AMERICAN SOCIETY OF ECHOCARDIOGRAPHY — ECHO HAWAII 2022

Echo Beyond 2022: *Artificial Intelligence and Machine Learning*

Sanjiv J. Shah, MD, FASE
Stone Professor of Medicine
Director of Research, Bluhm Cardiovascular Institute
Director, Northwestern HFpEF Program
Division of Cardiology, Department of Medicine
Northwestern University Feinberg School of Medicine
sanjiv.shah@northwestern.edu • http://www.hfpef.org • Twitter: @HFpEF

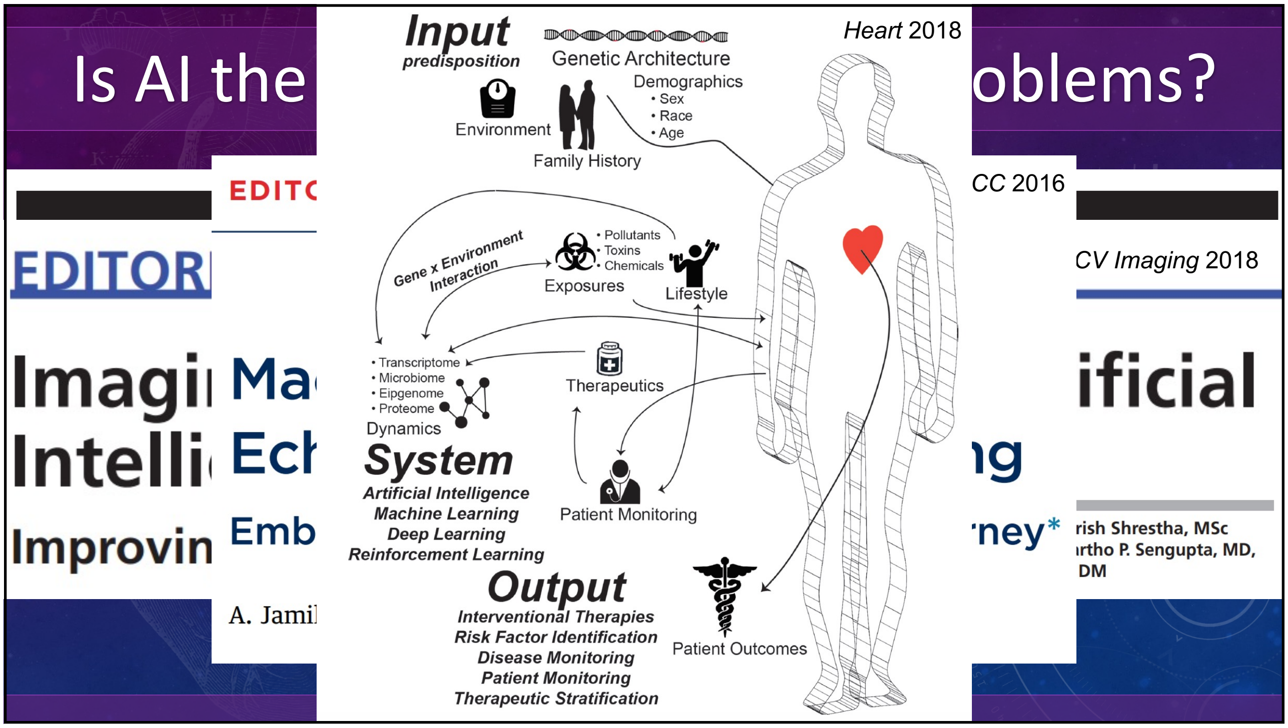
NORTHWESTERN UNIVERSITY FEINBERG SCHOOL OF MEDICINE

1

Disclosures

- Research funding:
  - ✓ NIH U54 HL160273, R01 HL107577, R01 HL127028, R01 HL140731
  - ✓ AHA #16SFRN28780016, #15CVGPSD27260148
  - ✓ Actelion, AstraZeneca, Corvia, Novartis, Pfizer
- Consulting / advisory board / steering committee:
  - ✓ Abbott, Actelion, AstraZeneca, Amgen, Aria CV, Axon Therapies, Bayer, Boehringer-Ingelheim, Boston Scientific, Bristol-Myers Squibb, Cardiora, Coridea, CVRx, Cyclerion, Cytokinetics, Edwards Lifesciences, Eidos, Eisai, Imara, Impulse Dynamics, Intellia, Ionis, Ironwood, Lilly, Merck, MyoKardia, NGMbio, Novartis, Novo Nordisk, Pfizer, Prothena, Regeneron, Rivus, Roche, Sanofi, Shifamed, Tenax, Tenaya, and United Therapeutics

2



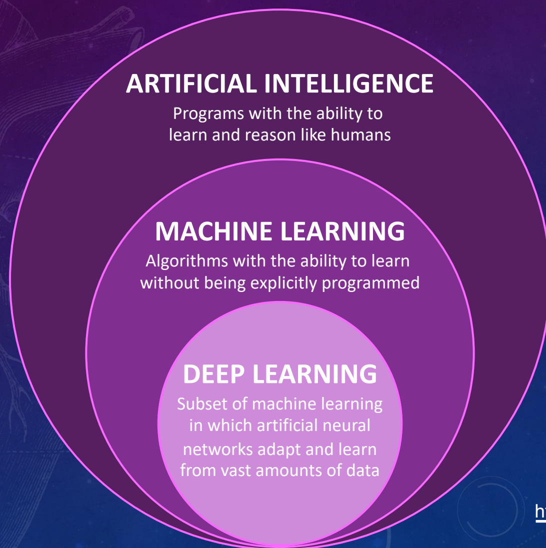
3

What problems can we solve?

- Diagnosis
  - ✓ Rare diseases, missed diagnoses, misdiagnoses
  - ✓ Patients in need of specialized treatment options
- Classification
  - ✓ Heterogeneous clinical syndromes
- Automation
- Risk prediction

4

AI vs. machine learning



https://www.qubole.com/blog/deep-learning-the-latest-trend-in-ai-and-ml/

5

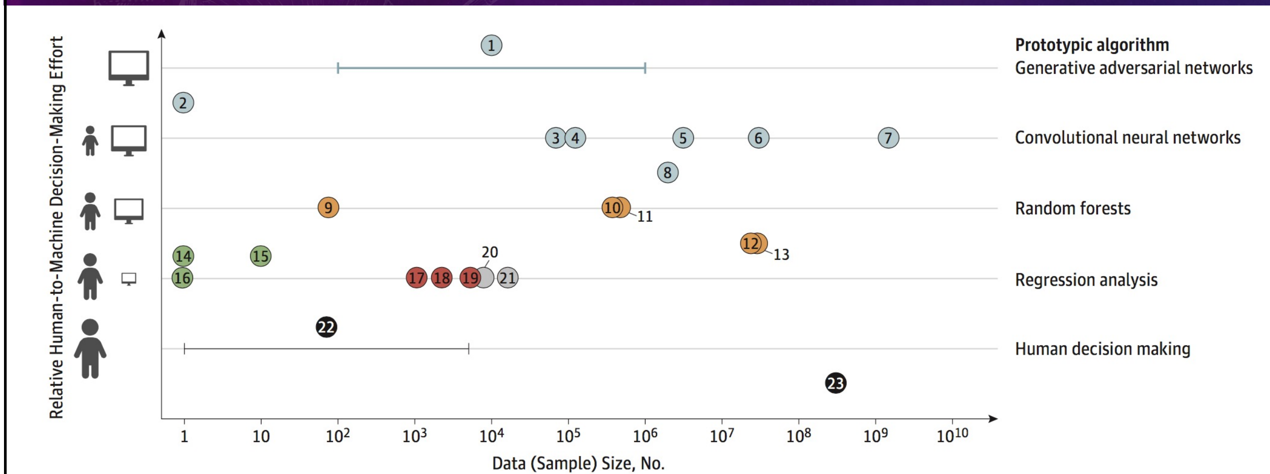
What is machine learning?

