Why medicine is creating exciting new frontiers for machine learning

Mihaela van der Schaar

John Humphrey Plummer Professor of
Machine Learning, Artificial Intelligence and Medicine, University of Cambridge
Chancellor's Professor, University of California Los Angeles


vanderschaar-lab.com


mv472@cam.ac.uk

@MihaelaVDS 

linkedin.com/in/mihaela-van-der-schaar/ 

Why is AI/ML for healthcare different?

AI/ML has accomplished wonders …. on well-posed problems where the notion of a "solution" is well-defined and solutions are verifiable

Healthcare is different – problems are not well-posed, notion of a "solution" is often not well-defined and solutions are hard to verify

This presents enormous challenges – and also enormous opportunities

Our goal: New ML problem formalisms, models, techniques for revolutionizing healthcare!

Our group's mission:
Develop AI & ML to turn medicine from art to science

1) **deliver** bespoke medicine
2) **understand** the basis and trajectories of health and disease
3) **empower** healthcare professionals and patients
4) **inform and improve** clinical pathways, better utilize resources & reduce costs
5) **transform** population health and public health policy
6) **enable** new discoveries – clinical, therapeutics

Our group's agenda:
New ML aimed at revolutionizing healthcare

- AutoML
- Time-series forecasting
- Causal machine learning & Causal effect inference
- RL & Inverse reinforcement learning
- ML - Interpretability & explainability
- ML - Trustworthiness- confidence estimates
- Statistical ML
- Etc. etc.

https://www.vanderschaar-lab.com/van-der-schaar-lab-at-neurips-2020-9-papers-accepted/

Developing Clinical Analytics: Challenges

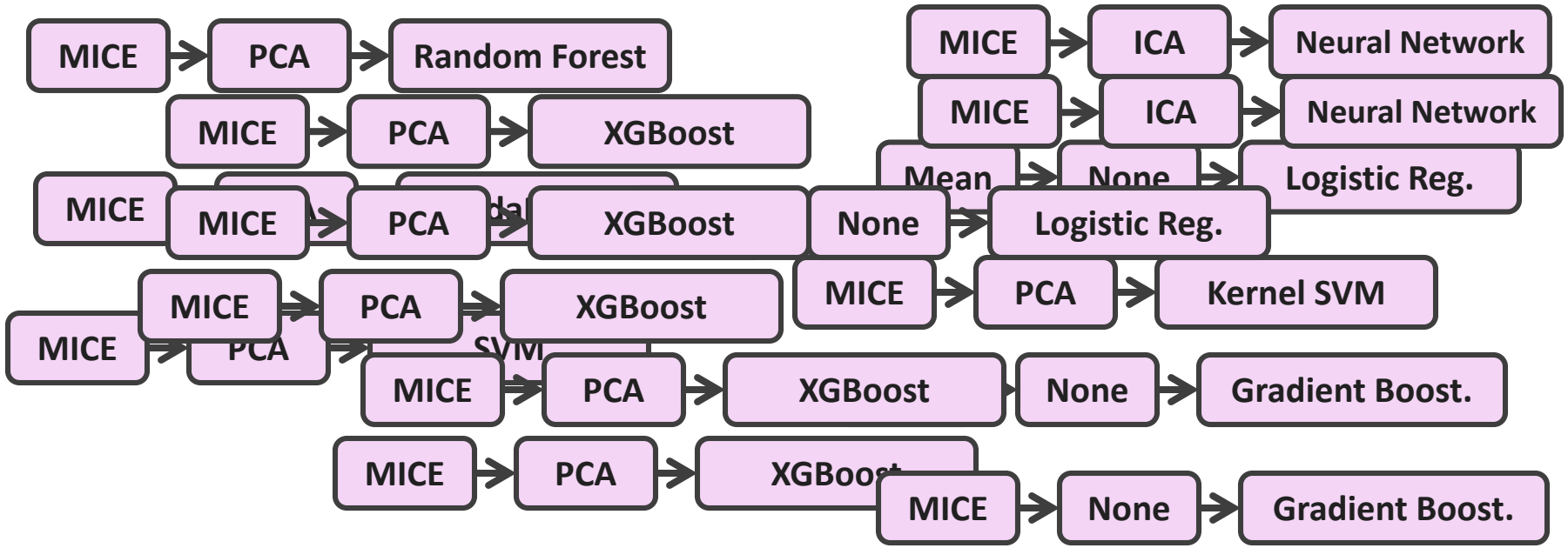
- Model
  - No “one-size-fits-all” solution
    - Which ML model to choose?
  - Reproducibility
  - Interpretability, explainability
  - Trustworthiness, uncertainty estimates

- **AutoML – what is different in healthcare?**

Automating risk prediction modeling

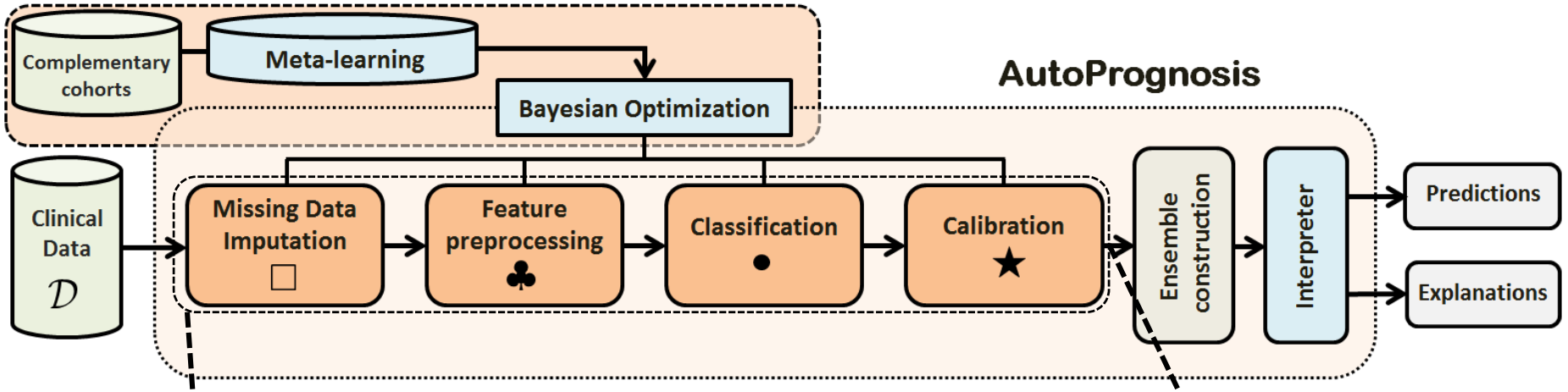
- **Goal:** develop holistic view of patient health
  - many risk scores for many clinical conditions
- **Challenge:** a HUGE design space!

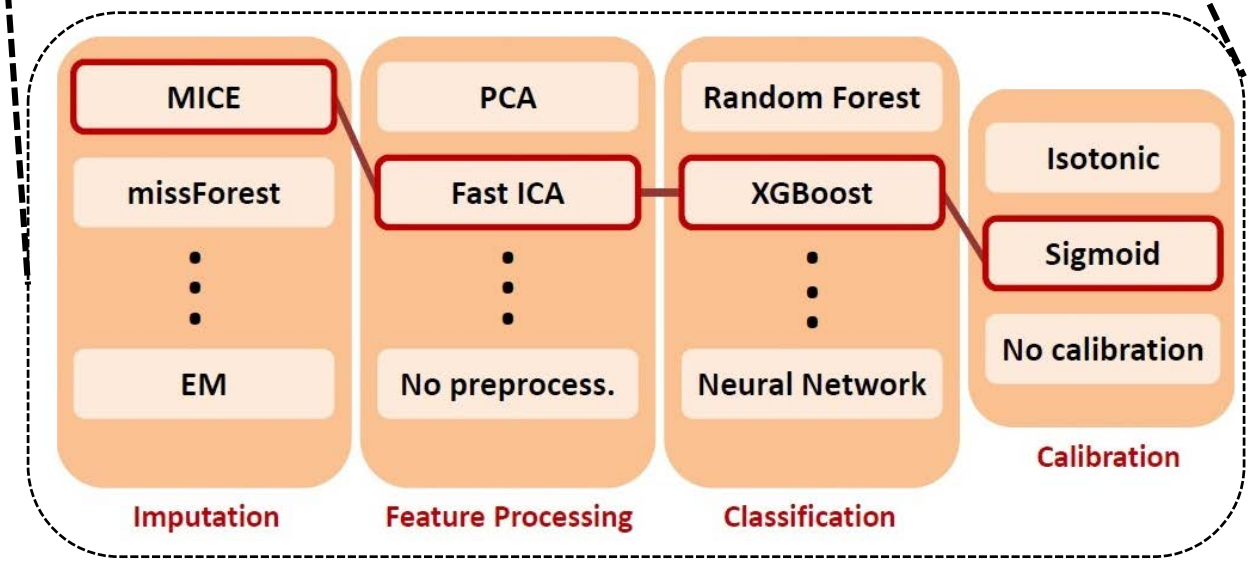
Many possible pipelines!



AutoPrognosis [Alaa & vdS, ICML 2018]:

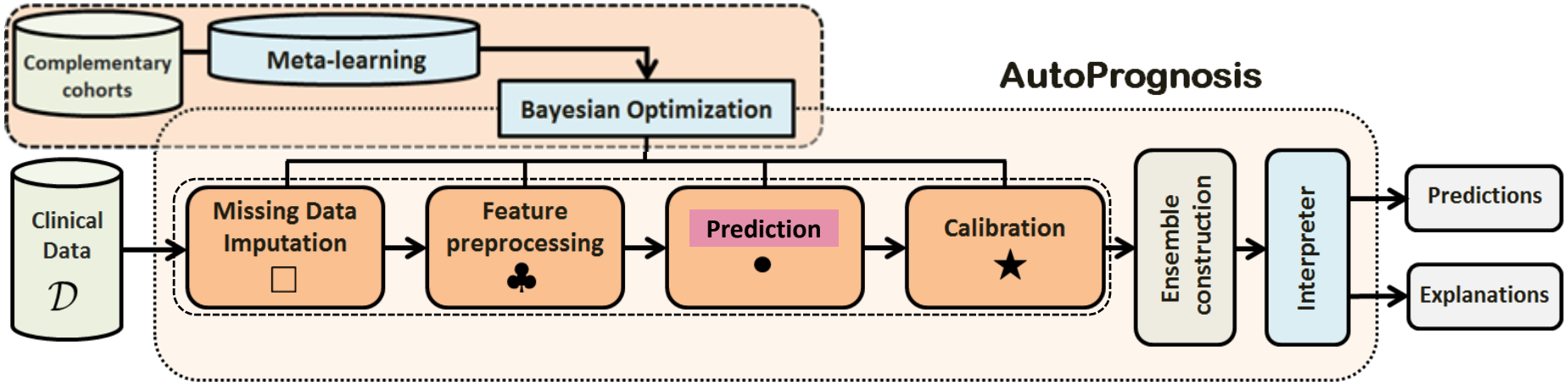
A tool for crafting Clinical Scores





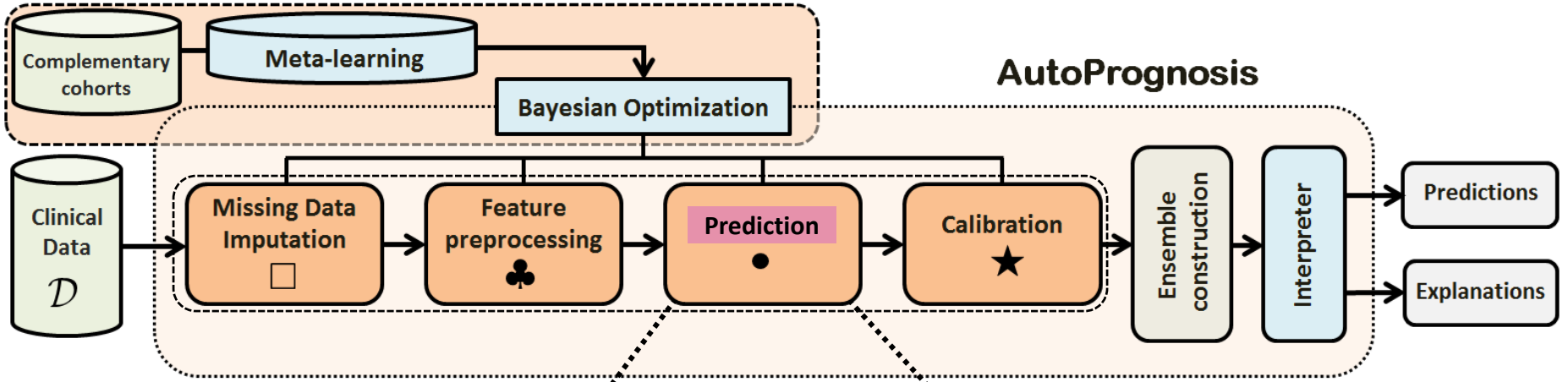
We need an entire pipeline!

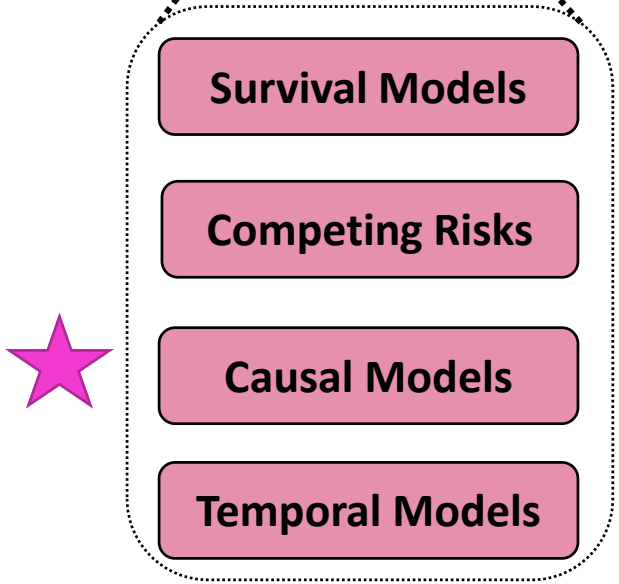
AutoPrognosis in practice



Cardiovascular - ICML 2018
Cystic Fibrosis - Scientific Reports - 2018
UK Biobank - Plos One 2018
Breast Cancer – 2019
Covid - 2020

AutoML: Beyond classification





Lee, Alaa, Zame, vdS, AISTATS 2019

Alaa, vdS, NIPS 2017
Lee, Yoon, vdS, TBME 2019

Alaa, vdS, ICML 2019

Zhang, Jarett, vdS, AISTATS 2020

Personalized treatment recommendations

Goal: estimate the effect of a treatment/intervention on an individual

Bob

Diagnosed with Disease X

Which treatment is best for Bob?

