervised ML

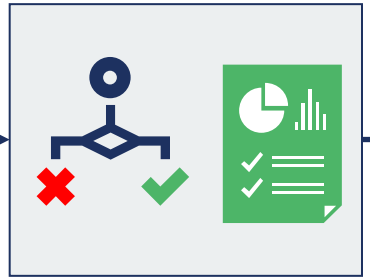
ML clusters data into set number of groups without knowing diagnosis

Signature Identification



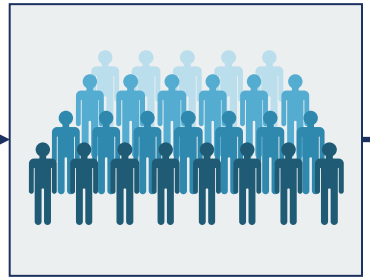
ML plots all available data in thousands of dimensions and combinations to create clusters

Algorithm Lockdown



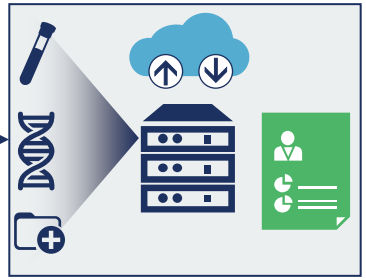
The clustering algorithm re-iterates until it correctly clusters patients by diagnosis, then it is locked

Algorithm Validation



Algorithm validated for its predictive ability on many thousands of patients

Final Product



Full sequencing data is inputted into the algorithm; the algorithm can continue training and evolve

Note: Sequencing-based tests include unbiased proteomic profiling.
Source: Health Advances analysis, Decoded Science, company websites.

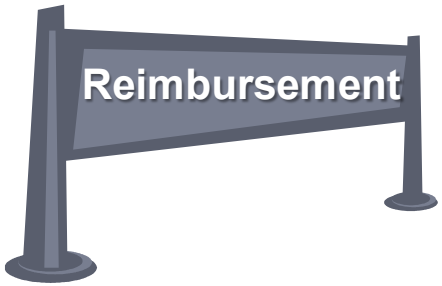
Key Barriers for Market Entry

ML-derived diagnostic tests must gain sufficient evidence in large validation studies, have clear scientific rationale, and change the course of clinical care to successfully enter the market.

- Large number of samples needed in large, global clinical trials
- Appropriately defined controls (e.g., heavy smokers for lung cancer diagnosis)
- Independent training and validation sets
- Prospective clinical trials (retrospective or banked samples can be use for discovery, training)

- Almost all ML-derived tests will operate as specialty labs
- VALID Act and precedent may result in greater FDA scrutiny
- Sequencing-based tests can be evolved without regulatory input in CLIA environment

- Physicians are not satisfied with 80-90% predictive sensitivity in some ML tests
- Physicians want to understand the scientific logic behind the biomarker signatures identified by ML



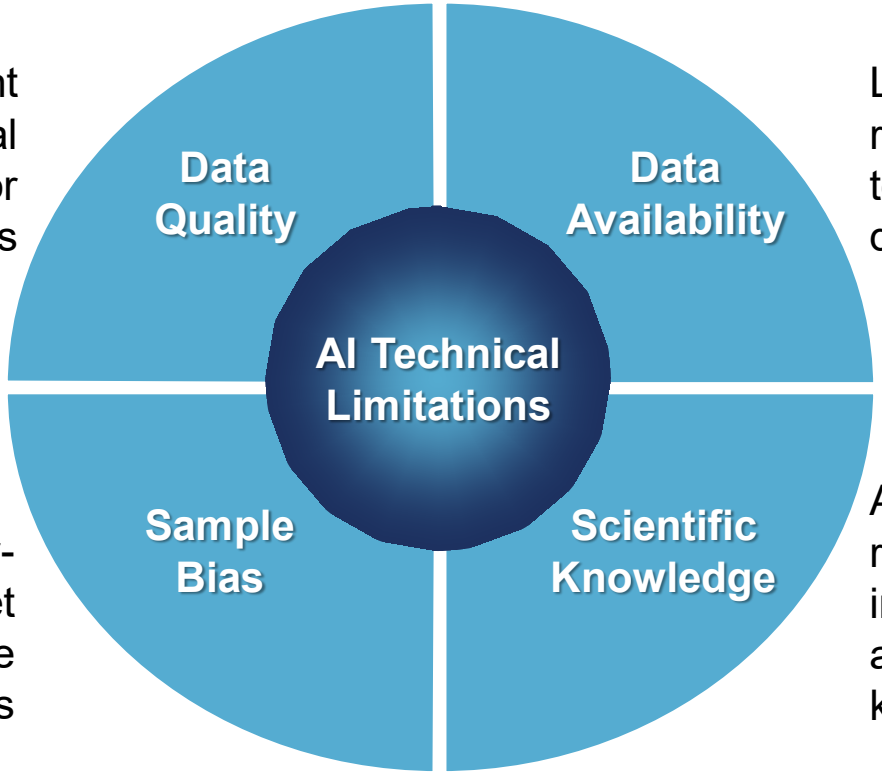
- Payers are demanding prospective clinical validation
- Payers want to see that the test ultimately changes the standard of care and clinical outcomes (e.g., reduces procedures or improves outcomes)
- Value based pricing possible when clinical outcome is significantly improved

Source: Health Advances interviews and analysis.

Major Technical Limitations

While AI/ML algorithms are largely developed and ready for use, the current major technical constraints are centered around the inputted data.

High-quality patient sample data/clinical history required for analysis



Large sample sizes required restrict use to higher prevalence disease areas

Algorithm may over-train on sample set and not be accurate on other populations

Algorithm cannot replace human inference of scientific and medical knowledge

Source: Health Advances interviews and analysis.

Contact Information

Kristine Mechem, PhD
Vice President
kmechem@healthadvances.com

Health Advances LLC
9 Riverside Road
Weston, MA 02493

101 Second Street, Suite 800
San Francisco, CA 94105

781-647-3435

415.834.0800
www.healthadvances.com