- Machine learning: A program that learns to perform a task or make a decision automatically from data rather than having to be explicitly programmed
  - Merges statistics + computer science
  - Statistics: seeks to learn relationships from data
  - Computer science: Optimizes efficiency of computer algorithms

Beam & Kohane. *JAMA* 2018; Deo RC. *Circulation* 2015

6

Machine learning spectrum



Beam & Kohane. *JAMA* 2018

7

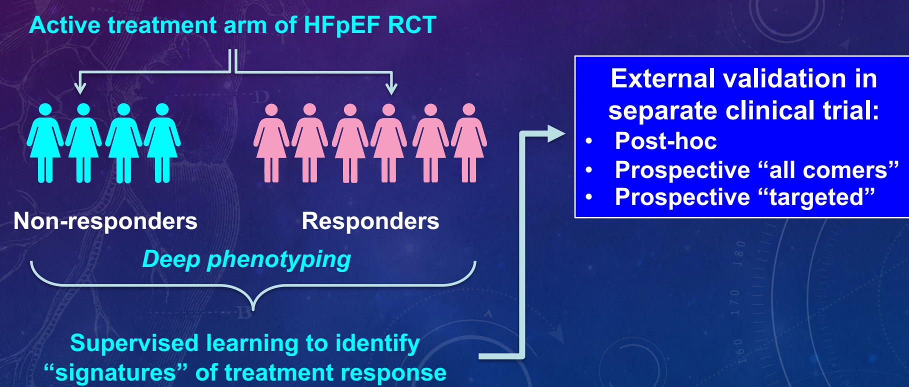
Machine learning: Key concepts

- Types of machine learning
  - Supervised learning: learning based on labeled data
  - Unsupervised learning: pattern recognition in unlabeled data
  - Deep learning: neural networks to handle high-density data
- Bias-variance trade-off + regularization
- Bigger data ≠ better data
- Feature selection
- Train-validate-test

8

Supervised learning

- Examples: logistic regression, support vector machines, random forests



9

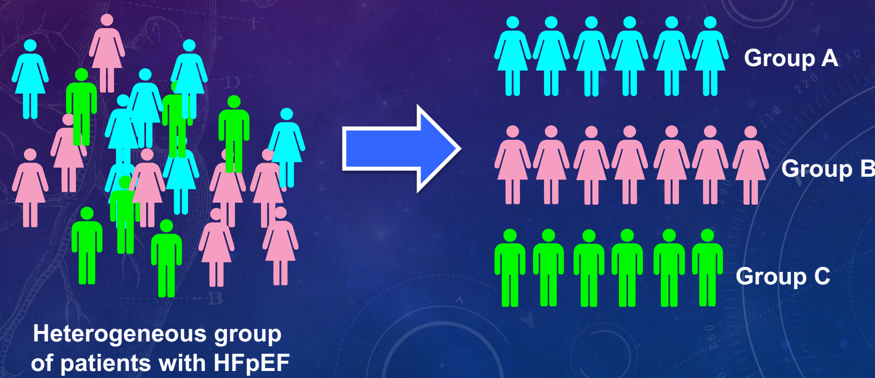
Unsupervised learning

- Examples: hierarchical clustering, model-based clustering, tensor factorization

10

Unsupervised learning

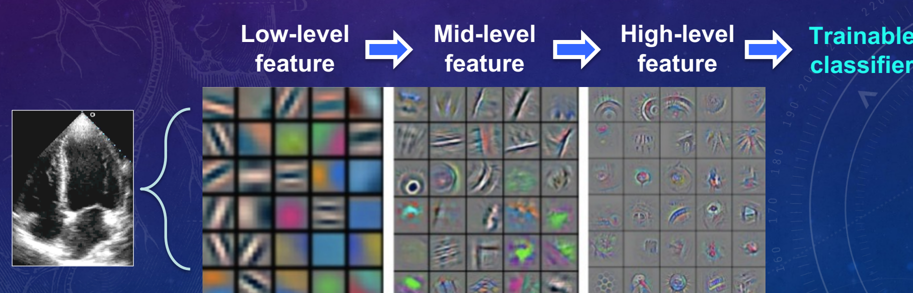
- Examples: hierarchical clustering, model-based clustering, tensor factorization



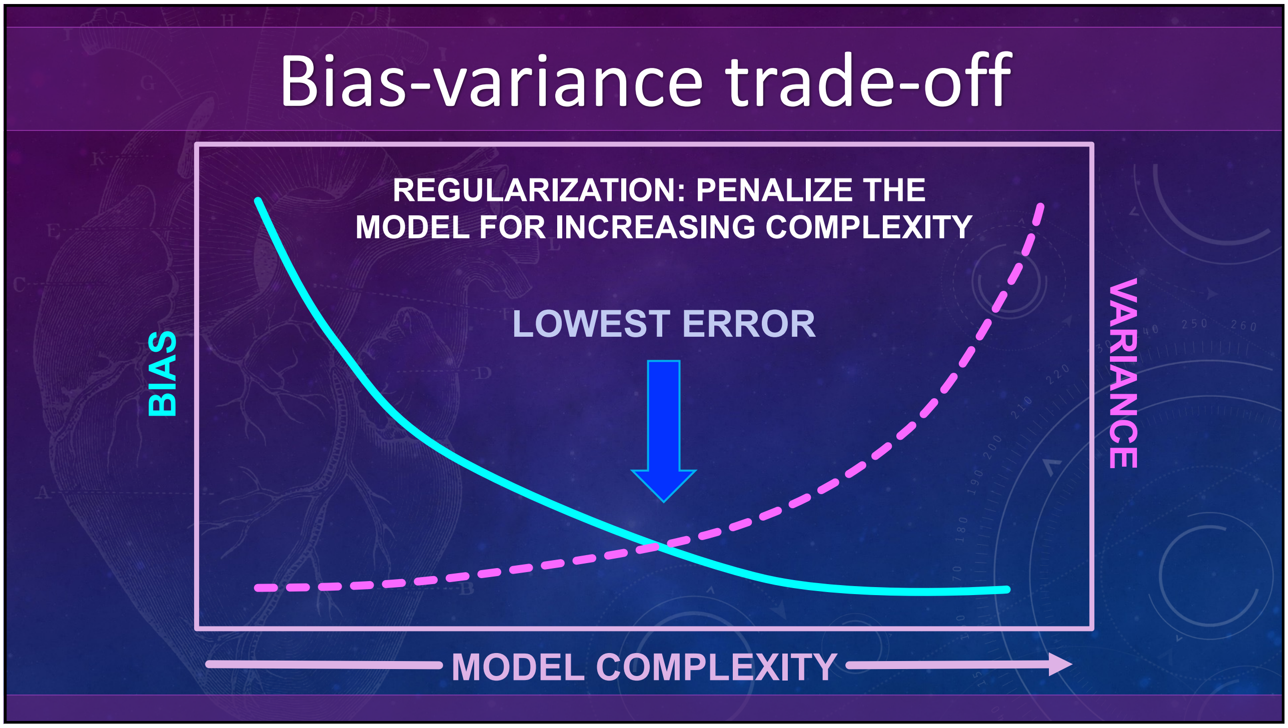
11

Deep learning

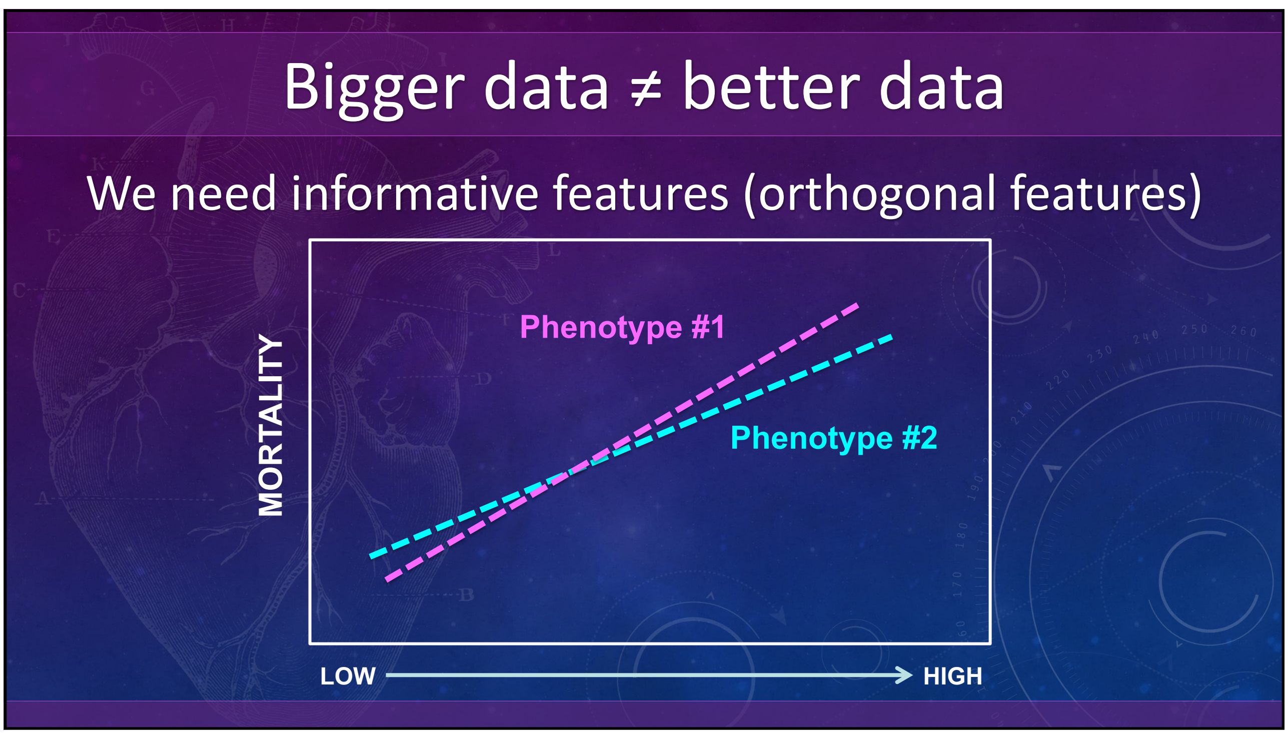
- Deep learning: *supervised or unsupervised*
  - Useful for very large datasets (e.g., imaging)
  - Neural network with multiple layers of nodes for feature identification and classification



12


Bias-variance trade-off
REGULARIZATION: PENALIZE THE
MODEL FOR INCREASING COMPLEXITY
LOWEST ERROR
BIAS
VARIANCE
MODEL COMPLEXITY