Challenge: treatment effects are often heterogeneous

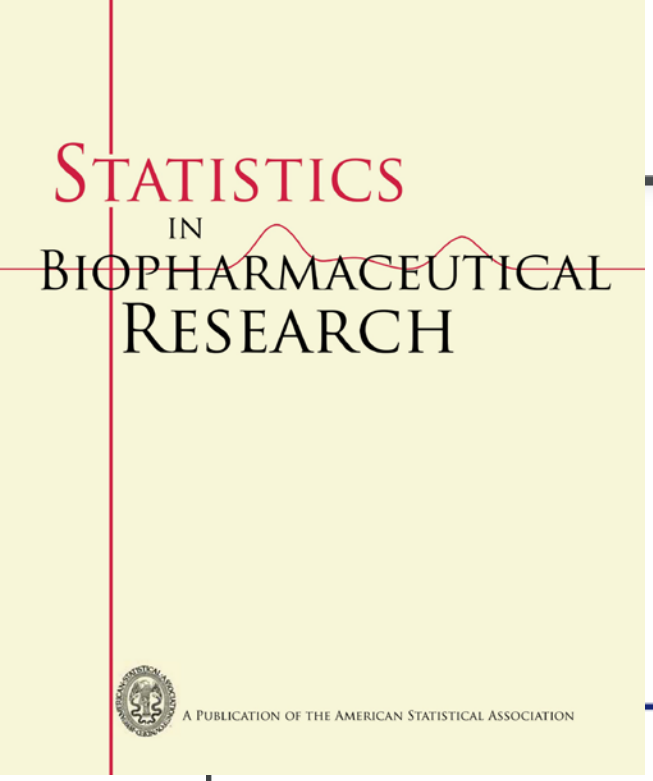
Personalized therapeutics: Adaptive Clinical Trials

Randomized Control Trials

Population-level

Adaptive Clinical Trials

[Atan, Zame, vdS, AISTATS 2019]
[Shen, Villar, vdS, ICML 2020]
[Lee, Zhang, Shen, Zame, vdS, NeurIPS 2020]
[Bica, Jordon, Alaa, vdS, NeurIPS 2020]
[Lee, Shen, Zame, vdS, AISTATS 2021]



Taylor & Francis
Taylor & Francis Group

Statistics in Biopharmaceutical Research

ISSN: (Print) 1946-6315 (Online) Journal homepage: https://www.tandfonline.com/loi/usbr20

Machine learning for clinical trials in the era of COVID-19

William R. Zame, Ioana Bica, Cong Shen, Alicia Curth, Hyun-Suk Lee, Stuart Bailey, James Weatherall, David Wright, Frank Bretz & Mihaela van der Schaar

To cite this article: William R. Zame, Ioana Bica, Cong Shen, Alicia Curth, Hyun-Suk Lee, Stuart Bailey, James Weatherall, David Wright, Frank Bretz & Mihaela van der Schaar (2020): Machine learning for clinical trials in the era of COVID-19, Statistics in Biopharmaceutical Research, DOI: 10.1080/19466315.2020.1797867

Personalized therapeutics:
Individualized treatment effects

[Alaa, vdS, NIPS 2017, ICML 2018, ICML 2019]
[Yoon, Jordon, vdS, ICLR 2018]
[Lim, Alaa, vdS, NeurIPS 2018]
[Zhang, Bellot, vdS, AISTATS 2020]
[Bica, Alaa, vdS, ICLR 2020, ICML 2020]
[Curth, vdS, AISTATS 2021]

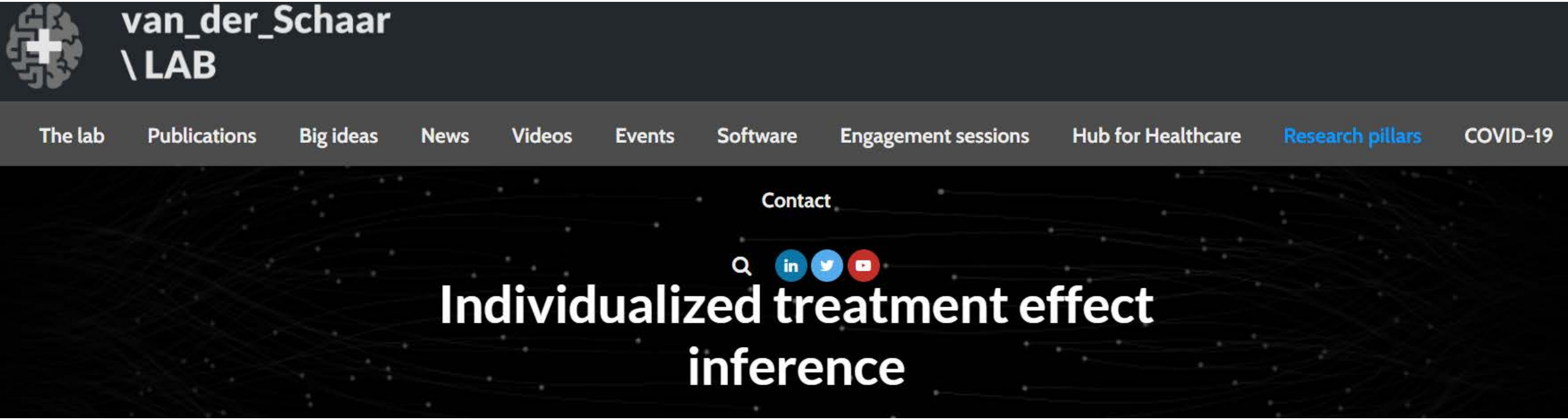
Machine Learning:
Individualized Treatment Effects

Patient-centric

Real-world observational data
Scalable & adaptive implementation
Fast deployment
Cost-effective

Want to learn more?

https://www.vanderschaar-lab.com/individualized-treatment-effect-inference/

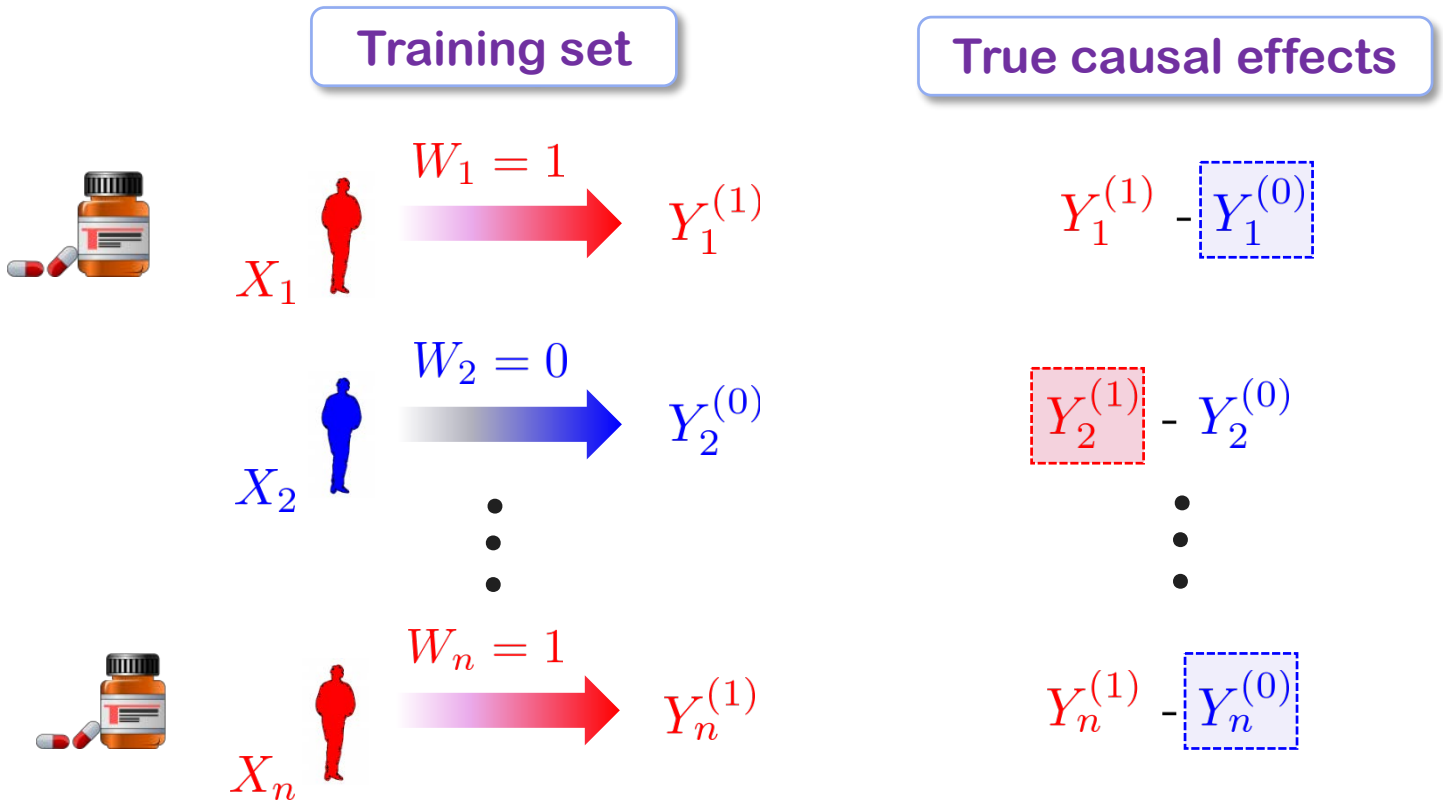


Causal effect inference: a complicated ML problem

Counterfactuals – answering "What if" questions

Challenge: we never observe counterfactual outcomes!

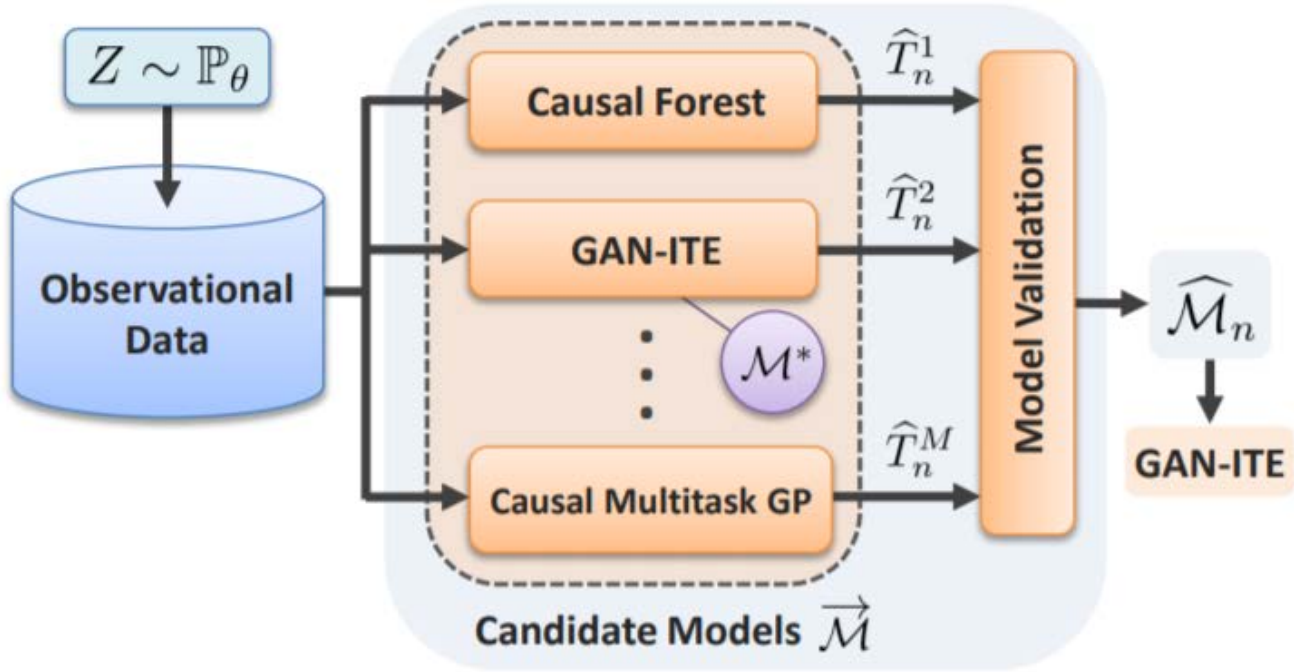
Not a simple, "supervised" ML problem - no explicit label!



How to select the best model?
Automating causal inference! [Alaa&vdS, ICML 2019]

- Numerous causal models in ICML, NeurIPS and ICLR
- Goal: Select best model for each observational study

	BNN	ICML 2016
→	CMGP	NIPS 2017
	TARNet	ICML 2017
	CFR Wass.	ICML 2017
	CFR MMD	ICML 2017
→	NSGP	ICML 2018
→	GANITE	ICLR 2018
	SITE	NIPS 2018
	BART	
	Causal Forest	



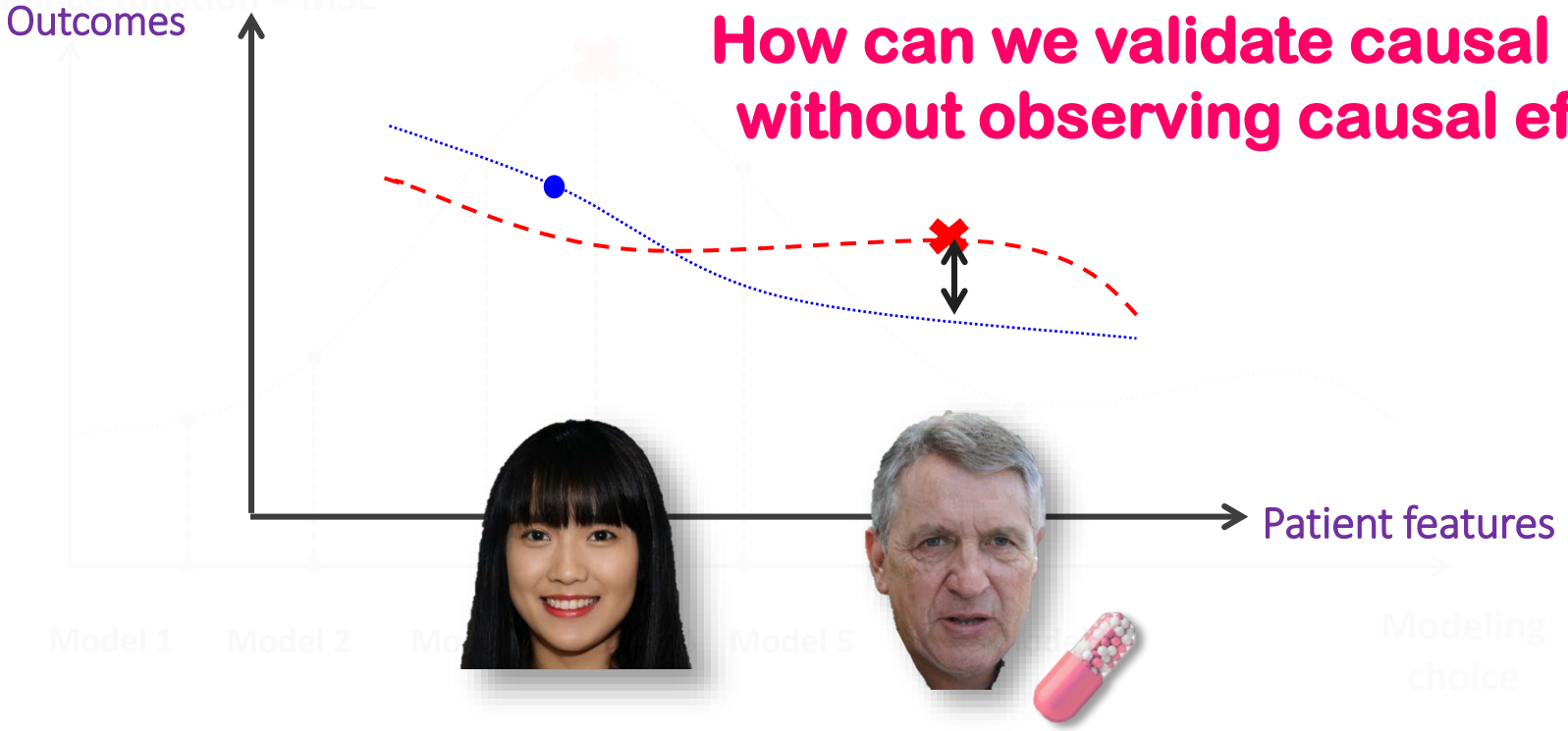
Key Challenge: How to do cross-validation?

https://www.vanderschaar-lab.com/individualized-treatment-effect-inference/

Automating causal inference! [Alaa&vdS, ICML 2019]