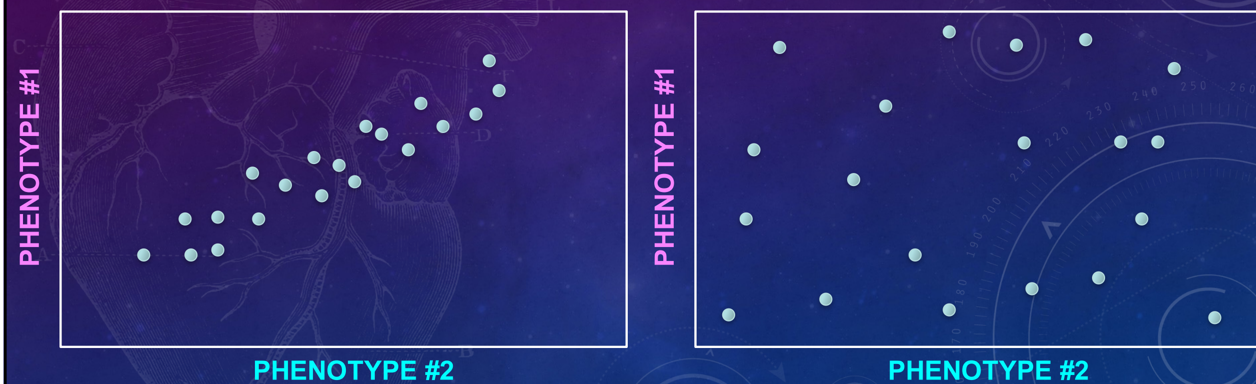
13


Bigger data ≠ better data
We need informative features (orthogonal features)
Phenotype #1
Phenotype #2
MORTALITY
LOW
HIGH

14

Bigger data ≠ better data

We need informative features (orthogonal features)



NOT INFORMATIVE   INFORMATIVE

15

Feature selection

- Which features (variables) should be included in the machine learning model?
- When presented with a large number of features, how can we select the "best" features?
  - Supervised learning to rank features (Random Forests)
  - Unsupervised learning to reduce the dimensions of the data (Principal Components Analysis)

16

Train-validate-test

- The ML model will always look "great" in the training dataset
- Use internal validation to tune the model and make it more generalizable
- External testing in a completely separate study, cohort is critical: *think about this in the study design phase*

17

Key steps in machine learning

- Identify rich dataset for training and a separate, similar dataset for testing
- Determine which variables to include in the machine learning analysis
- Handle data missingness and dimension reduction
- Decide on type of ML technique and determine optimal parameters for model
- Regularization (prevents overfitting)
- Validation and testing

18

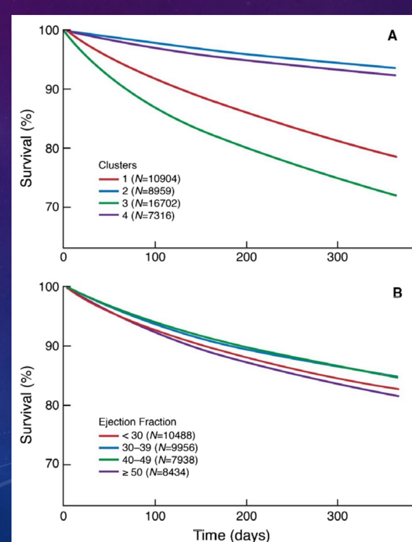
How to evaluate a machine learning study

Category	Evaluation Criteria
Study question and design	✓ Does ML offer specific advantages over conventional statistics?
Data	✓ Are data being collected primarily for research or clinical purposes? ✓ Are there issues of biases or data quality?
Approach	✓ Is there good rationale for the type of ML used? ✓ Internal validation? ✓ External testing? ✓ Is model performance superior to conventional, simpler models?
Clinical relevance	✓ Do the results have clinical relevance or provide mechanistic insight? ✓ How well should we expect the study population to generalize to the target population?

19

Machine learning in HF: growing pains

- Ahmad T, et al (*JAHA* 2018):
  - 44,086 Swedish HF patients (all LV ejection fractions)
  - Supervised learning of mortality (Random Forests): selected top 8 predictors
  - Unsupervised learning (K-means) of the 8 top predictors found 4 clusters
  - Validation: *clusters differ markedly by mortality*



SURVIVAL BY CLUSTER

SURVIVAL BY LVEF GROUP

20

Machine learning in HF: growing pains

- Frizzell JD, et al (*JAMA Cardiol* 2018):
  - Complex ML models no better than simple statistical models for prediction of 30-day readmissions in HF patients

Table. Comparison of C Statistics Judging Discriminatory Capacity in Predicting HF 30-Day Readmissions in Nationally Representative Models

Model	Study Population	No.	C Statistic[a]
TAN[b]	CMS + GWTG-HF	56 477	0.62
LR[b]	CMS + GWTG-HF	56 477	0.62
LASSO[b]	CMS + GWTG-HF	56 477	0.62
RF[b]	CMS + GWTG-HF	56 477	0.61
GBM[b]	CMS + GWTG-HF	56 477	0.61
EHR 2016[b]	CMS + GWTG-HF	56 477	0.59
EHR 2013[c]	CMS + GWTG-HF	33 349	0.59
CMS[d]	CMS	567 447	0.6

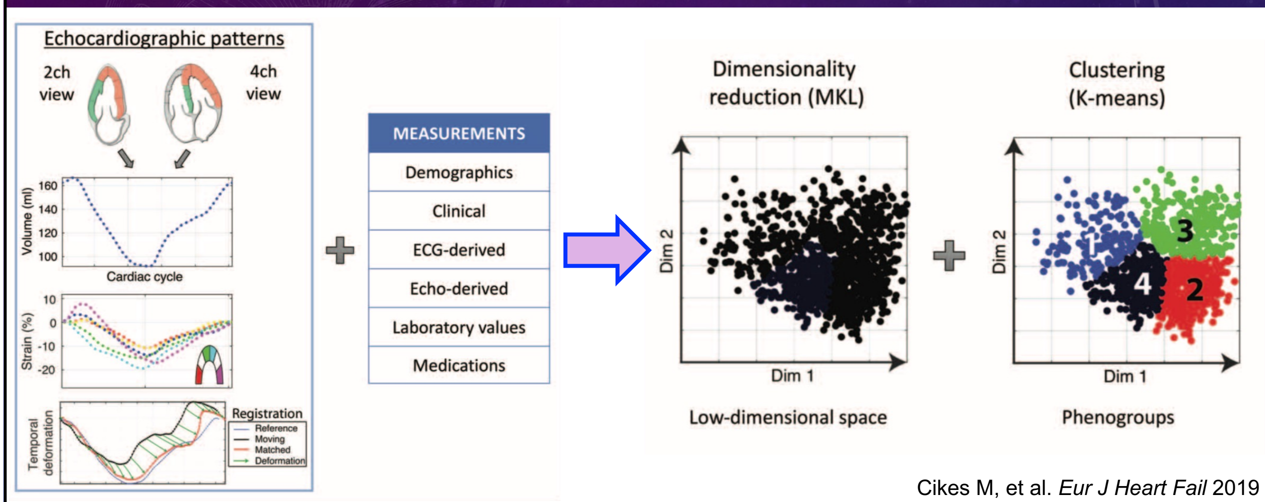
Logistic regression

Random forests

21

Machine learning to predict CRT response

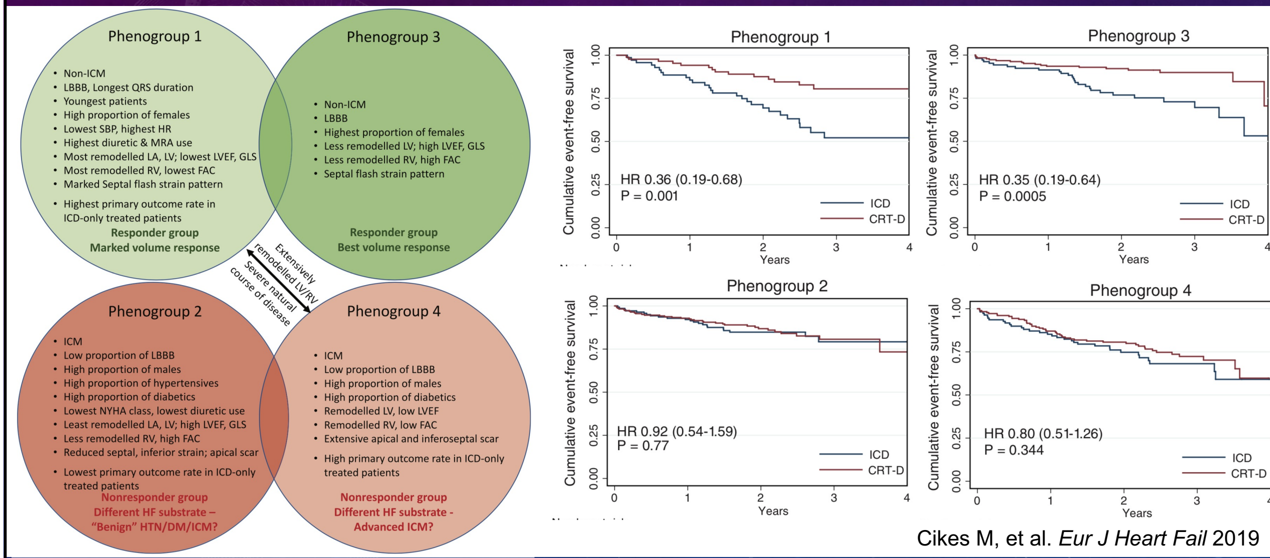
INPUT DATA **UNSUPERVISED MACHINE LEARNING**



Cikes M, et al. *Eur J Heart Fail* 2019

22

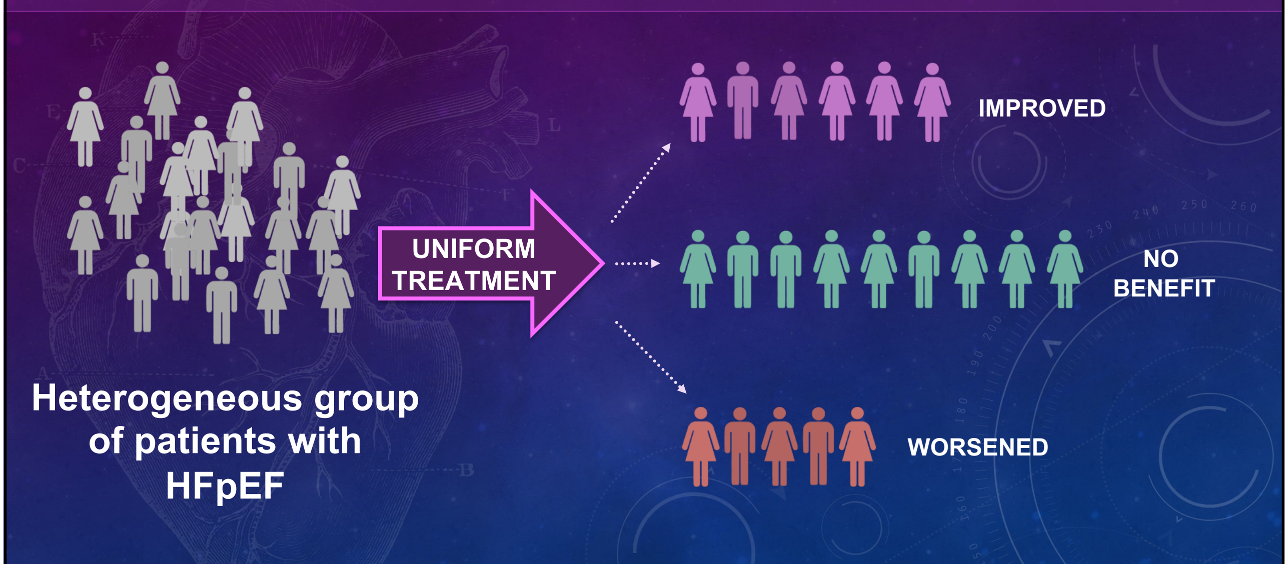
Machine learning to predict CRT response



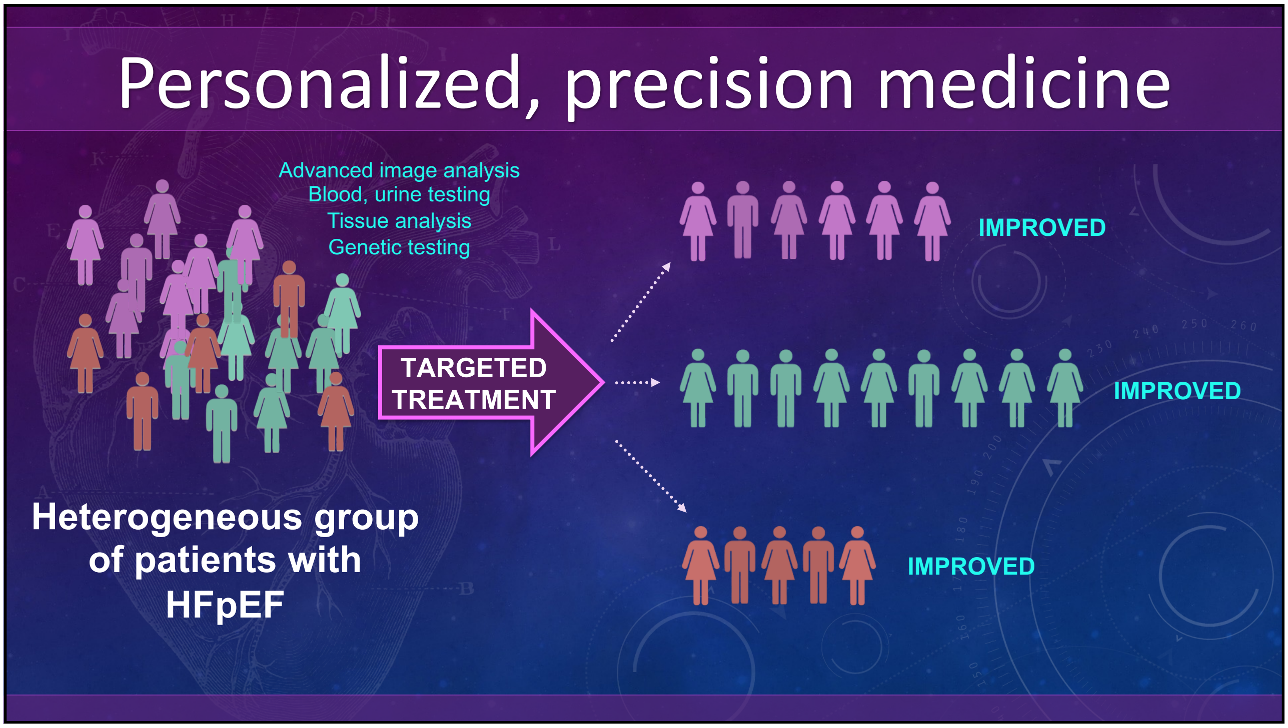
Cikes M, et al. *Eur J Heart Fail* 2019

23

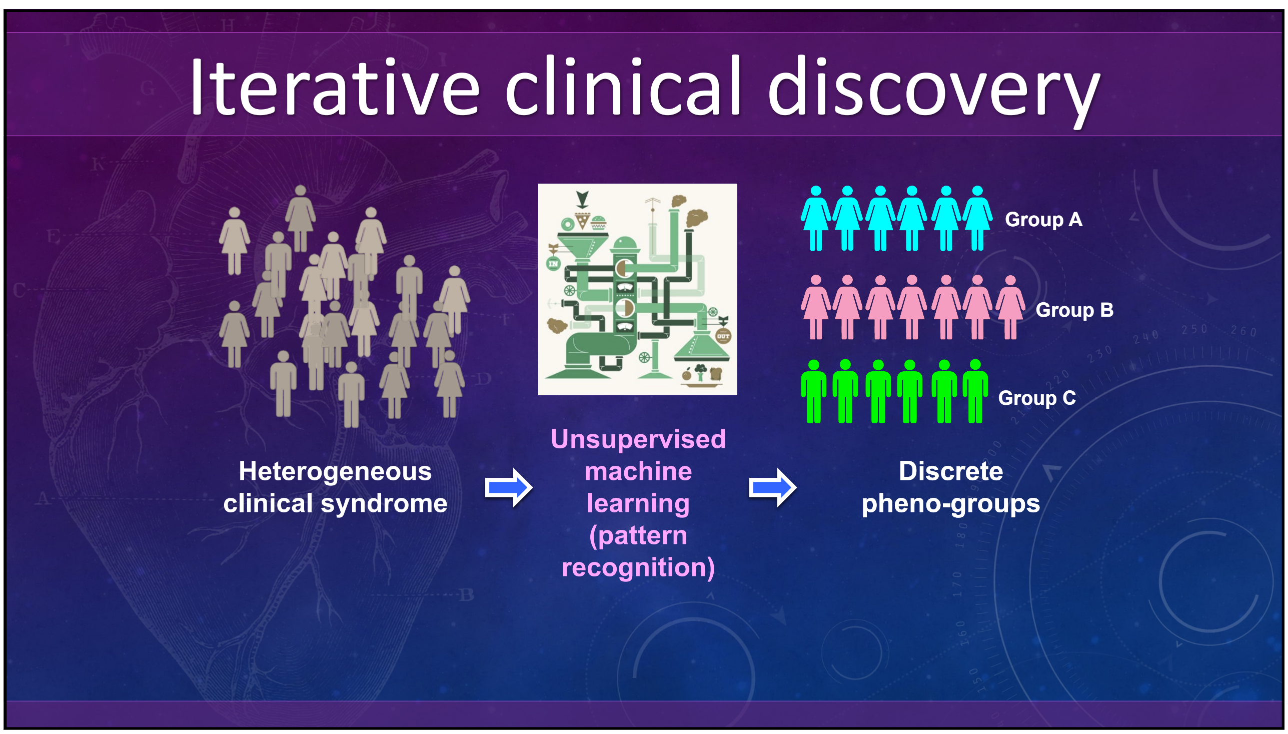
HF treatment: One-size-fits-all approach



24


Personalized, precision medicine
Advanced image analysis
Blood, urine testing
Tissue analysis
Genetic testing
TARGETED TREATMENT
IMPROVED
IMPROVED
IMPROVED
Heterogeneous group of patients with HFpEF

25


Iterative clinical discovery
Group A
Group B
Group C
Heterogeneous clinical syndrome
Unsupervised machine learning (pattern recognition)
Discrete pheno-groups