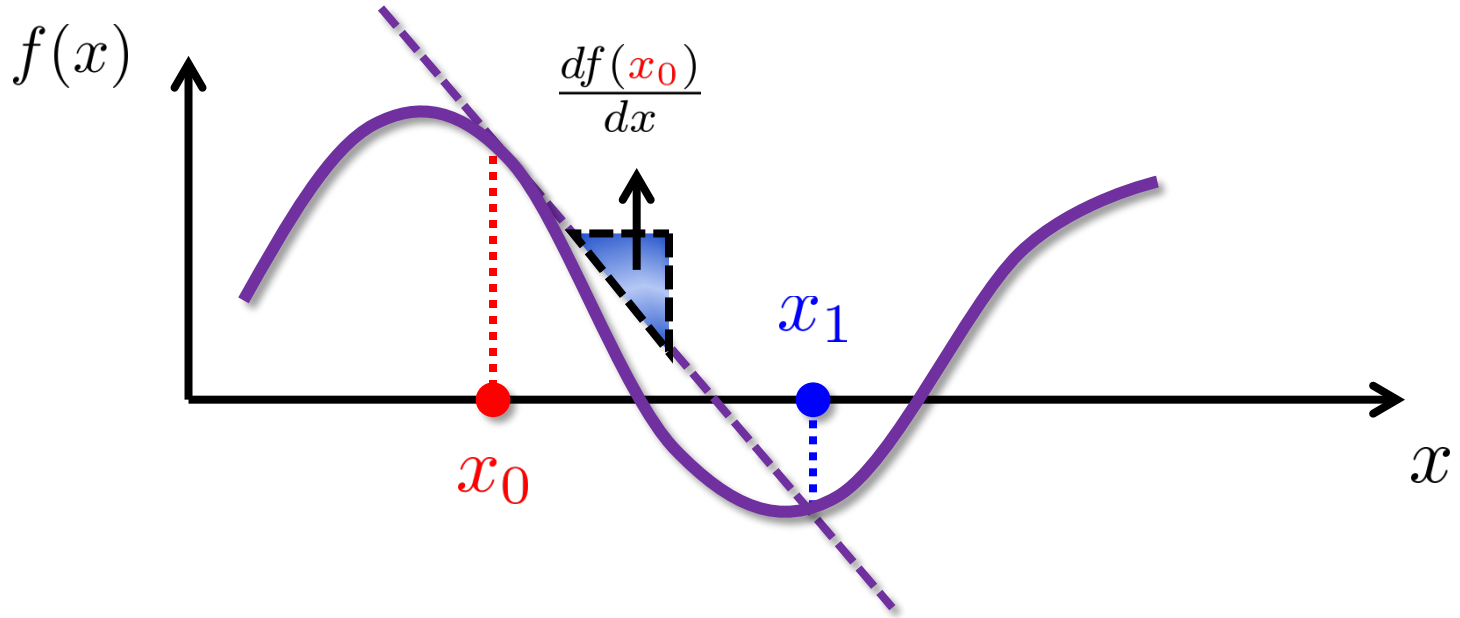
- **PEHE**: MSE of a causal model

$$\ell_\theta(\widehat{T}) = \|T(X) - \widehat{T}(X)\|_\theta^2$$



Taylor series approximation

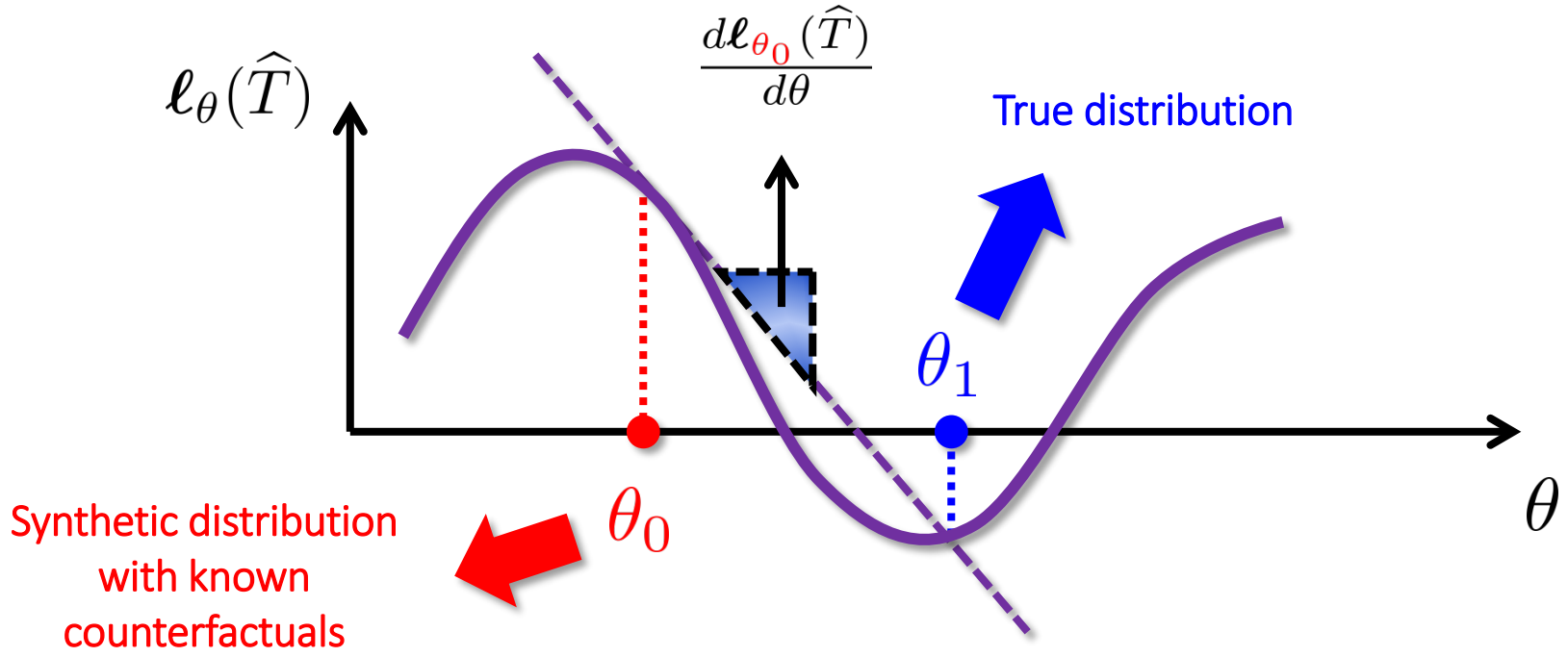
- Value of a function at a given input can be predicted using its value and (higher-order derivatives) at a proximal input



$$f(x_1) = f(x_0) + (x_1 - x_0)\frac{df(x_0)}{dx} + \frac{1}{2!}(x_1 - x_0)^2\frac{d^2f(x_0)}{dx^2} + \dots$$

Analogy with Taylor series approximation

- Performance of a causal inference model is a functional of data-generating distribution $\mathbb{P}_\theta$



Functional calculus: von-Mises expansion

- A "distributional" analog of Taylor expansion [Fernholz, 1983]

$$\ell_{\theta_1}(\widehat{T}) = \ell_{\theta_0}(\widehat{T}) + \int \dot{\ell}^{(1)}_{\theta_0}(z;\widehat{T}) d(\mathbb{P}_{\theta_1} - \mathbb{P}_{\theta_0}) + \frac{1}{2!}\int \dot{\ell}^{(2)}_{\theta_0}(z;\widehat{T}) d(\mathbb{P}_{\theta_1} - \mathbb{P}_{\theta_0})^2 + \dots$$

- **Influence Functions ↔ Derivatives**

We can predict performance of a causal inference model using the influence functions (IF) of its loss on a "similar" synthetic dataset.

How to estimate a model's performance?

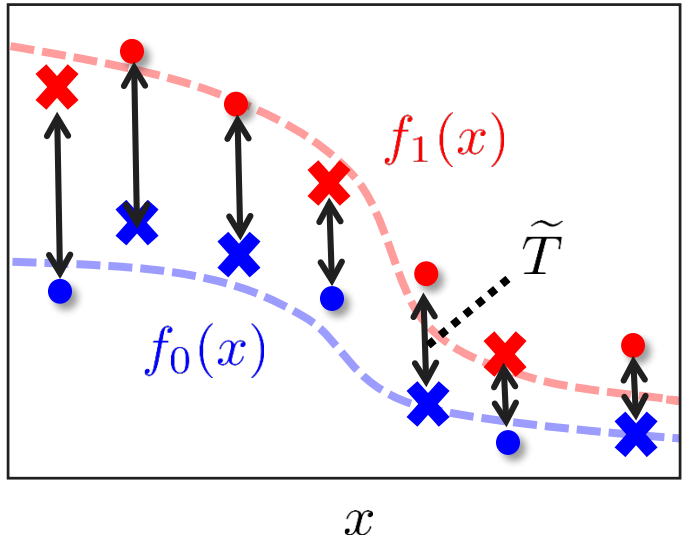
- **Synthesize counterfactuals!**

Step 1: Plug-in estimation

- Plug-in model $\widetilde{T}$
- Plug-in PEHE loss $\ell_{\tilde{\theta}}(\widehat{T})$

Step 2: Bias correction

$$\ell_{\theta}(\widehat{T}) = \ell_{\tilde{\theta}}(\widehat{T}) + \int \dot{\ell}^{(1)}_{\tilde{\theta}}(z;\widehat{T}) d\mathbb{P}_{\theta}$$



Estimating a causal model's performance

- First-order "Taylor approximation"

Inaccessible empirical measure

True distribution

$$\ell_\theta(\widehat{T}) \approx \ell_{\tilde{\theta}}(\widehat{T}) + \int \dot{\ell}^{(1)}_{\tilde{\theta}}(z;\widehat{T})\, d(\mathbb{P}_\theta - \mathbb{P}_{\tilde{\theta}})$$

Synthetic (Plug-in) distribution

Influence function

Accessible empirical measure

AutoML for causal inference! [Alaa&vdS, ICML 2019]

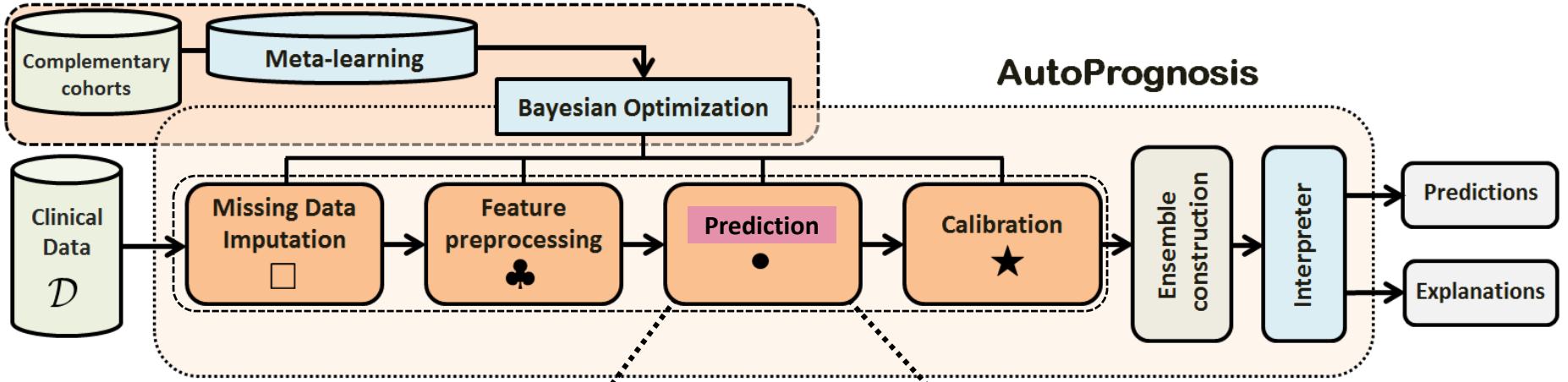
- Average performance on the 77 benchmark datasets.
- No absolute single winner on all datasets.
- Our IF-based selection is better than any single model.

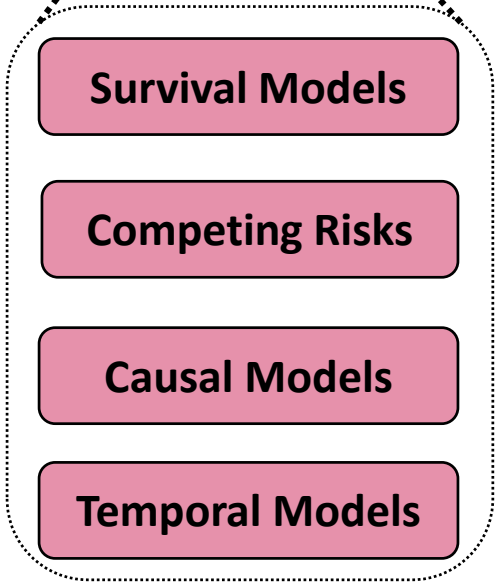
Method	% Winner
BNN	3%
CMGP	12%
NSGP	17%
TARNet	8%
CFR Wass.	9%
CFR MMD	12%
GANITE	7%
SITE	7%
BART	15%
C. Forest	7%
AutoML IF-based	72%

https://www.vanderschaar-lab.com/publications/
ML-subfield: Causal Inference

https://www.vanderschaar-lab.com/individualized-treatment-effect-inference/

AutoML: Beyond classification





Lee, Alaa, Zame, vdS, AISTATS 2019

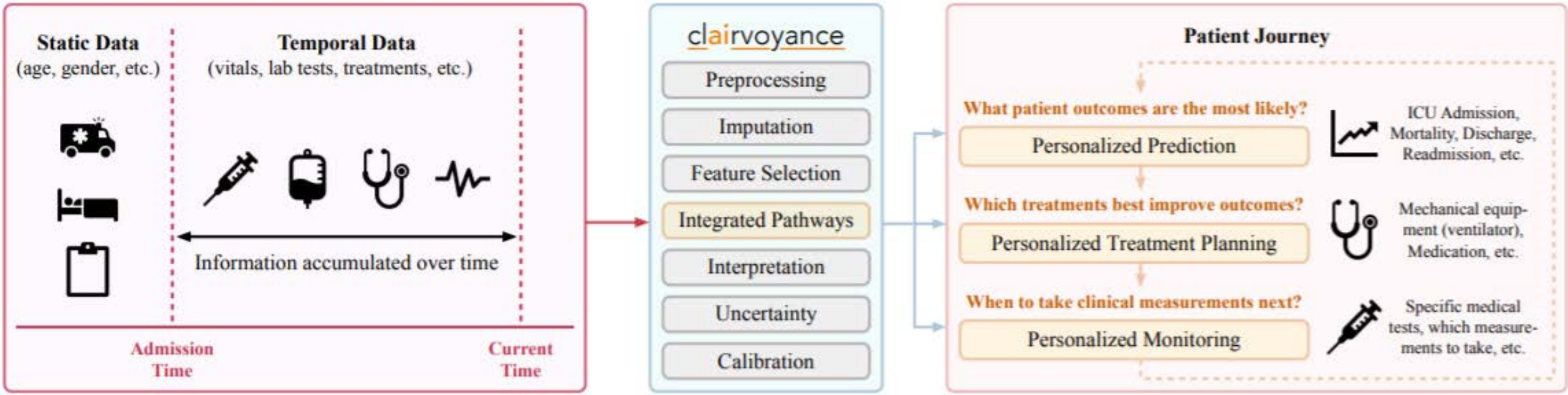
Alaa, vdS, NIPS 2017
Lee, Yoon, vdS, TBME 2019