26

Iterative clinical discovery

Group A

Group B

Group C

Further investigation → Machine learning → Discrete pheno-groups

New clinical, biological insights

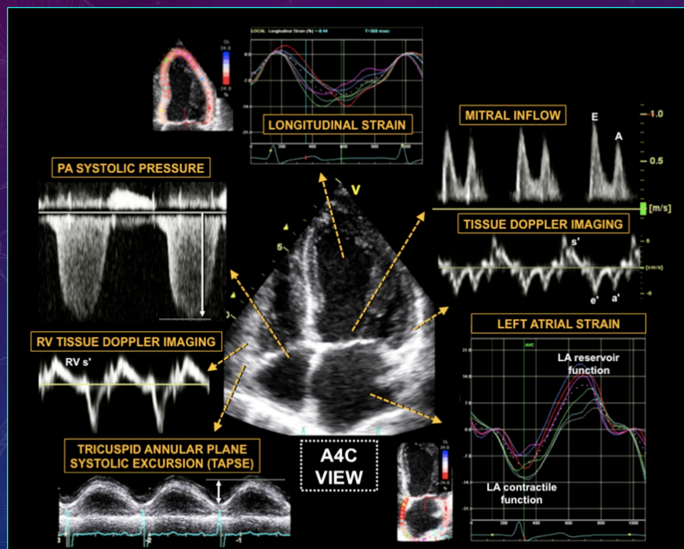
27

Heart failure vs. cancer

Sub-optimal one-size-fits-all approach	Optimal targeted approach
HF	**Cancer**
Imaging, ECG, PEX Phenotypic analysis • Quantify LVEF • Functional class • Fluid status • QRS duration	Tissue biopsy Imaging Phenotypic analysis • Tumor size, extent • Histologic analysis Gene expression
Non-targeted therapy	**Targeted therapy**

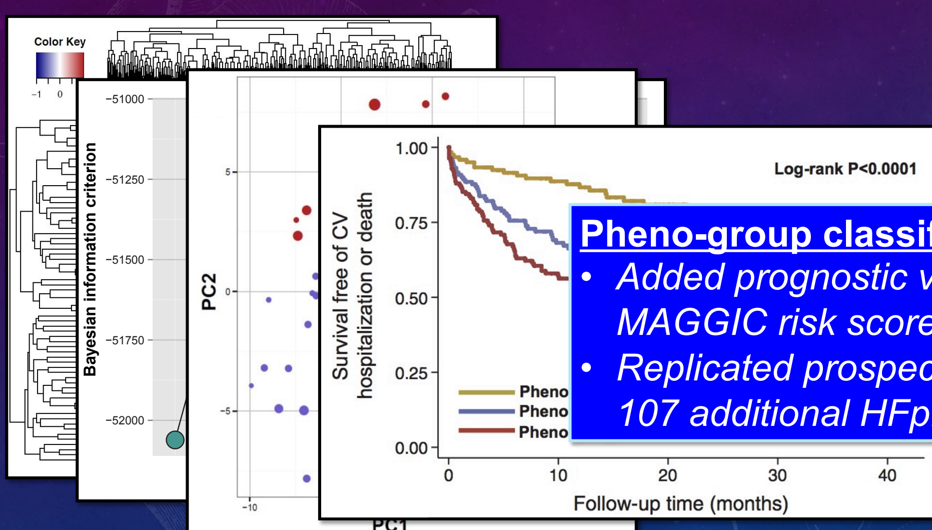
28

Echocardiography: deep phenotyping



29

Phenomapping of HFpEF

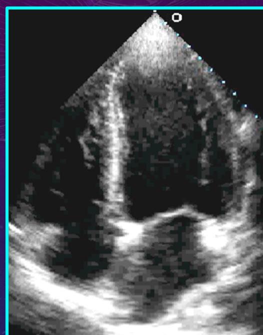
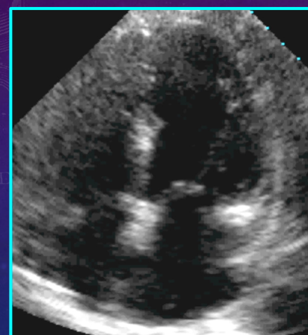
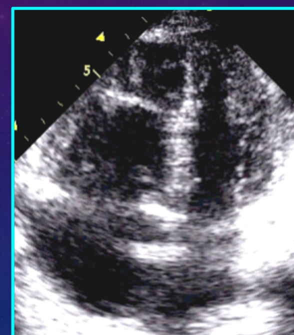


<u>Pheno-group classification</u>:
- *Added prognostic value over MAGGIC risk score and BNP*
- *Replicated prospectively in 107 additional HFpEF patients*

Shah SJ…Deo RC. *Circulation* 2015

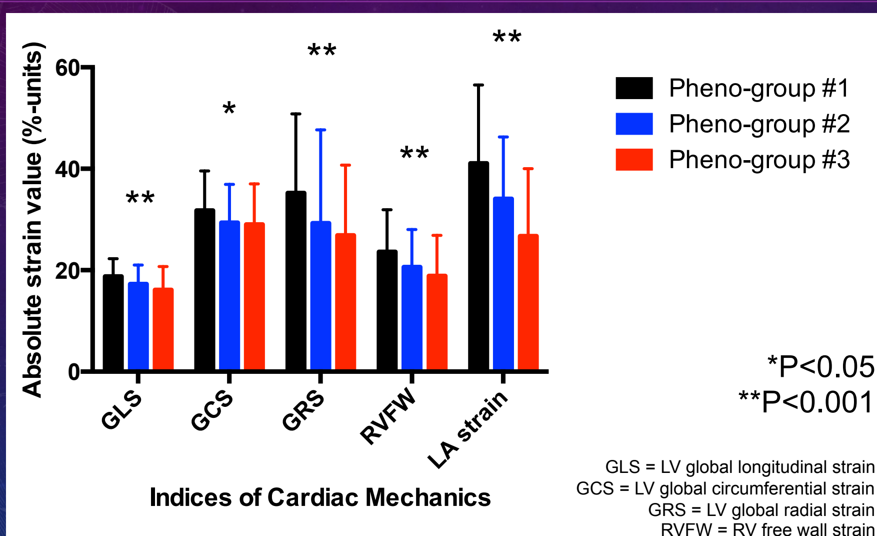
30

HFpEF pheno-groups

Pheno-group #1	Pheno-group #2	Pheno-group #3
		
Least cardiac remodeling/dysfxn Lowest BNP	Most severely impaired myocardial relaxation Highest prevalence of diabetes	Most severe electrocardiac remodeling, renal dysfunction

31

HFpEF pheno-groups: Cardiac mechanics



32

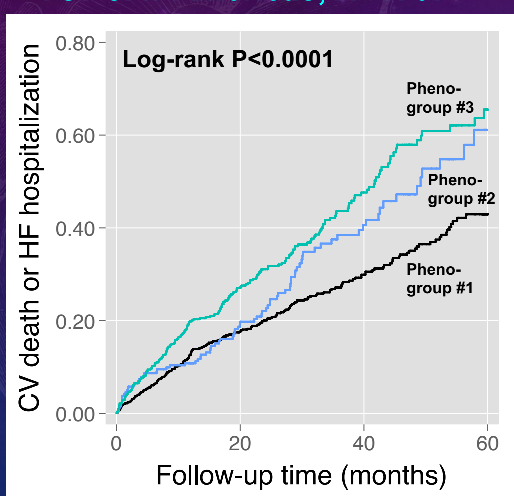
HFpEF pheno-groups: CPET

Parameter	Group 1 (n=54)	Group 2 (n=50)	Group 3 (n=46)	P-value
Exercise time (s)	469±241	310±272	356±195	0.003
Peak SBP (mmHg)	181±27	181±33	164±32	0.009
Heart rate reserve (%)	129±23	114±21	105±26	0.001
Chronotropic incompetence	50%	71%	76%	0.052
VO_2 max, ml/min/kg	16.4±6.6	11.3±2.6	13.2±4.3	<0.001
VE/VCO_2 at AT	31.1±5.1	32.4±4.9	34.4±5.3	0.015

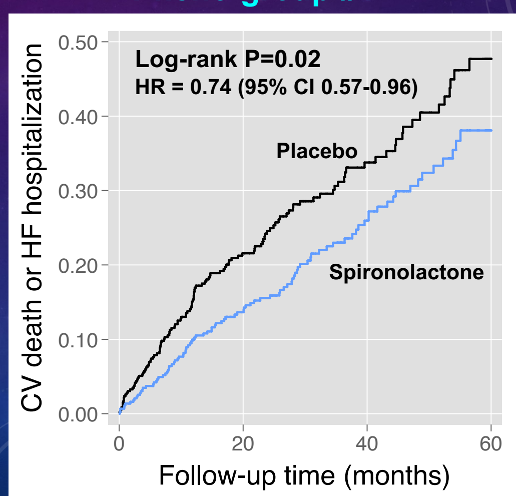
33

HFpEF phenomapping in TOPCAT

TOPCAT Americas, N=1767

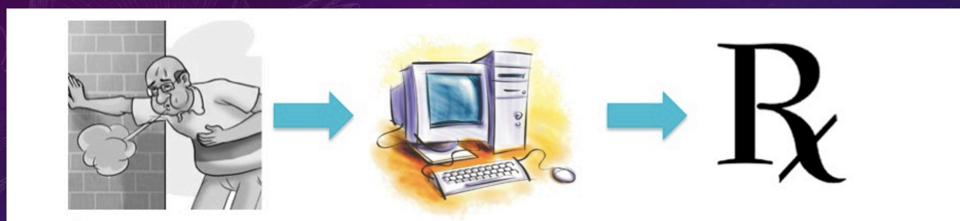


Pheno-group #1



34

Phenomapping in the clinic?