Alaa, vdS, ICML 2019

Zhang, Jarett, vdS, AISTATS 2020

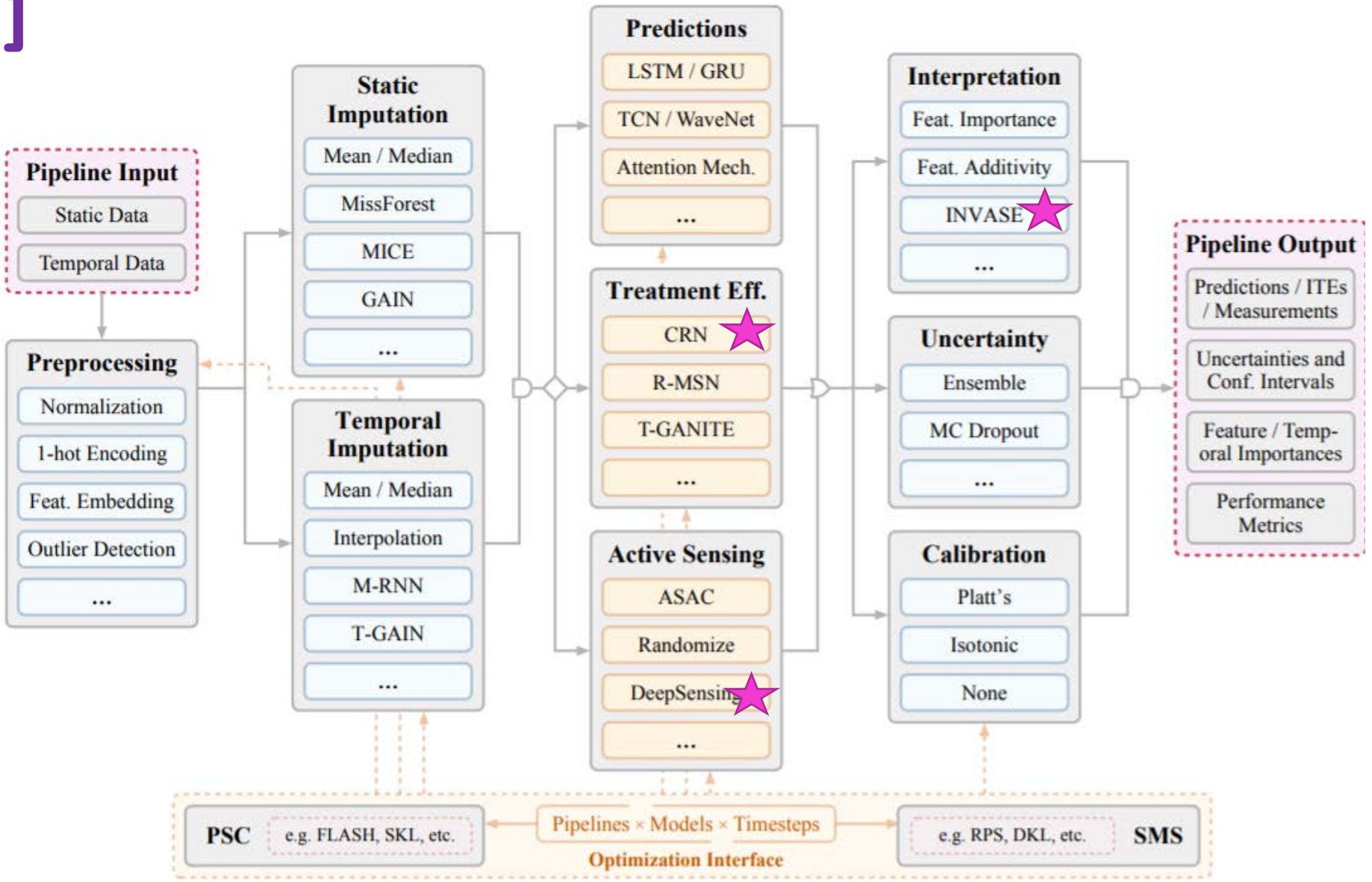
Clairvoyance [ICLR 2021]

Clairvoyance:
A unified, end-to-end pipeline for clinical-decision support
[Jarrett, Yoon, Bica, Qian, Ercole, vdS, ICLR 2021]

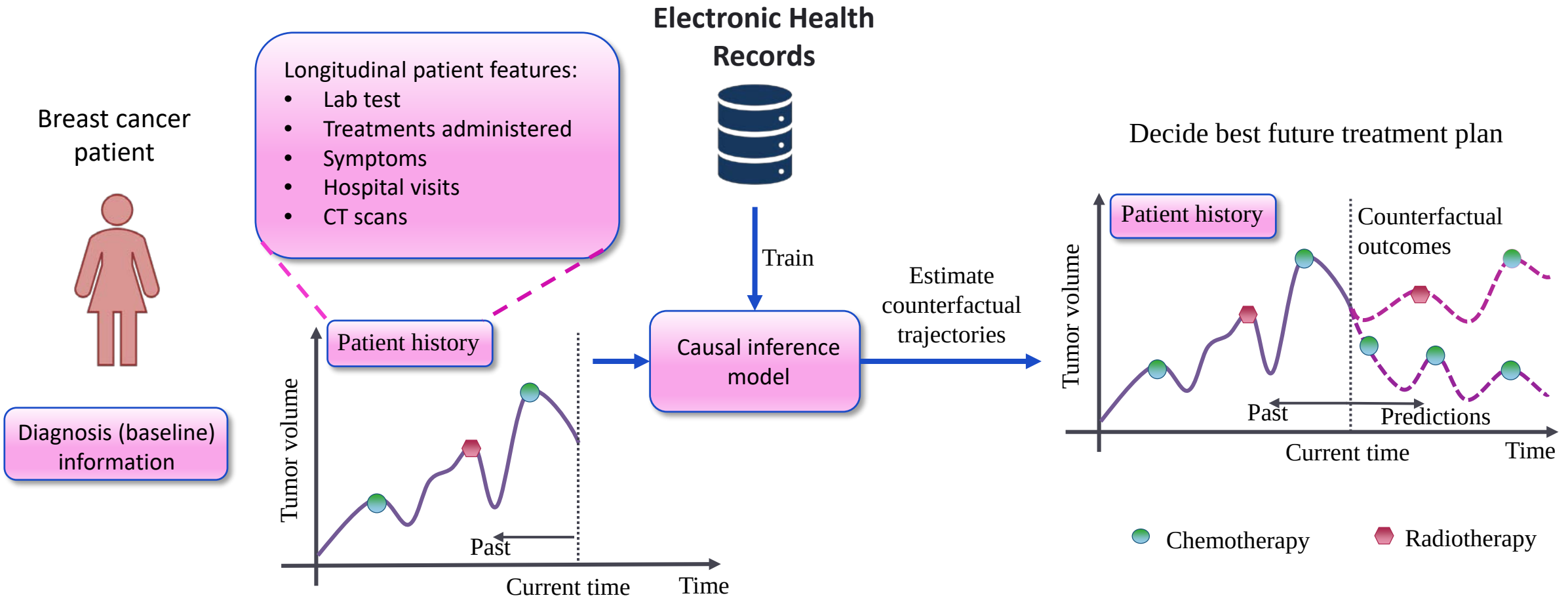


Important milestone in ML for healthcare: REPRODUCIBILITY!

Clairvoyance:
A unified, end-to-end pipeline for clinical-decision support [ICLR 2021]



Individualized Treatment effects over time

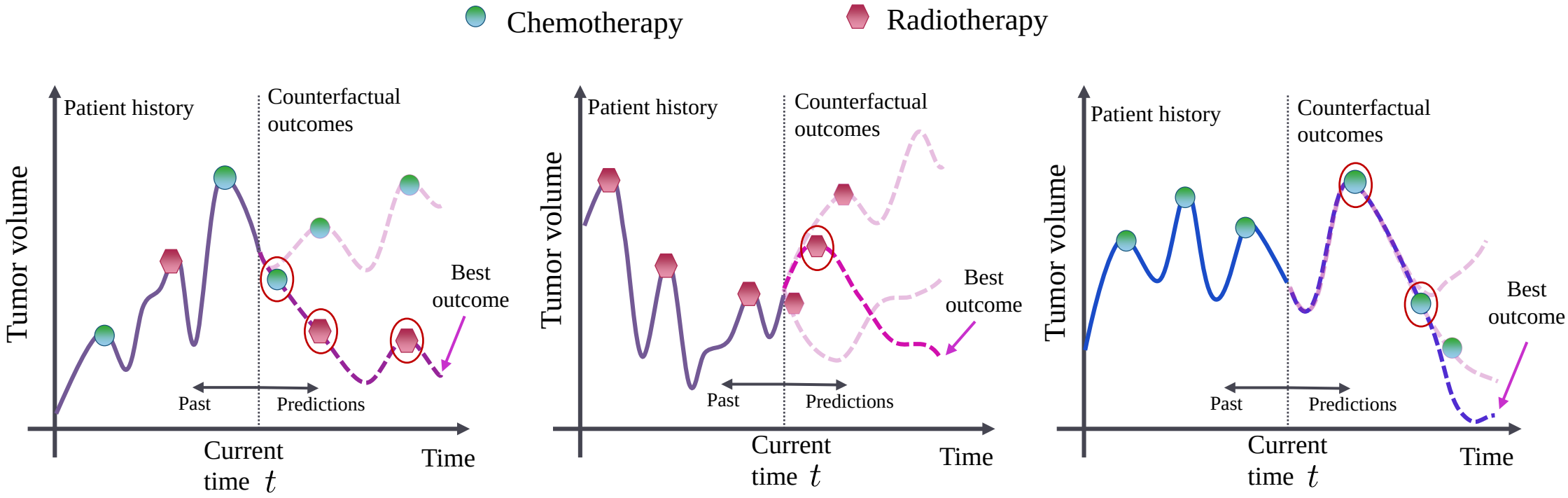


Individualized Treatment effects over time

How to treat?

When to give treatment?

When to stop treatment?

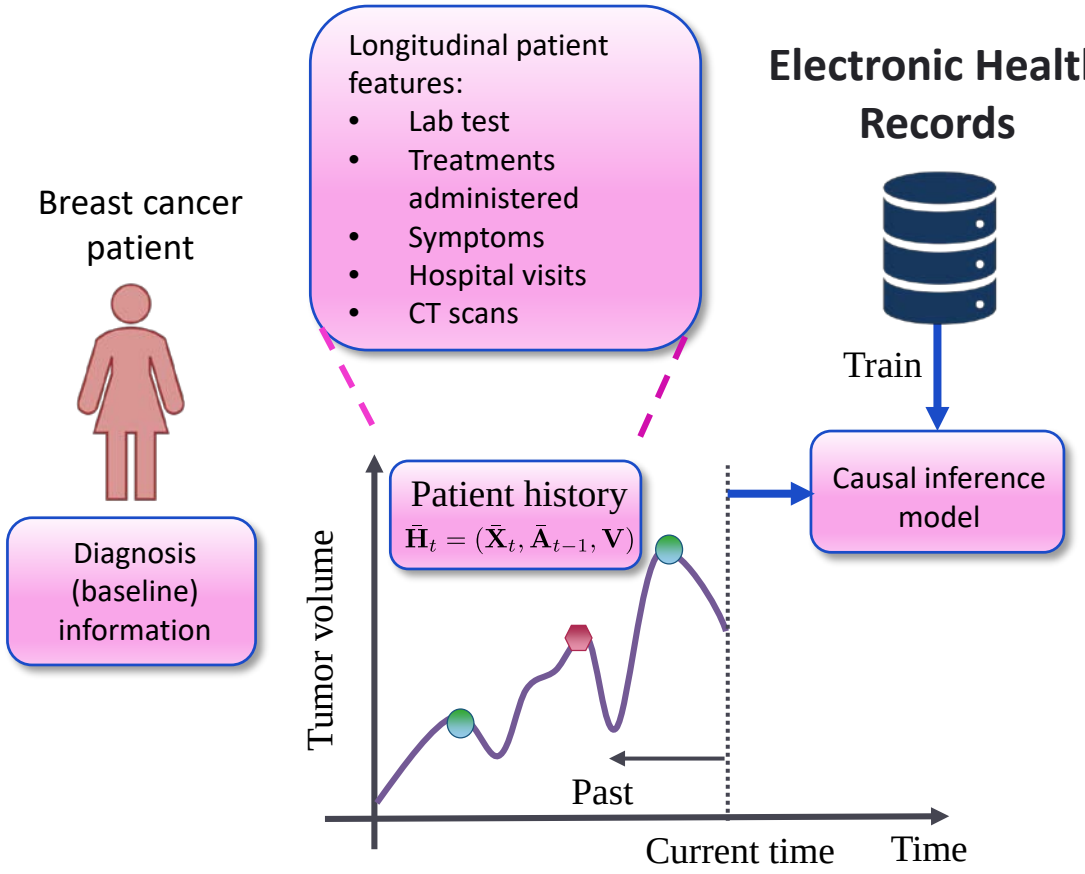


(a) Decide treatment plan

(b) Decide optimal time of treatment

(c) Decide when to stop treatment

Causal effect inference based on longitudinal patient observational data



Longitudinal patient observational data

- Time-dependent patient features: $\bar{\mathbf{X}}_t = (\mathbf{X}_1, \ldots \mathbf{X}_t)$
- Time-dependent treatments: $\bar{\mathbf{A}}_t = (\mathbf{A}_1, \ldots \mathbf{A}_t)$ where

$$\mathbf{A}_t \in \{A_1, \ldots, A_K\}$$

- Static patient features: $\mathbf{V}$

Patient history: $\bar{\mathbf{H}}_t = (\bar{\mathbf{X}}_t, \bar{\mathbf{A}}_{t-1}, \mathbf{V})$

Observed (factual) outcome for treatment $\mathbf{A}_t$ given patient history $\bar{\mathbf{H}}_t$: $\mathbf{Y}_{t+1}$

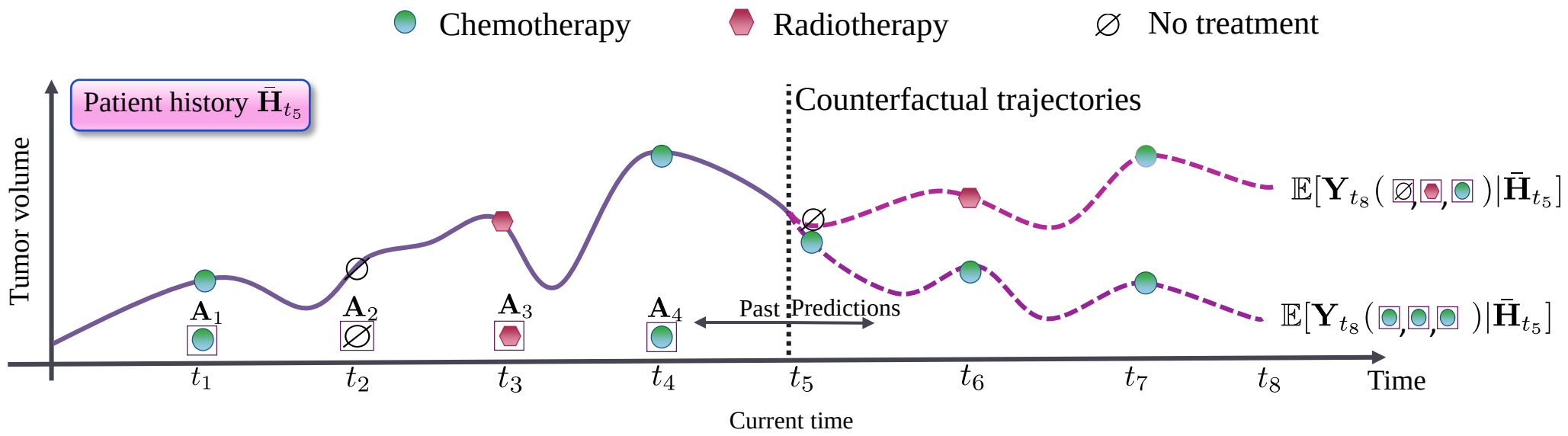
Estimate potential outcomes [Rubin (1978),Neyman (1923),Robins & Hernan (2008)]

Potential outcomes under planned sequence of future treatments:

$$\mathbb{E}[\mathbf{Y}_{t+\tau}(\bar{\mathbf{a}}(t, t+\tau-1))|\bar{\mathbf{H}}_t]$$

Assumptions

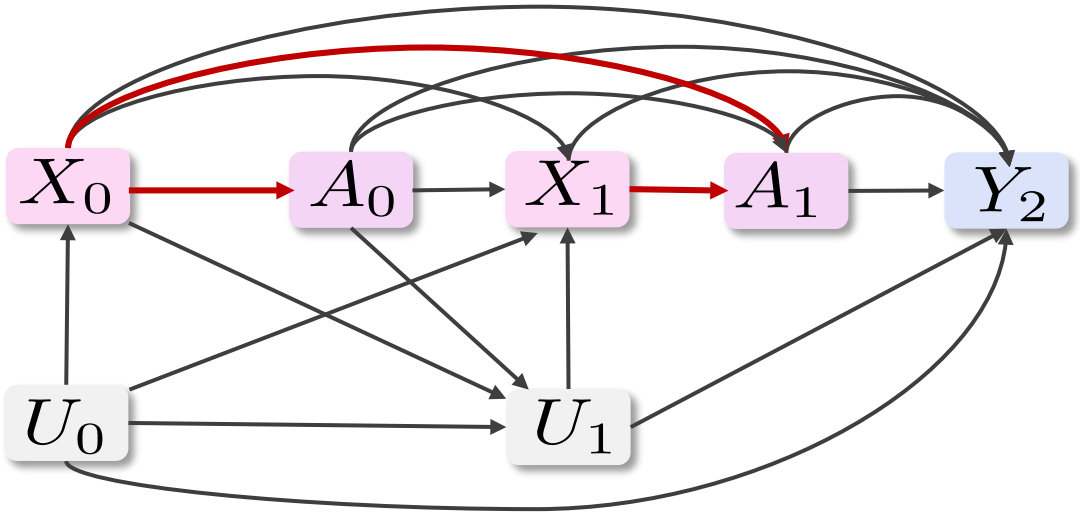
- **Sequential overlap:** $P(\mathbf{A}_t = \mathbf{a}_t \mid \bar{\mathbf{H}}_t = \bar{\mathbf{h}}_t) > 0, \forall \mathbf{a}_t, \forall t$
- **Sequential strong ignorability:** $\mathbf{Y}(\bar{\mathbf{a}}_{\geq t}) \perp\!\!\!\perp \mathbf{A}_t \mid \bar{\mathbf{H}}_t$, for all possible treatment plans $\bar{\mathbf{a}}_{\geq t}$ and $\forall t$



Challenges in using longitudinal observational data for estimating individualized outcomes

The patient history $\bar{\mathbf{H}}_t = (\bar{\mathbf{X}}_t, \bar{\mathbf{A}}_{t-1}, \mathbf{V})$ contains time-dependent confounders which bias the treatment assignment $\mathbf{A}_t$ in the observational dataset.

Patient covariates - affected by past treatments which then influence future treatments and outcomes



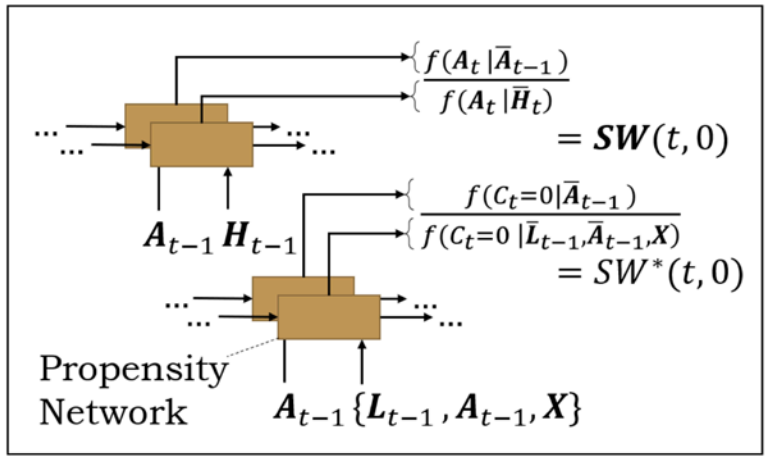
Bias from time-dependent confounders.

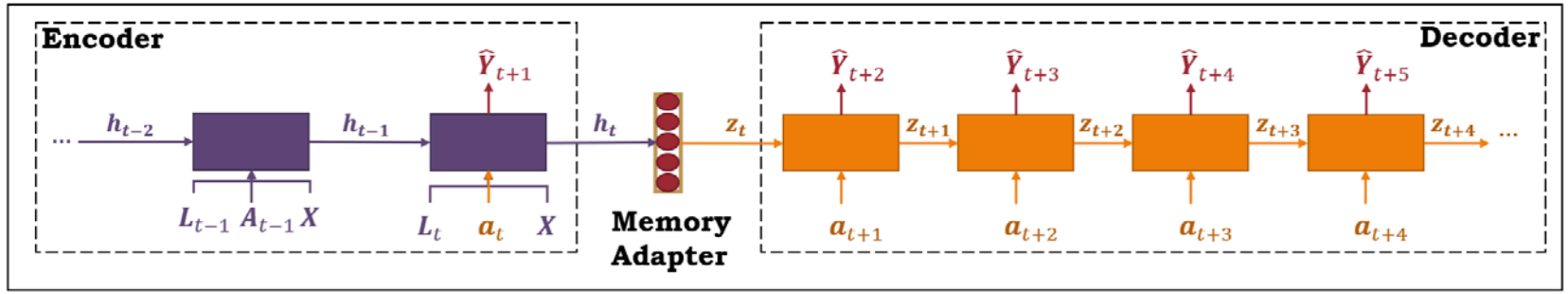
Handling time-dependent confounding bias

Inverse probability of treatment weighting

→ **Marginal Structural Models** [Robins, Hernan, Brumback, Epidemiology 2000]

→ **Recurrent Marginal Structural Networks** [Lim, Alaa, van der Schaar, NeurIPS 2018]





$$e(i,t,\tau) = \mathbf{S}\tilde{\mathbf{W}}_i(t,\tau-1) \times \tilde{SW}_i^*(t,\tau-1) \times \|\mathbf{Y}_{t+\tau,i} - g(\tau, a(t,\tau-1), \bar{\mathbf{H}}_t)\|^2$$

$$\mathbf{SW}(t,\tau) = \prod_{n=t}^{t+\tau} \frac{f(\mathbf{A}_n|\bar{\mathbf{A}}_{n-1})}{f(\mathbf{A}_n|\bar{\mathbf{H}}_n)} = \prod_{n=t}^{t+\tau} \frac{\prod_{k=1}^{\Omega_a} f(A_n(k)|\bar{\mathbf{A}}_{n-1})}{\prod_{k=1}^{\Omega_a} f(A_n(k)|\bar{\mathbf{H}}_n)}$$

Handling time-dependent confounding bias

Inverse probability of treatment weighting

→ Marginal structural models [Robins, Hernan, Brumback, Epidemiology 2000]

→ Recurrent Marginal Structural Networks [Lim, Alaa, van der Schaar, NeurIPS 2018]

Numerically unstable

High variance

Representation Learning

→ Counterfactual Recurrent Network [Bica, Alaa, Jordon, van der Schaar, ICLR 2020]

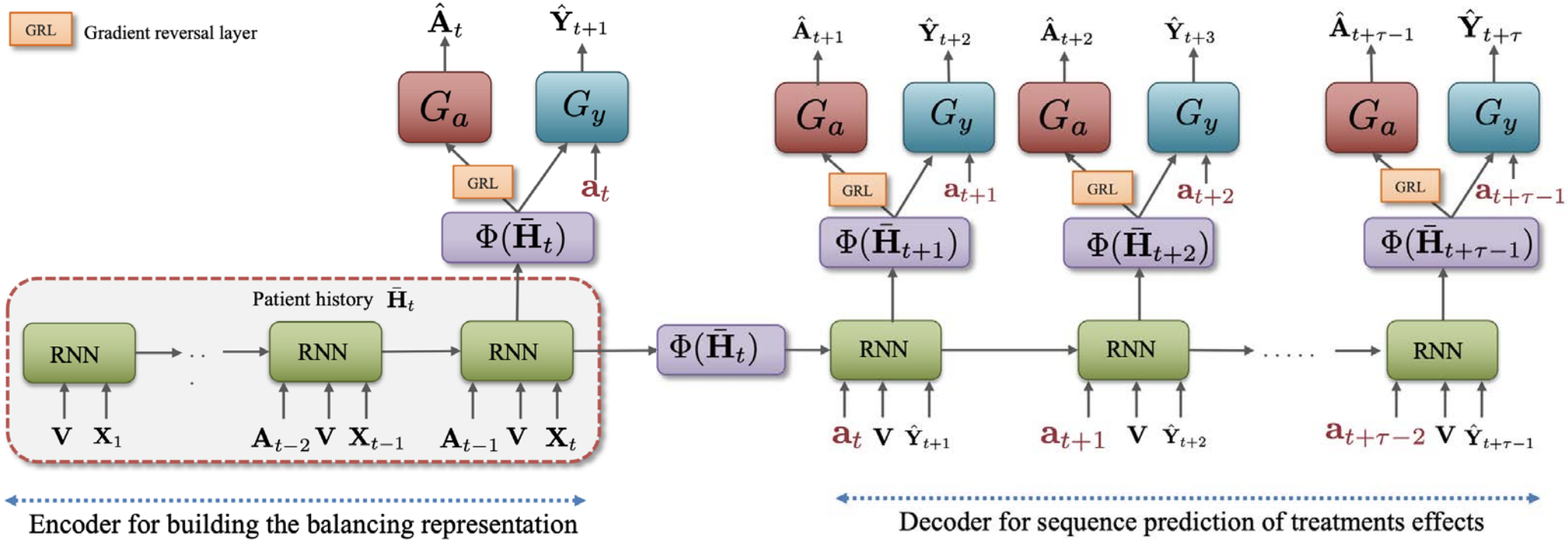
$$P(\Phi(\bar{\mathbf{H}}_t) \mid \mathbf{A}_t = A_1) = \cdots = P(\Phi(\bar{\mathbf{H}}_t) \mid \mathbf{A}_t = A_K)$$

Balanced representations/
Treatment invariant representations

Counterfactual Recurrent Network

[Bica, Alaa, Jordon & van der Schaar, ICLR 2020]

- Builds treatment invariant representations using domain adversarial training [Ganin et al., 2016].
- Estimates counterfactual trajectories using sequence-to-sequence architecture.



Adversarially balanced representations

- **Treatment (domain) loss**

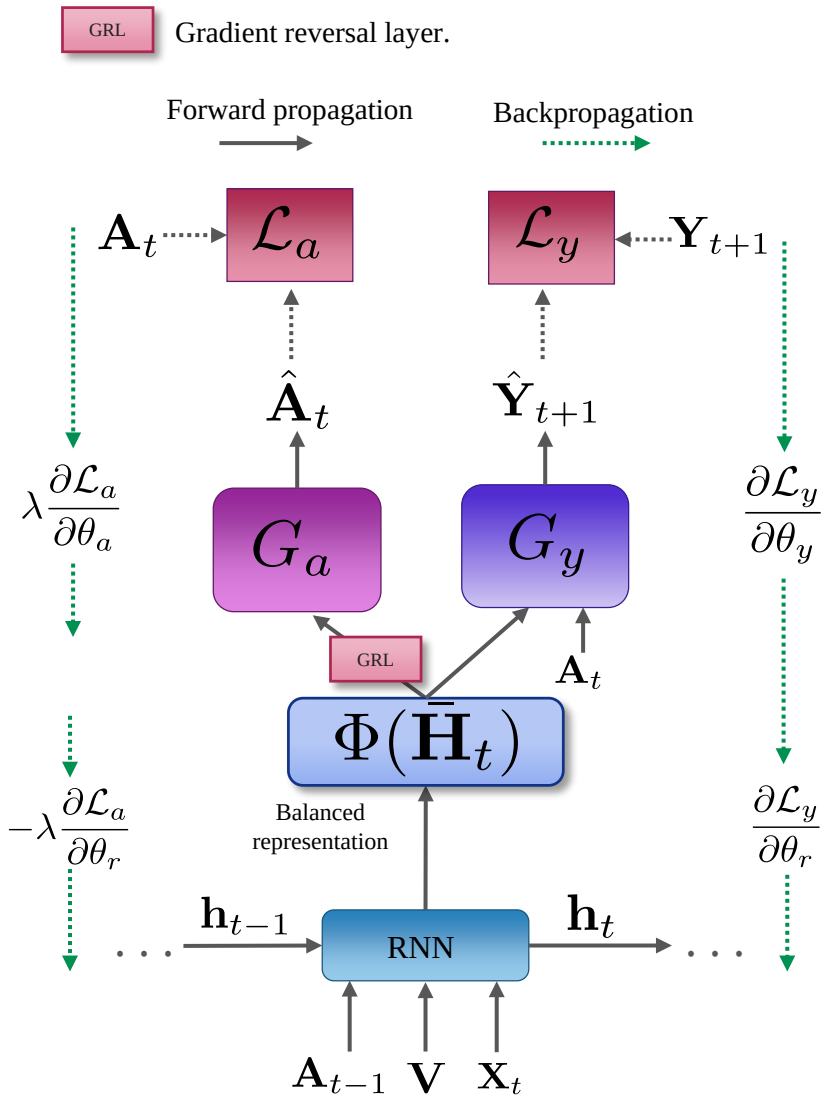
$$\mathcal{L}_{t,a}(\theta_r, \theta_a) = -\sum_{j=1}^{K} \mathbb{I}_{\{\mathbf{a}_t = a_j\}} \log(G_a^j(\Phi(\bar{\mathbf{H}}_t; \theta_r); \theta_a))$$

- **Outcome loss**

$$\mathcal{L}_{t,y}(\theta_r, \theta_y) = \|\mathbf{Y}_{t+1} - (G_y(\Phi(\bar{\mathbf{H}}_t; \theta_r), \theta_y))\|^2$$

- **Overall loss at timestep t**

$$\mathcal{L}_t(\theta_r, \theta_y, \theta_a) = \mathcal{L}_{t,y}(\theta_r, \theta_y) - \lambda \mathcal{L}_{t,a}(\theta_r, \theta_a)$$



Experiments using model of tumour growth

- Tumour volume $t+1$ days after diagnosis

$$V(t+1) = \Big(1 + \underbrace{\rho \log\big(\frac{K}{V(t)}\big)}_{\text{Tumor growth}} - \underbrace{\beta_c C(t)}_{\text{Chemotherapy}} - \underbrace{\big(\alpha_r d(t) + \beta_r d(t)^2\big)}_{\text{Radiotherapy}} + \underbrace{e_t}_{\text{Noise}}\Big) V(t)$$

- Chemotherapy and radiotherapy treatment assignments

$$p_c(t) = \sigma\left(\frac{\gamma_c}{D_{\max}}(\bar{D}(t) - \delta_c)\right) \qquad p_r(t) = \sigma\left(\frac{\gamma_r}{D_{\max}}(\bar{D}(t) - \delta_r)\right)$$

- Parameters γ_c and γ_r control the amount of time-dependent confounding.

Results: Treatment plans – how and when to treat?

Chemotherapy application

$$\begin{aligned}
\mathbf{Y}_{t+\tau} \quad &| \quad \mathbf{a}_t = A_1, \mathbf{a}_{t+1} = A_0, \ldots \mathbf{a}_{t+\tau-1} = A_0, \bar{\mathbf{H}}_t \\
\mathbf{Y}_{t+\tau} \quad &| \quad \mathbf{a}_t = A_0, \mathbf{a}_{t+1} = A_1, \ldots \mathbf{a}_{t+\tau-1} = A_0, \bar{\mathbf{H}}_t \\
&\ldots \\
\mathbf{Y}_{t+\tau} \quad &| \quad \mathbf{a}_t = A_0, \mathbf{a}_{t+1} = A_0, \ldots \mathbf{a}_{t+\tau-1} = A_1, \bar{\mathbf{H}}_t
\end{aligned}$$