- ✓ History/physical
- ✓ Labs
- ✓ ECG
- ✓ Echo

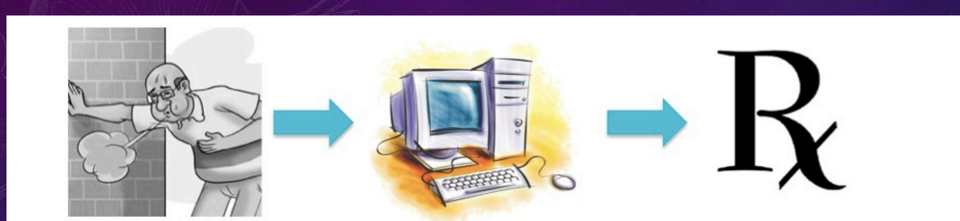
Assign pheno-group

Targeted treatment, clinical trials

Statistical learning

35

Phenomapping in the clinic?



- ✓ History/physical
- ✓ Labs
- ✓ ECG
- ✓ Echo

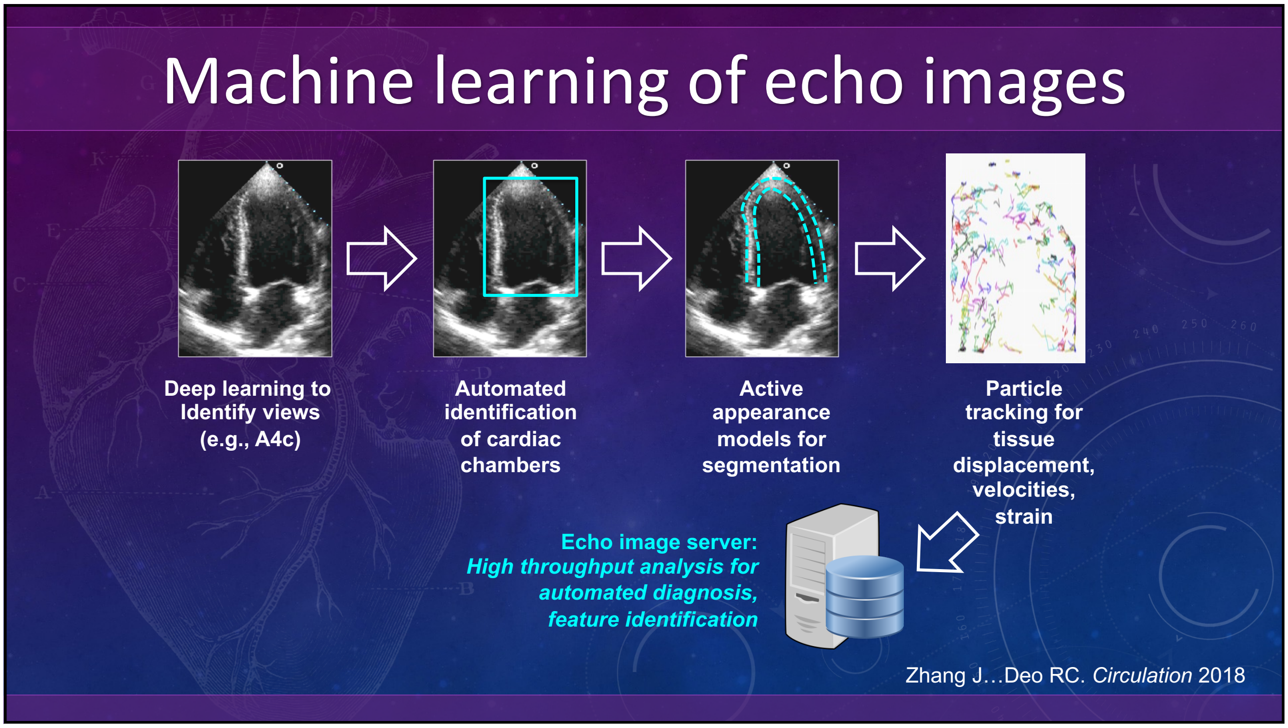
Deep learning

Statistical learning

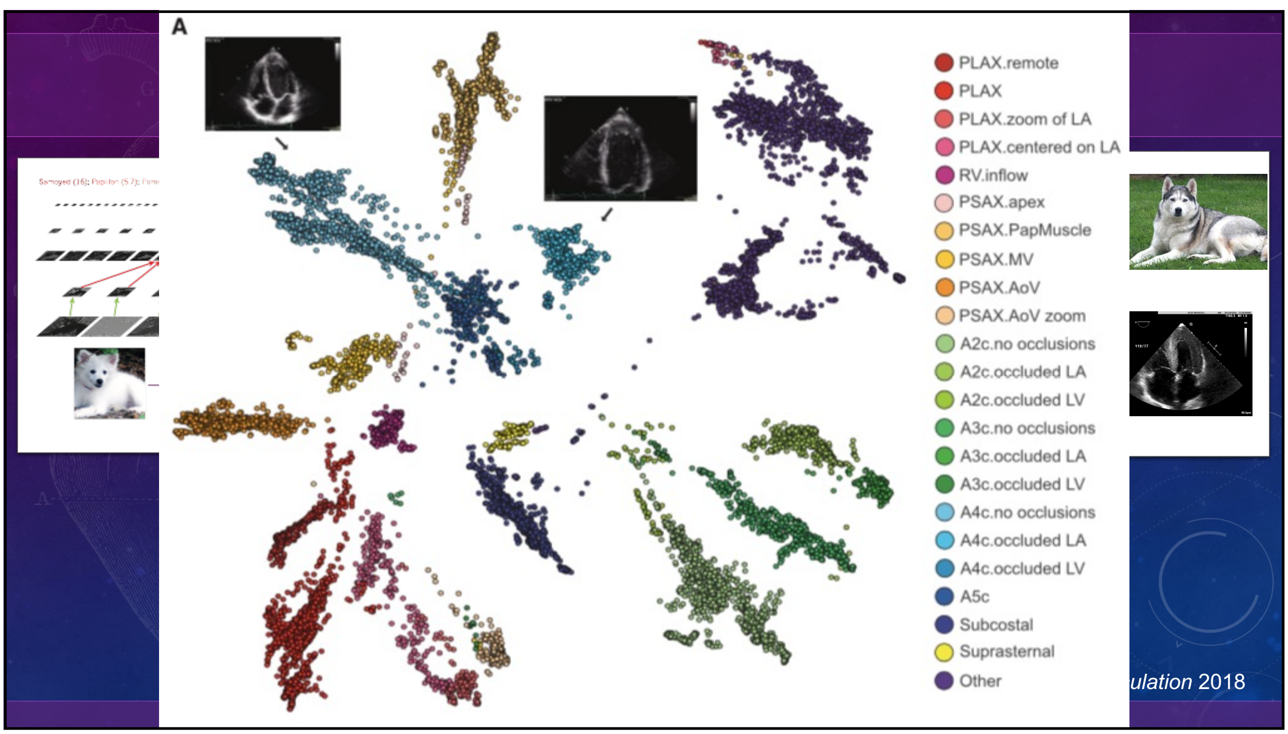
Assign pheno-group

Targeted treatment, clinical trials

36


Machine learning of echo images
Deep learning to Identify views (e.g., A4c)
Automated identification of cardiac chambers
Active appearance models for segmentation
Particle tracking for tissue displacement, velocities, strain
Echo image server:
High throughput analysis for automated diagnosis, feature identification
Zhang J…Deo RC. Circulation 2018

37


A
PLAX.remote
PLAX
PLAX.zoom of LA
PLAX.centered on LA
RV.inflow
PSAX.apex
PSAX.PapMuscle
PSAX.MV
PSAX.AoV
PSAX.AoV zoom
A2c.no occlusions
A2c.occluded LA
A2c.occluded LV
A3c.no occlusions
A3c.occluded LA
A3c.occluded LV
A4c.no occlusions
A4c.occluded LA
A4c.occluded LV
A5c
Subcostal
Suprasternal
Other
ulation 2018

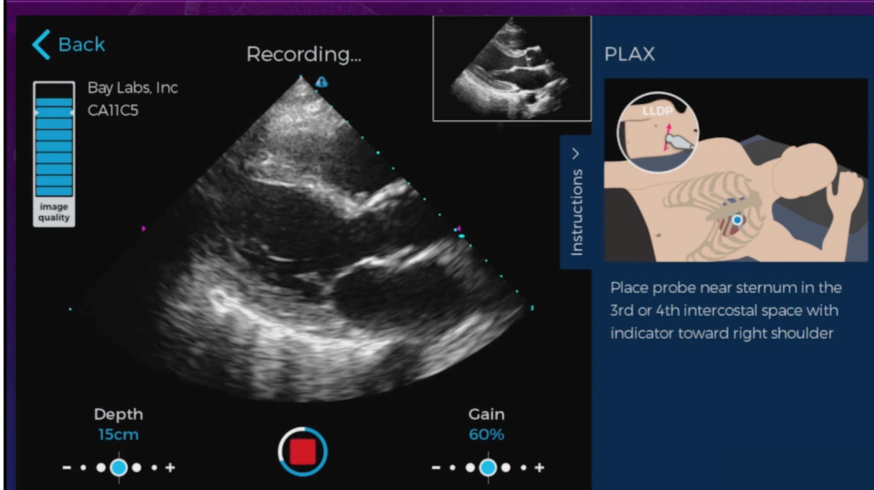
38

ML for precision diagnosis of heart disease

Can we use the same technology as self-driving cars to diagnose heart disease?