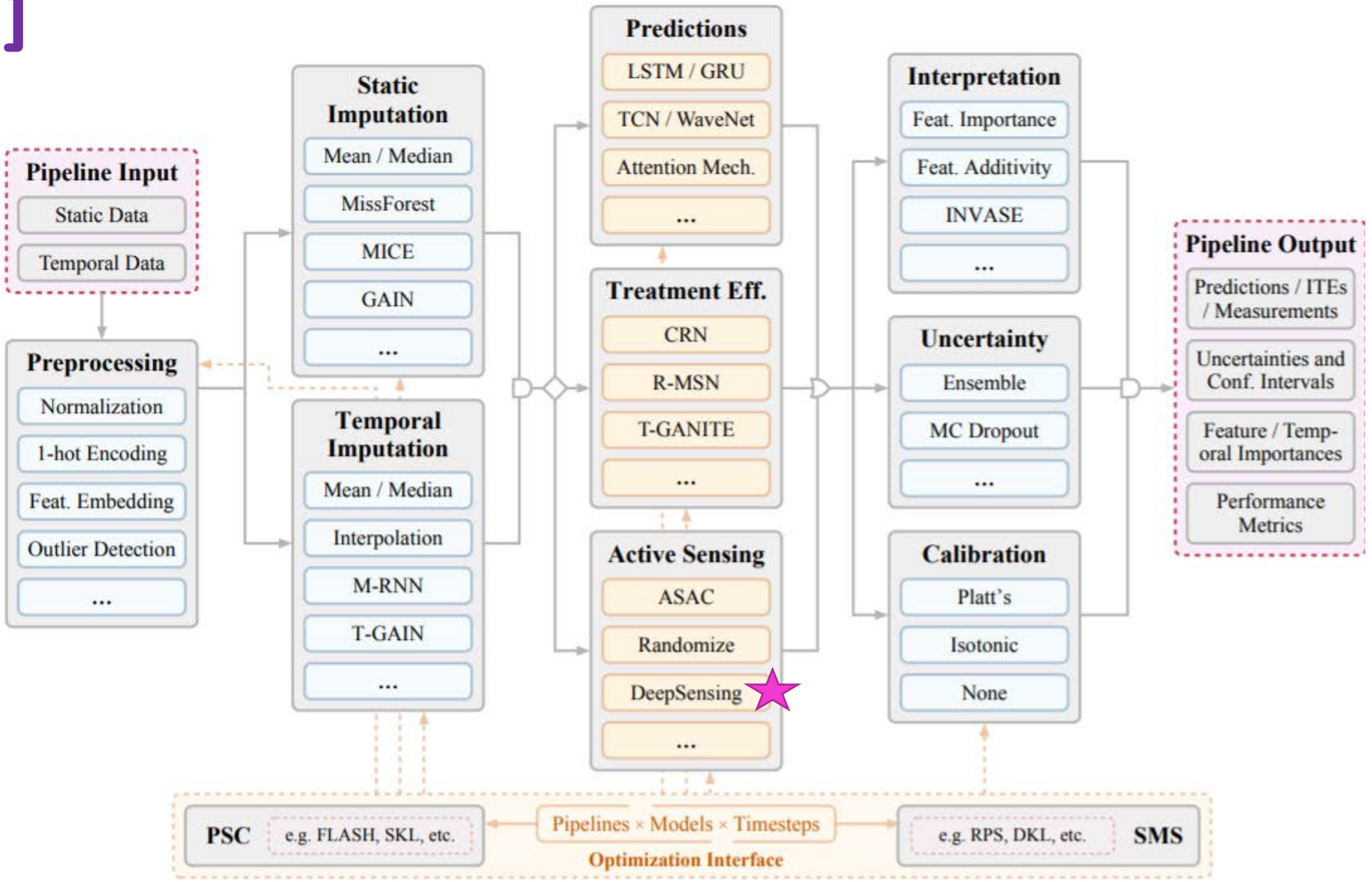
Radiotherapy application

$$\begin{aligned}
\mathbf{Y}_{t+\tau} \quad &| \quad \mathbf{a}_t = A_2, \mathbf{a}_{t+1} = A_0, \ldots \mathbf{a}_{t+\tau-1} = A_0, \bar{\mathbf{H}}_t \\
\mathbf{Y}_{t+\tau} \quad &| \quad \mathbf{a}_t = A_0, \mathbf{a}_{t+1} = A_2, \ldots \mathbf{a}_{t+\tau-1} = A_0, \bar{\mathbf{H}}_t \\
&\ldots \\
\mathbf{Y}_{t+\tau} \quad &| \quad \mathbf{a}_t = A_0, \mathbf{a}_{t+1} = A_0, \ldots \mathbf{a}_{t+\tau-1} = A_2, \bar{\mathbf{H}}_t
\end{aligned}$$

$A_0 = \text{no treatment}$ $A_1 = \text{chemotherapy}$ $A_2 = \text{radiotherapy}$

		$\gamma_c = 5, \gamma_r = 5$			$\gamma_c = 5, \gamma_r = 0$			$\gamma_c = 0, \gamma_r = 5$		
	τ	CRN	RMSN	MSM	CRN	RMSN	MSM	CRN	RMSN	MSM
Treatment Accuracy	3	**83.1%**	75.3%	73.9%	**83.2%**	78.6%	77.1%	**92.9%**	87.3%	74.9%
	4	**82.5%**	74.1%	68.5%	**81.3%**	77.7%	73.9%	**85.7%**	83.8%	74.1%
	5	**73.5%**	72.7%	63.2%	**78.3%**	77.2%	72.3%	**83.8%**	82.1%	72.8%
	6	**69.4%**	66.7%	62.7%	**79.5%**	76.3%	71.8%	**78.6%**	69.7%	64.5%
	7	**71.2%**	68.8%	62.4%	**72.7%**	71.8%	71.6%	**71.9%**	69.3%	61.2%
Treatment Timing Accuracy	3	**79.6%**	78.1%	67.6%	**80.5%**	76.8%	77.5%	**79.8%**	75.7%	60.6%
	4	**73.9%**	70.3%	63.1%	**79.0%**	77.2%	73.4%	**75.4%**	71.4%	58.2%
	5	**69.8%**	68.6%	62.4%	**78.3%**	73.3%	63.6%	**66.9%**	31.3%	29.5%
	6	**66.9%**	66.2%	62.6%	**73.5%**	72.1%	63.9%	**65.8%**	24.2%	15.5%
	7	**64.5%**	63.6%	62.2%	**70.6%**	57.4%	44.2%	**63.9%**	25.6%	12.5%

Clairvoyance:
A unified, end-to-end pipeline for clinical-decision support [ICLR 2021]

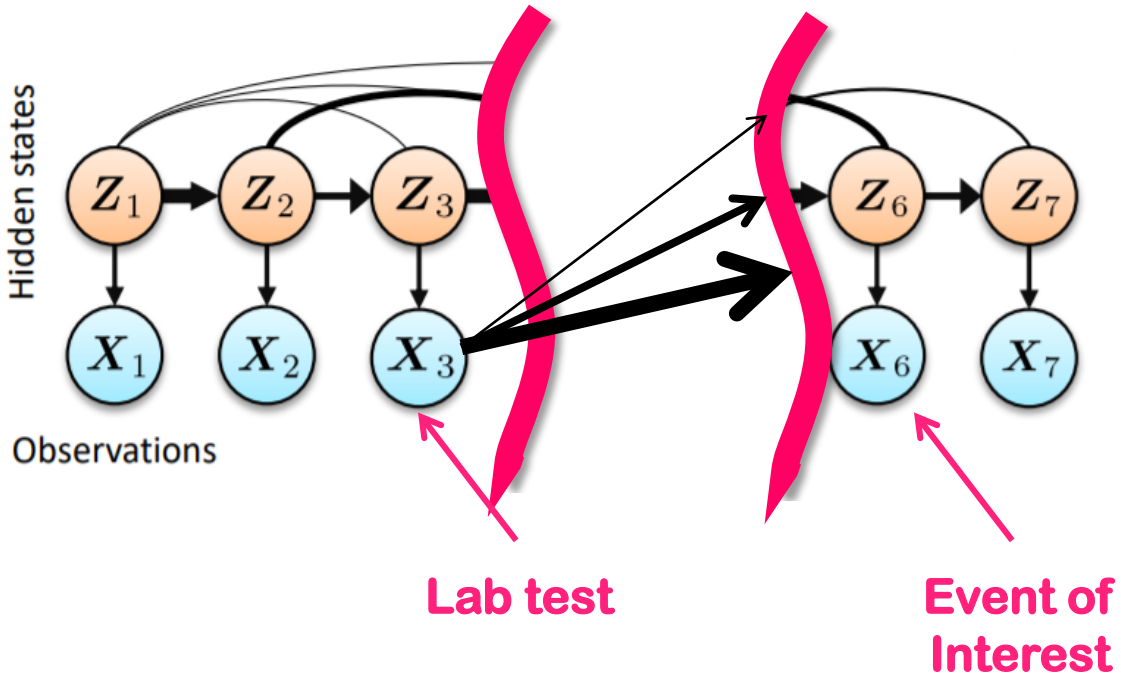


Personalized Screening/Monitoring:
Who to Screen? When to Screen? What to Screen?

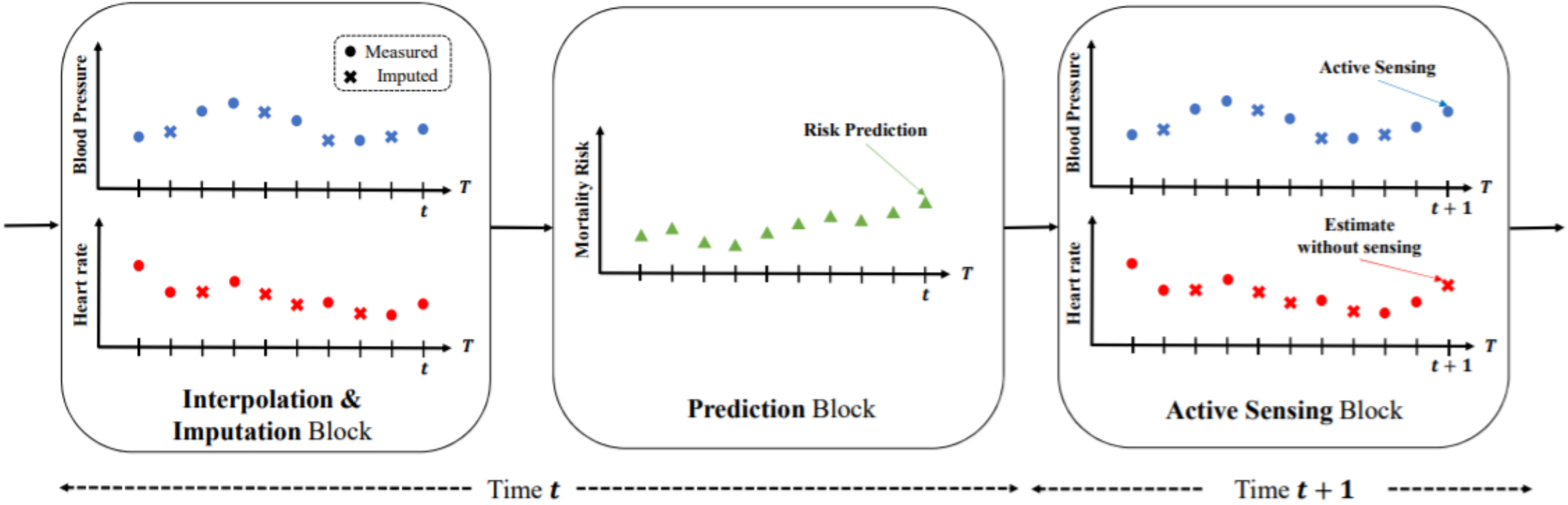
Deep Sensing [Yoon, Zame, vdS, ICLR 2018]

Disease Atlas [Lim, vdS, ML4HC 2018]

ASAC [Yoon, Jordon, vdS, ML4HC 2019]



Deep Sensing: Active Sensing using Multi-directional Recurrent Neural Networks [Yoon, Zame, vdS, ICLR 2018]

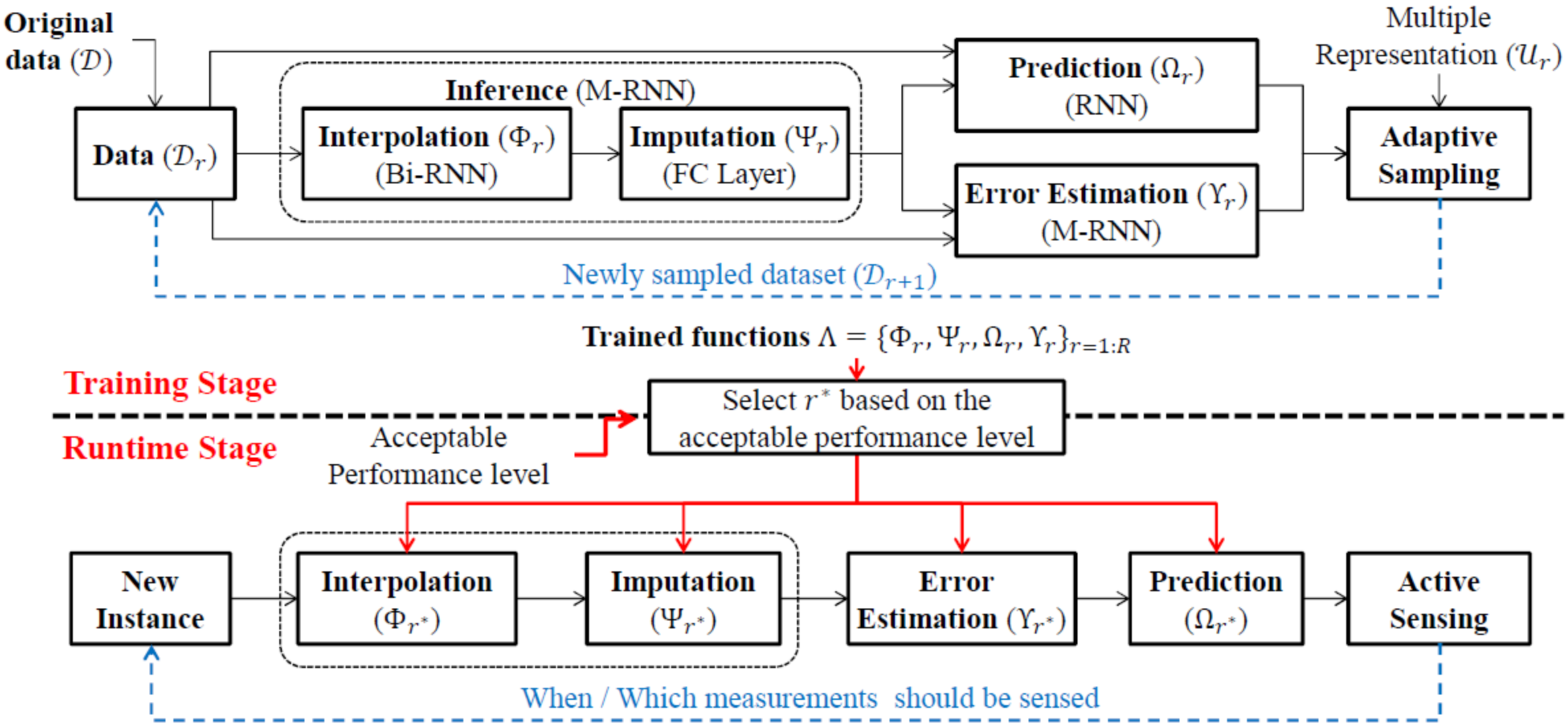


Challenges: Value of information is unknown & dynamically changing – needs to be learned!

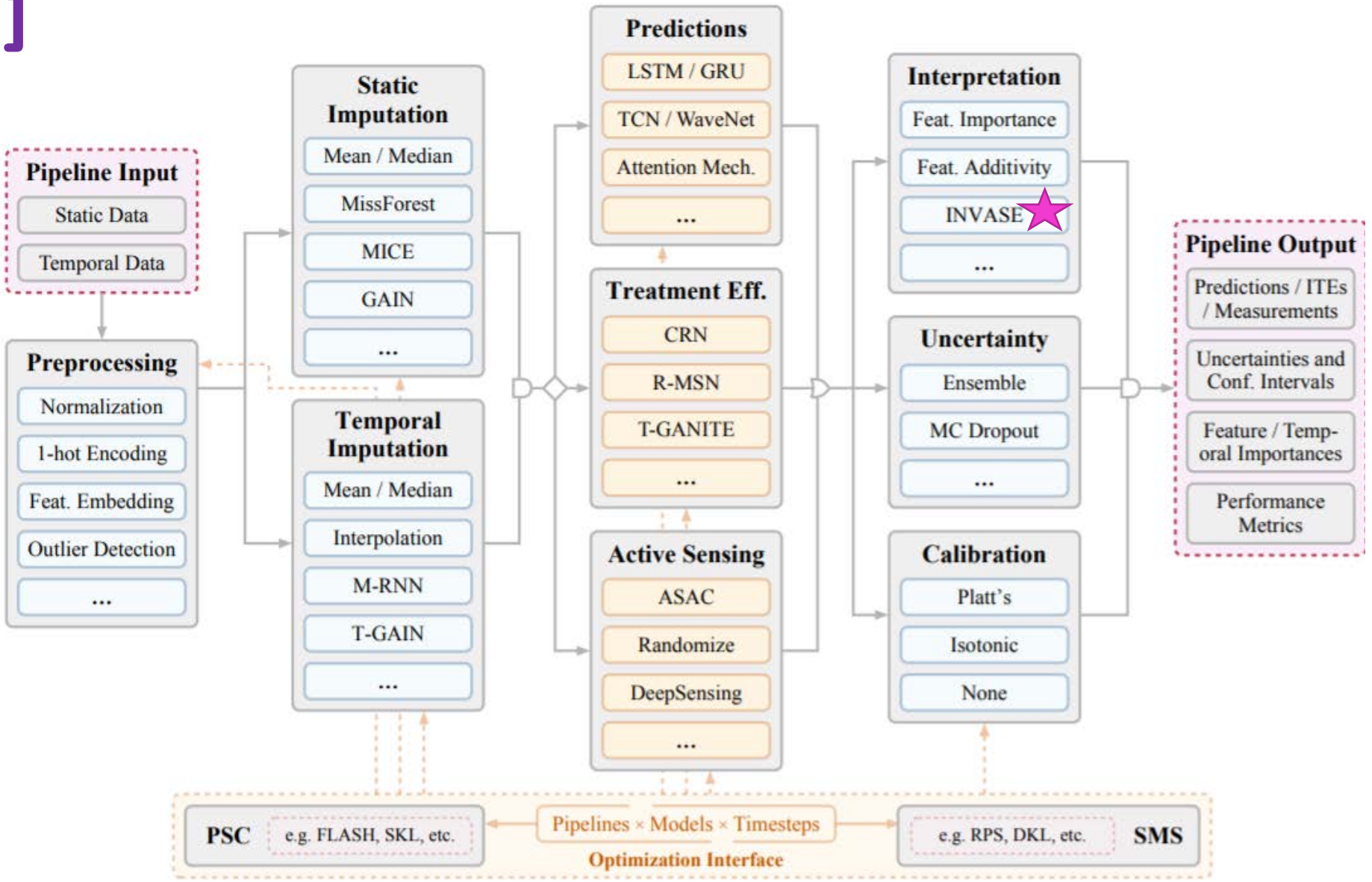
Deep Sensing: Active Sensing using Multi-directional Recurrent Neural Networks [Yoon, Zame, vdS, ICLR 2018]

- Challenges:
  - Value of information is unknown & dynamically changing – needs to be learned!
- Ideas:
  - Learn at training time value of information
  - A neural network learns to predict at various cost-performance points
  - How?
  - It creates multiple representations at various performance levels associated with different measurement rates (costs).
  - Each representation is learned and constructed recursively and adaptively learned by deliberately introducing missing data

Deep Sensing Architecture [Yoon, Zame, vdS, ICLR 2018]



Clairvoyance:
A unified, end-to-end pipeline for clinical-decision support [ICLR 2021]

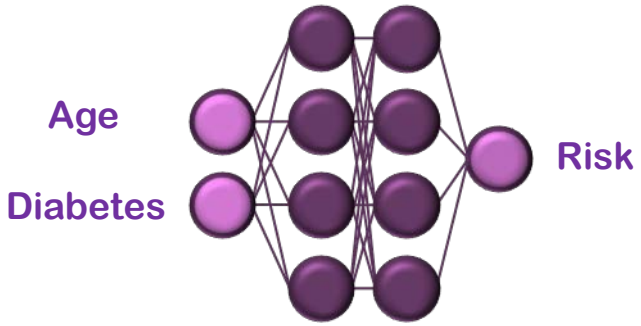


Interpretability, explainability and trustworthiness

Understand

why a prediction is made by the model

Interpretability



Interpretation 1

$$\boldsymbol{Risk} \approx \beta_o\, Age + \beta_1\, Diabetes$$

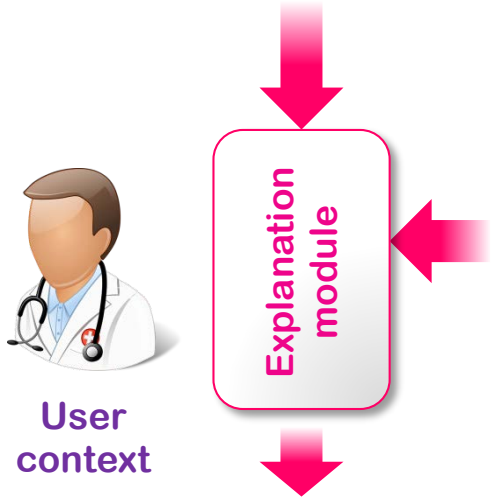
Interpretation 2

Feature importance: β_o, β_1

what can we learn from the model

Explainability

All possible interpretations



Interpretation 2

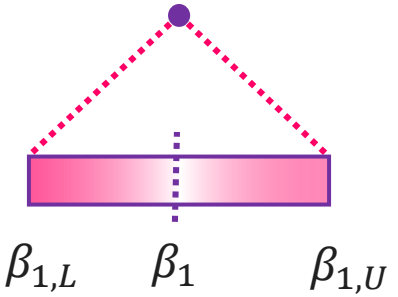
Feature importance: β_o, β_1

how trustworthy is the model's prediction

Trustworthiness

$$\boldsymbol{Risk} \approx \beta_o\, Age + \beta_1\, Diabetes$$

Confidence interval



Which features of an *individual* are relevant for a prediction?

Mary

Age, Gender, Diabetes, Hypertension, SBP, ….

Black Box Predictive Model

Mortality due to Covid-19: 0.78

INVASE

Age, Diabetes, Hypertension, SBP

[Yoon, Jordon, vdS, ICLR 2019]

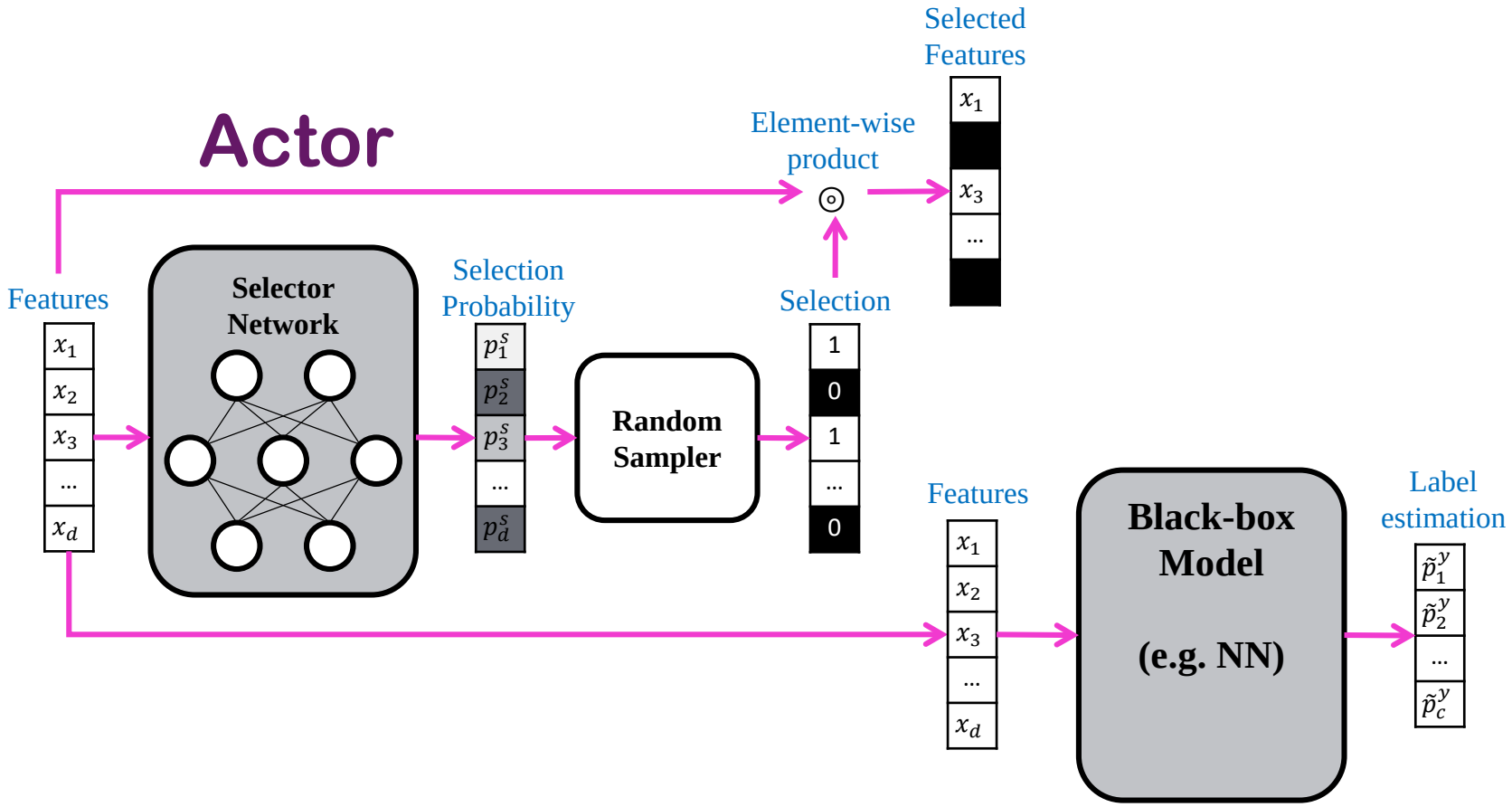
INVASE [Yoon, Jordon, vdS, ICLR 2019]

- How can we learn individualized feature importance?
- Key idea: **Use Reinforcement Learning (RL)**
  - Make observations
  - Select "actions" on the basis of these observations
  - Determine "rewards" for these actions

  - Ultimately learn a policy which selects the best actions
    - i.e. actions that maximize rewards given observations

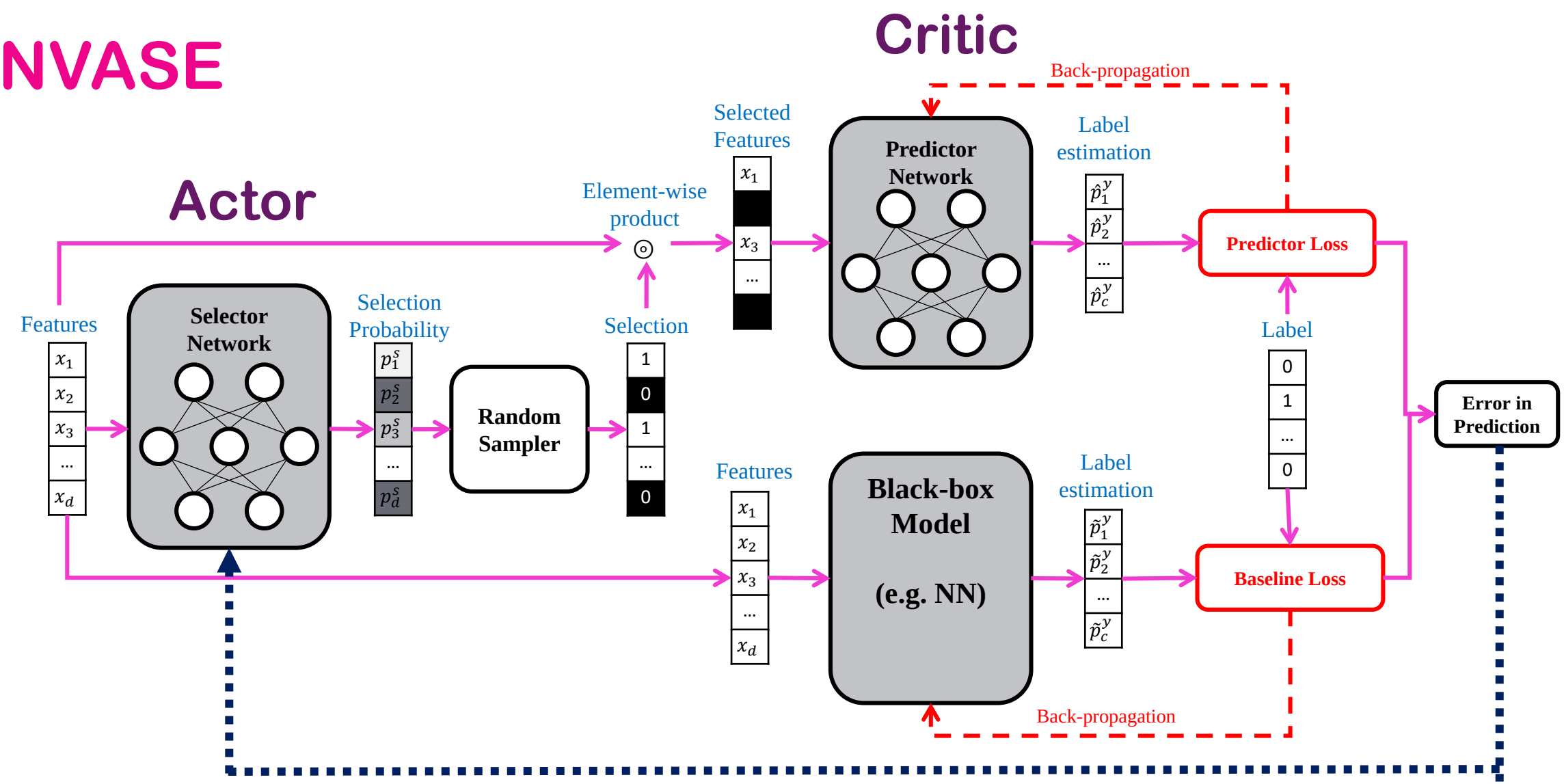
- We use the Actor-Critic approach to RL

INVASE



- **Selector network (actor)** takes instances and outputs vector of selection probabilities.

INVASE



- **Predictor network (critic)** receives the selected features, makes predictions and provides **feedback to the actor**.

Limitations of past methods for model interpretability

Method	Feature importance	Individualized feature importance	Model-independent	Identifying the set of relevant features for each instance
LASSO [Tibshirani, 1996]	✓		✓	
Knock-off [Candes et al, 2016]	✓		✓	
L2X [Chen et al, 2018]	✓	✓	✓	
LIME [Ribeiro et al, 2016]	✓	✓	✓	
SHAPE [Lundberg et al, 2017]	✓	✓	✓	
DeepLIFT [Shrikumar et al, 2017]	✓	✓		
Saliency [Simonyan et al, 2013]	✓	✓		
TreeSHAP [Lundberg et al, 2018]	✓	✓		
Pixel-wise [Batch et al, 2015]	✓	✓		
INVASE [Yoon, Jordon and van der Schaar, 2019]	✓	✓	✓	✓

INVASE discovers *the number* of relevant features for each instance

Are we done? What we are aiming for?

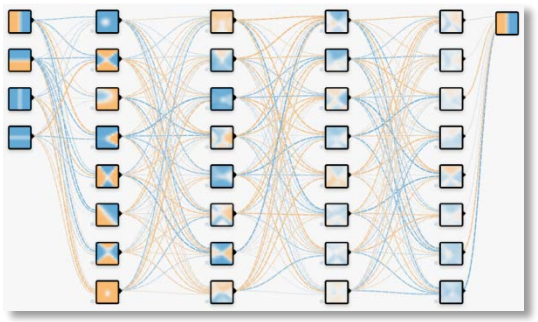
- Understand what the model discovered:
  feature importance, instance-wise feature importance, feature/statistical interactions, model non-linearity, etc.
- Produce a transparent risk equation describing the model for approval in practice guidelines
- Enable model explainability, not only interpretability

Can we have it all??

YES!

Demystifying Black-box Models with Symbolic Metamodels
[A. Alaa & vdS, NeurIPS 2019]

Black-box ML model



$f(\mathbf{x})$

Symbolic Metamodeling

White-box (transparent) model

$$\alpha_1 X_1 + \alpha_2 X_2^2 + \alpha_3 X_1 X_2$$
$$\alpha_4 X_3^3 + \alpha_5 \log(X_4)$$

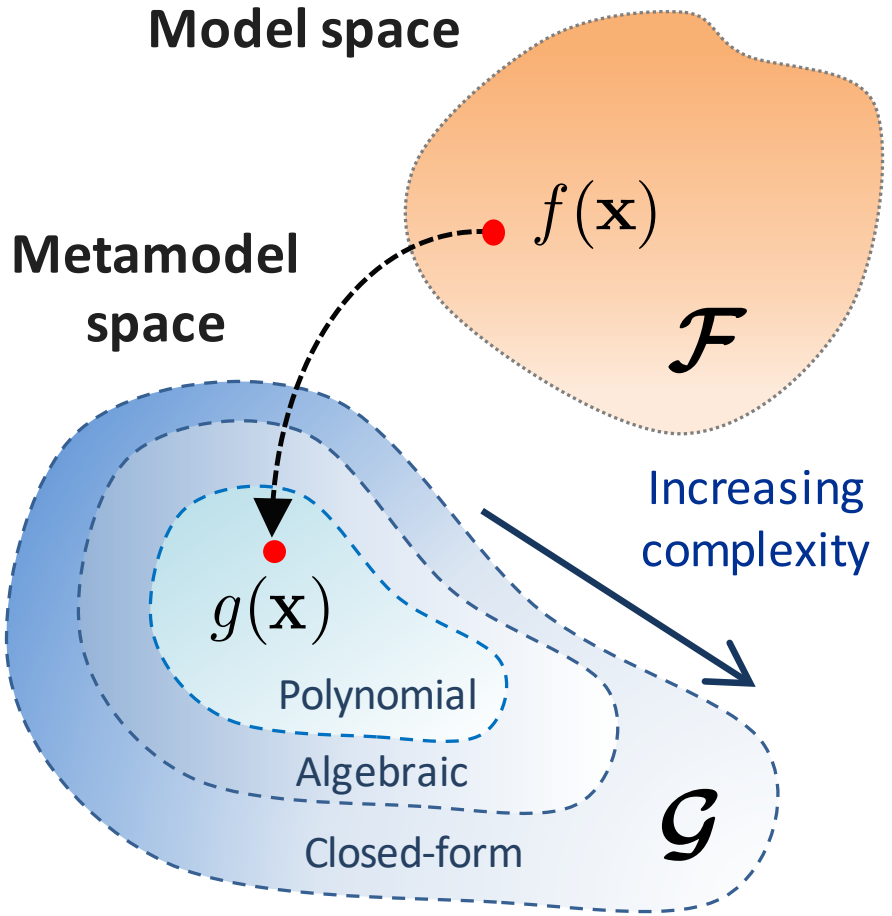
$g(\mathbf{x})$

- **Metamodel =** a model of a model.
- **A symbolic metamodel** outputs a **transparent function** ***describing*** the predictions of the black box model
- **Metamodeling** needs only **query access** to **trained black-box model**.