39

AI to guide echo acquisition

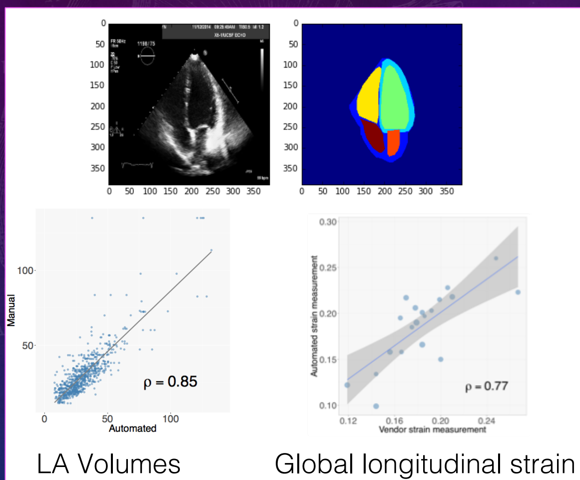


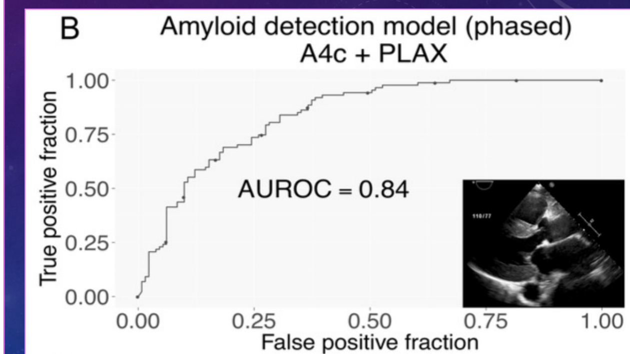
- **Improve access to echo in primary care and rural settings ("democratize echo")**
- **Deep learning algorithm guides echo acquisition by novices (nurses, MAs)**
- **8 nurses, n=240 patients**
- **Studies judged to be of diagnostic quality for LV size, function and pericardial effusion in 99% of cases and RV size in 93% of cases**
- **No major differences between nurses and sonographers**

Narang A…Thomas JD. *JAMA Cardiology* 2021

40

Automated measurements + disease detection

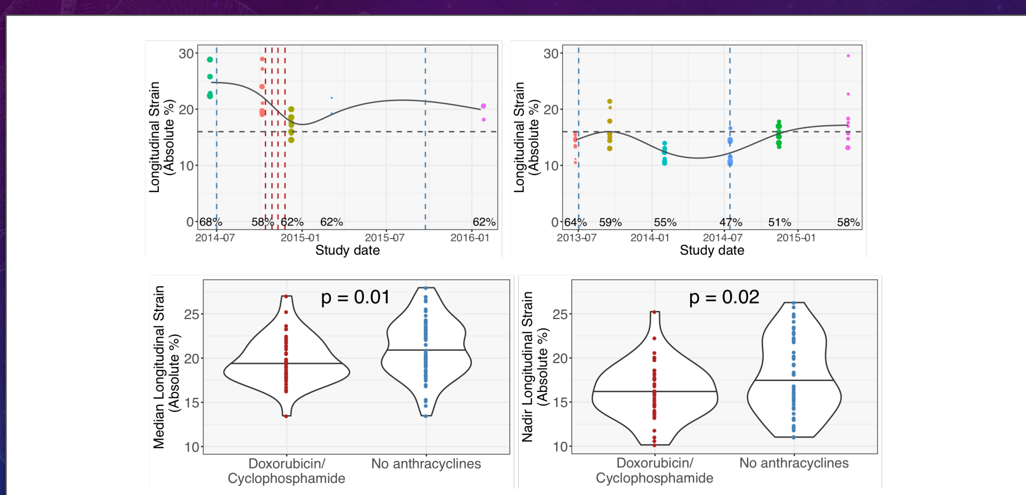




Zhang J…Deo RC. *Circulation* 2018

41

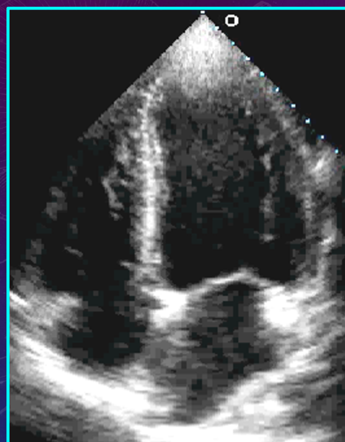
Automated disease tracking over time



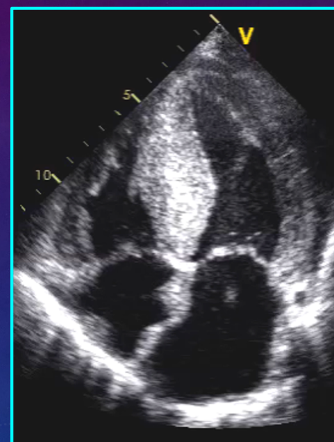
Zhang J…Deo RC. *Circulation* 2018

42

ML for precision diagnosis of heart disease



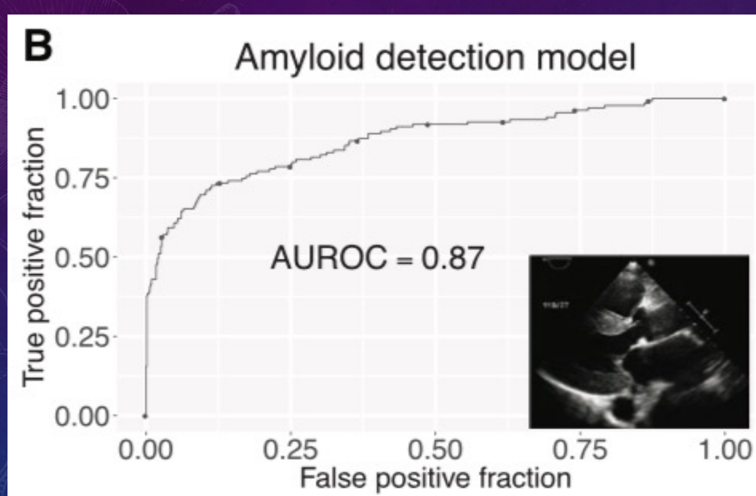
Typical HFpEF



Cardiac amyloidosis
(often misdiagnosed and requires specific treatment)

43

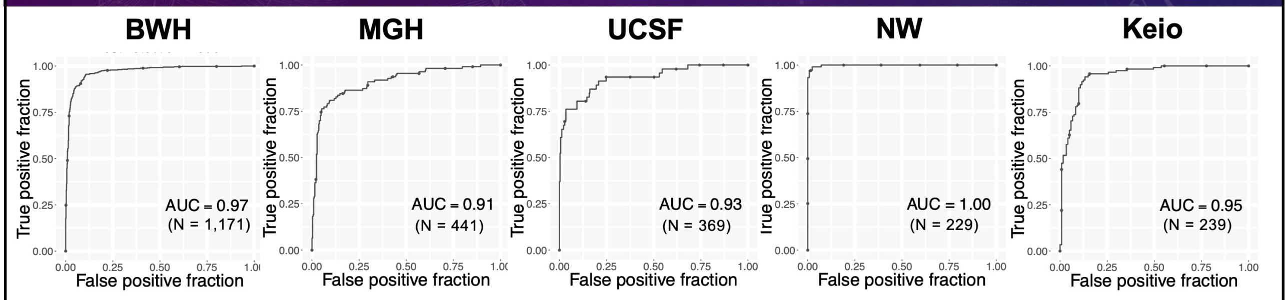
Automated measurements + disease detection



Zhang J…Deo RC. *Circulation* 2018

44

Deep learning to detect cardiac amyloidosis



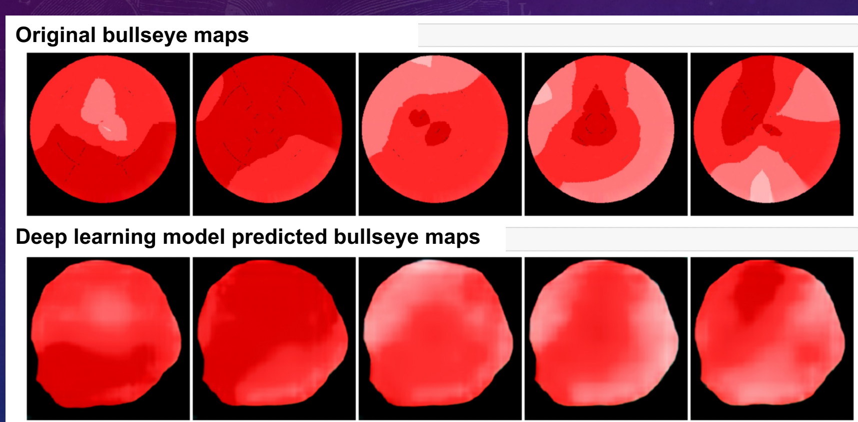
Newest deep learning model: AUC for cardiac amyloid 0.91-1.00 (based on a single view of the heart on echocardiography [heart ultrasound])

Goto S…Deo RC.
Nature Communications (in press)

45

Deep learning of bullseye patterns

Deep learning of bullseye maps in the Multi-Ethnic Study of Atherosclerosis (MESA): Population-based study, n=3,032 who underwent echo 2016-2018



Goal: find new bullseye patterns to identify disease, risk patterns, novel biology

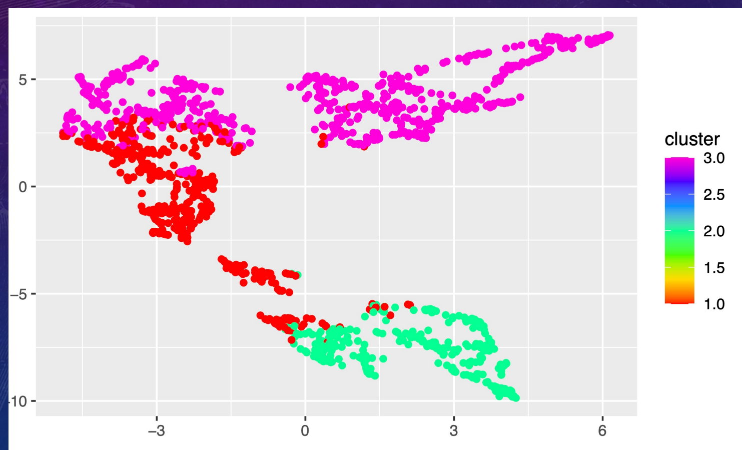
Step 1: Train deep learning model to identify "features"

46

Deep learning of bullseye patterns

Deep learning of bullseye maps in the Multi-Ethnic Study of Atherosclerosis (MESA): Population-based study, n=3,032 who underwent echo 2016-2018

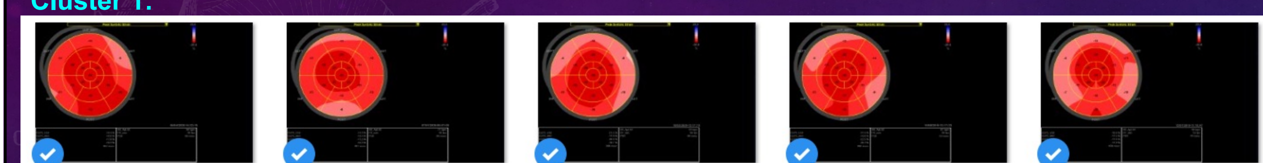
Step 2: Use statistical learning (e.g., model-based learning to cluster the features identified by deep learning model)



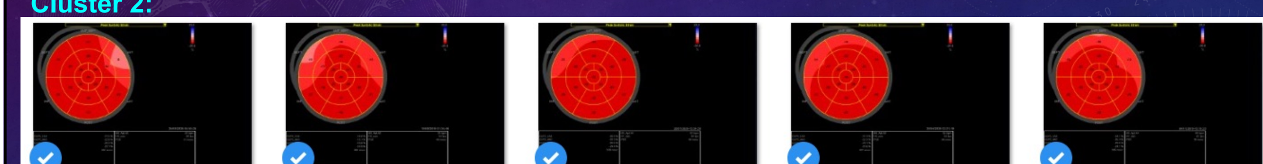
47

Deep learning of bullseye patterns

Cluster 1:



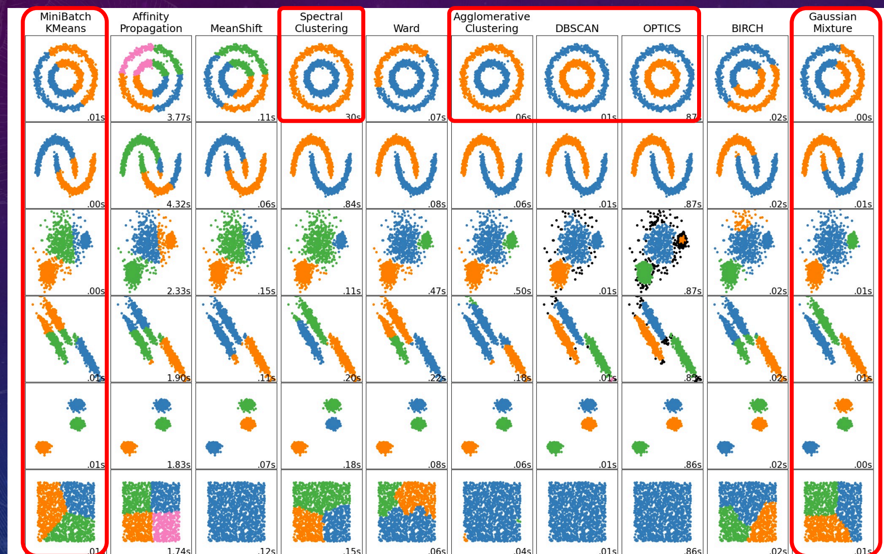
Cluster 2:



Cluster 3:

48

What went wrong?



49

HFpEF in the future

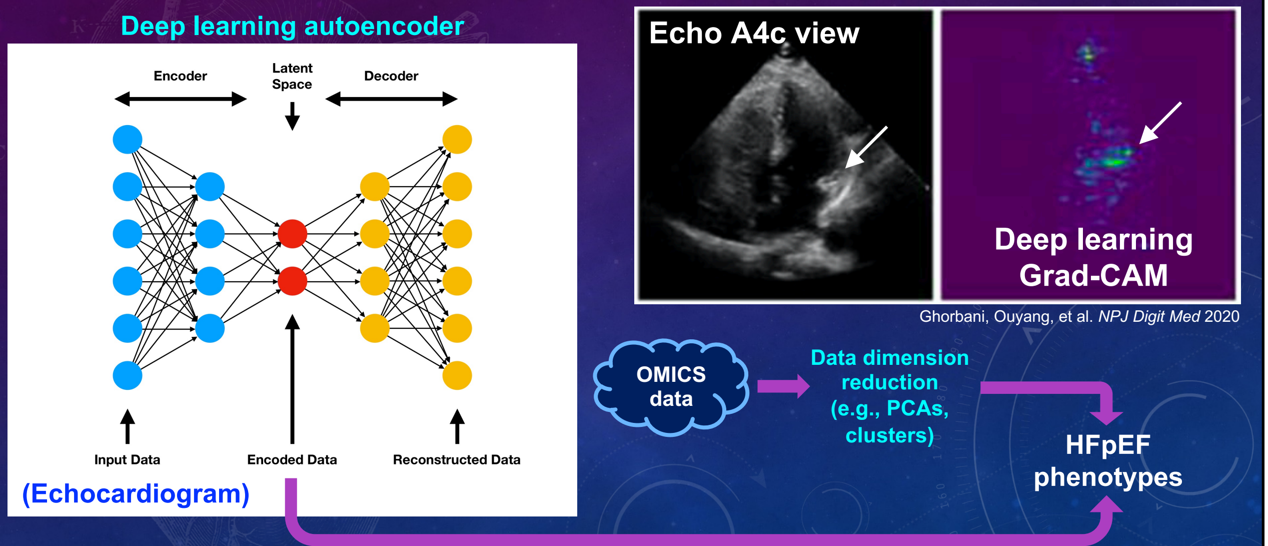
Heterogeneous HFpEF syndrome

Sub-phenotyping:
- Quantitative imaging
- Cardiac biopsy
- Liquid biopsy
- Exercise phenotyping
- Machine learning

Phenotype-specific treatment

50

Deep learning + multi-omics



51

Machine learning in HF: *Future directions*

- Differentiate types of learning tasks:
  - Mimic human behavior (machine replicates a task that humans do well)
  - Perform tasks that humans don't do well (find hidden meaning in data)
- Apply reinforcement learning, generational adversarial networks to healthcare problems
- Incorporate machine learning into clinical trials

52

Machine learning in HF: *Recommendations*

- Bigger data is NOT necessarily better data
  - Informative (orthogonal) features are key
  - Few precise features: better than lots of imprecise features
- Have a clear goal in mind at the onset of the study: resolve heterogeneity of complex phenotypes
- Think about validation/testing from study onset
- Deep learning: need to develop large repositories of images labeled by expert human readers

53

Take home points

- Data-driven analytics may be able to answer several unmet needs in echocardiography, especially:
  - Resolving the heterogeneity of complex CV syndromes
  - Early diagnosis of complex common and rare CV diseases
  - Automated measurements to improve workflow
  - AI guided ultrasound to "democratize" echocardiography
- There are 2 key types of machine learning:
  - Supervised learning
  - Unsupervised learning
- Machine learning is not perfect: *Don't use it blindly*
- Know how to properly evaluate studies that use ML

54


Shah Lab – Northwestern University
thank you
sanjiv.shah@northwestern.edu • http://www.hfpef.org • Twitter: @HFpEF

55