Symbolic Metamodeling

Black-box ML model

Model space (uninterpretable)



White-box model

Metamodel space

Metamodel space can be *chosen* by the user!

How are we going to achieve this?

- **Kolmogorov-Arnold Theorem** [Kolmogorov et al, 1961]

  Every multivariate continuous function can be written as a finite composition of **univariate** continuous functions

$$g(\mathbf{x}) = \sum_{q=0}^{r} g_q \left(\sum_{p=1}^{n} g_{q,p}(x_p) \right)$$

- **The symbolic metamodeling problem**

Metamodel representation based on interpretable basic fcts

$$g(\mathbf{x};\theta) = \sum_{q=0}^{2n} G \left(\sum_{p=1}^{n} G(x_p;\theta_{q,p});\theta_q \right)$$

Metamodel optimization

$$\theta^* = \arg\min_{\theta\in\Theta} \ell(f(\mathbf{x}), g(\mathbf{x};\theta))$$

What basic functions?

- **Meijer G-functions** [C. S. Meijer, 1936]

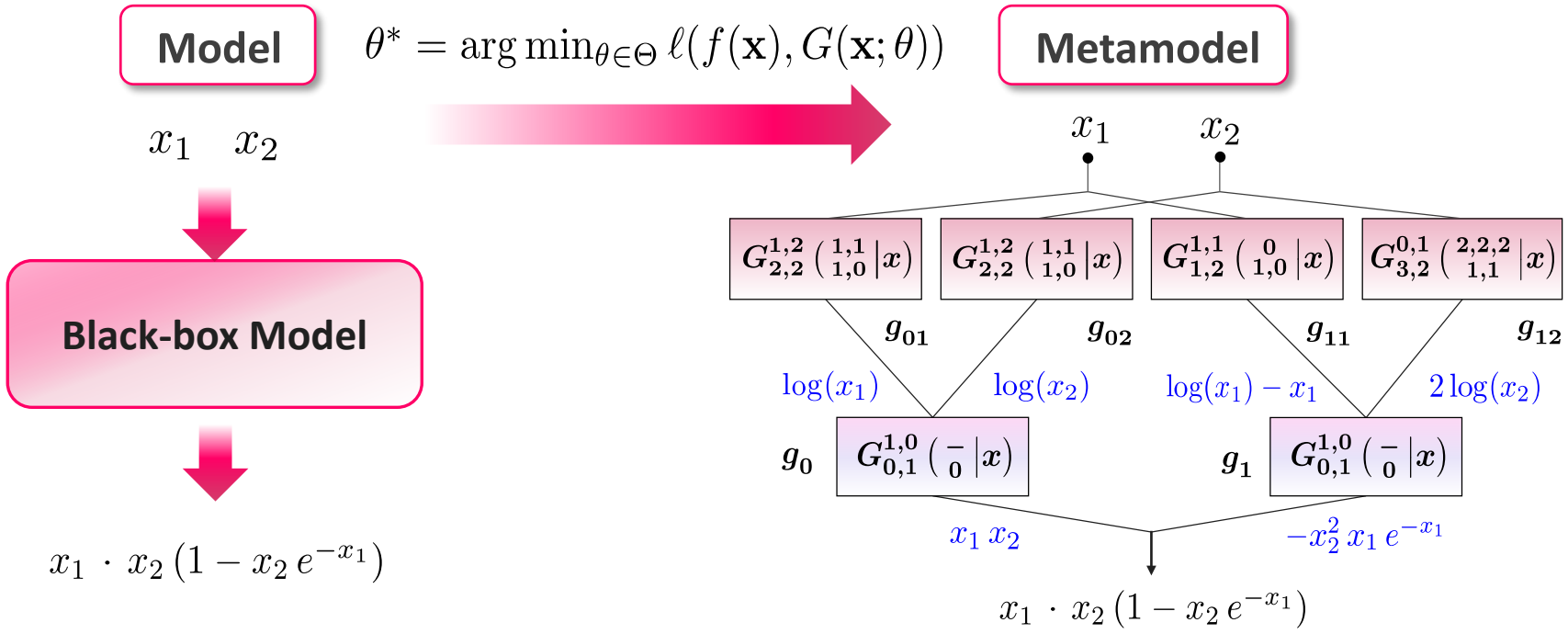
$$G_{p,q}^{m,n}\left(\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_q\end{matrix}\middle|x\right)=\frac{1}{2\pi i}\int_L \frac{\prod_{j=1}^{m}\Gamma(b_j-s)\prod_{j=1}^{n}\Gamma(1-a_j+s)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+s)\prod_{j=n+1}^{p}\Gamma(a_j-s)}\,x^s\,ds$$

- **Very general class of functions**
- **Parameter selection yields many familiar functions**

G-function	Equivalent function	G-function	Equivalent function
$G_{0,1}^{1,0}\left(\begin{matrix}-\\ 0\end{matrix}\middle|-x\right)$	e^x	$G_{2,2}^{1,2}\left(\begin{matrix}\frac{1}{2},1\\ \frac{1}{2},0\end{matrix}\middle|x^2\right)$	$2\arctan(x)$
$G_{2,2}^{1,2}\left(\begin{matrix}1,1\\ 1,0\end{matrix}\middle|x\right)$	$\log(1+x)$	$G_{1,2}^{2,0}\left(\begin{matrix}1\\ \alpha,0\end{matrix}\middle|x\right)$	$\Gamma(\alpha,x)$
$G_{0,2}^{1,0}\left(\begin{matrix}-\\ 0,\frac{1}{2}\end{matrix}\middle|\frac{x^2}{4}\right)$	$\frac{1}{\sqrt{\pi}}\cos(x)$	$G_{1,2}^{2,0}\left(\begin{matrix}1\\ 0,\frac{1}{2}\end{matrix}\middle|x^2\right)$	$\sqrt{\pi}\,\mathrm{erfc}(x)$
$G_{0,2}^{1,0}\left(\begin{matrix}-\\ \frac{1}{2},0\end{matrix}\middle|\frac{x^2}{4}\right)$	$\frac{1}{\sqrt{\pi}}\sin(x)$	$G_{0,2}^{1,0}\left(\begin{matrix}-\\ \frac{a}{2},\frac{-a}{2}\end{matrix}\middle|\frac{x^2}{4}\right)$	$J_a(x)$

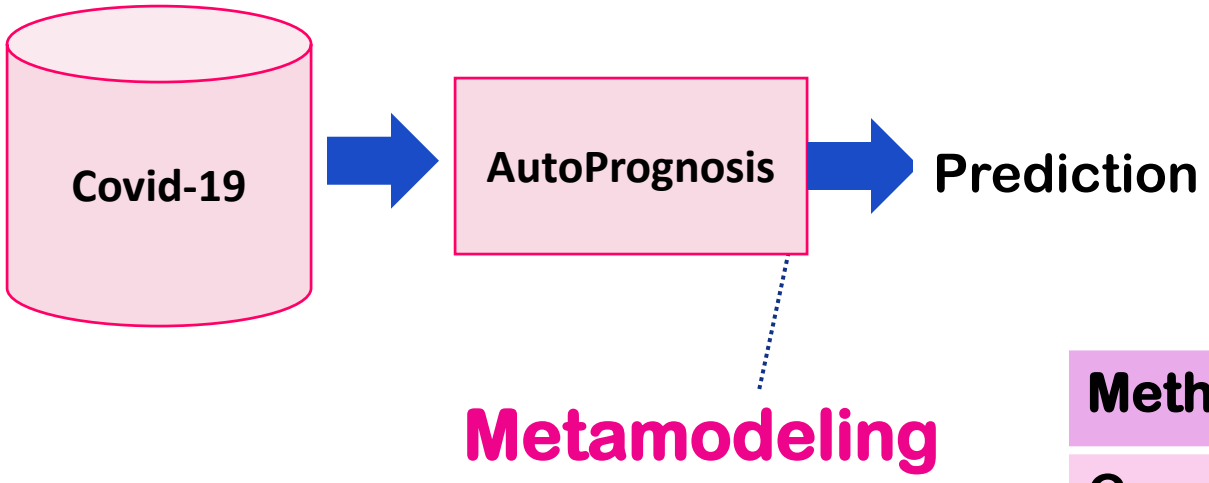
Building a symbolic metamodel

- **Metamodel construction is "analogous" to a 2-layer neural network**



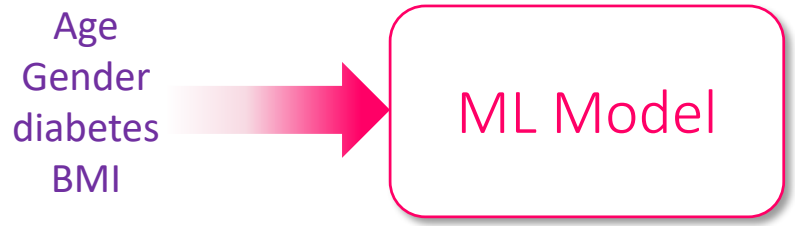
Parameters of a Meijer-G function can be learned by gradient descent!
This can be done very fast!

Interpretability using symbolic metamodeling in practice



Method	AUC-ROC
Cox	0.690 ± 0.002
AutoPrognosis	0.771 ± 0.002

Example: Use Metamodels for Individual-level feature importance



Original Metamodel

$$g(\mathbf{x})$$

$$\alpha_0\,\mathrm{Age} + \alpha_1\,\mathrm{BMI}^2 + \alpha_2 \mathrm{Age}\cdot\mathrm{BMI} + \alpha_3 \mathrm{Age}\cdot\mathrm{Gender}$$
$$\alpha_4 \mathrm{Gender}\cdot(1+\alpha_5\,\mathrm{Diabetes}) + \alpha_6\log(\mathrm{Age}\cdot\mathrm{Diabetes}+1)$$

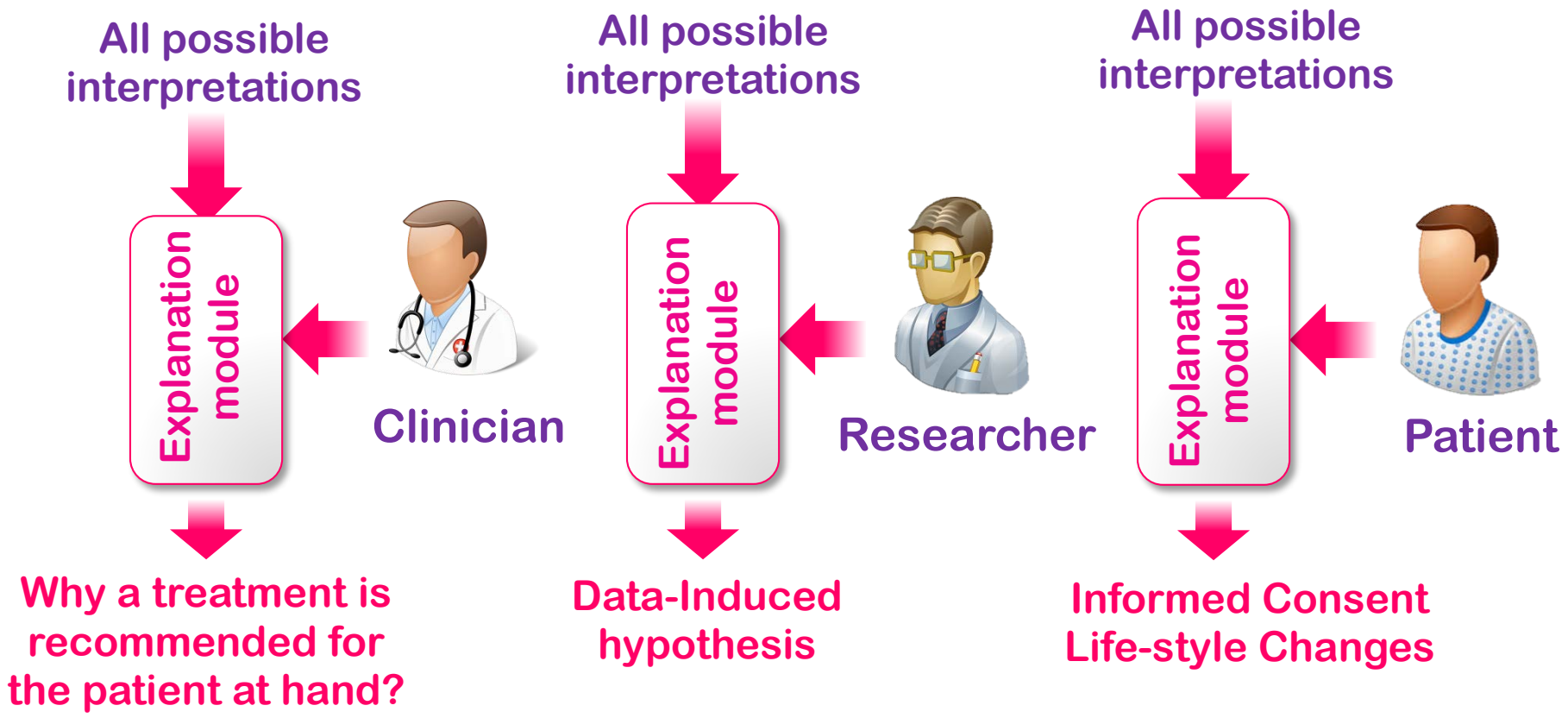
Individual-level feature importance

$$\frac{\partial\, g(\boldsymbol{x})}{\partial\,\mathrm{Age}} = \alpha_0 + \alpha_2\,\mathrm{BMI} + \alpha_3\,\mathrm{Gender} + \frac{\alpha_6\,\mathrm{Diabetes}}{\mathrm{Age}+1}$$

Beyond current feature importance

Explainability =User-dependent Interpretability

- Different users seek different forms of “understanding”…



Metamodels: How to use them?

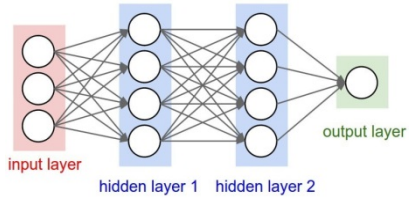
- Different forms of interpretations can be extracted from a Metamodel's forward and backward views!

Forward use

Input= features
Output=risk

- Treatment justification
- Hypothesis induction
- Variable interactions
- Variable importance

Black-box model



Symbolic Metamodel

Metamodel

g(Age, ER, HER2, Tumor size, Nodes)

Backward use

Input= reduced risk
Output=features

- Modifiable variables
- Dosage recommendation
- Policy design

Machine learning: from black boxes to white boxes

Machine learning: from black boxes to white boxes - Mihaela ...

The Alan Turing Institute

YouTube - Apr 14, 2020

https://www.youtube.com/watch?v=EVI5iMpX1cg

https://www.vanderschaar-lab.com/van-der-schaar-lab-at-neurips-2020-8-papers-accepted/

In addition to interpretability & explainability…. trustworthiness is key

Our approach: Post-hoc methodology with frequentist coverage guarantees

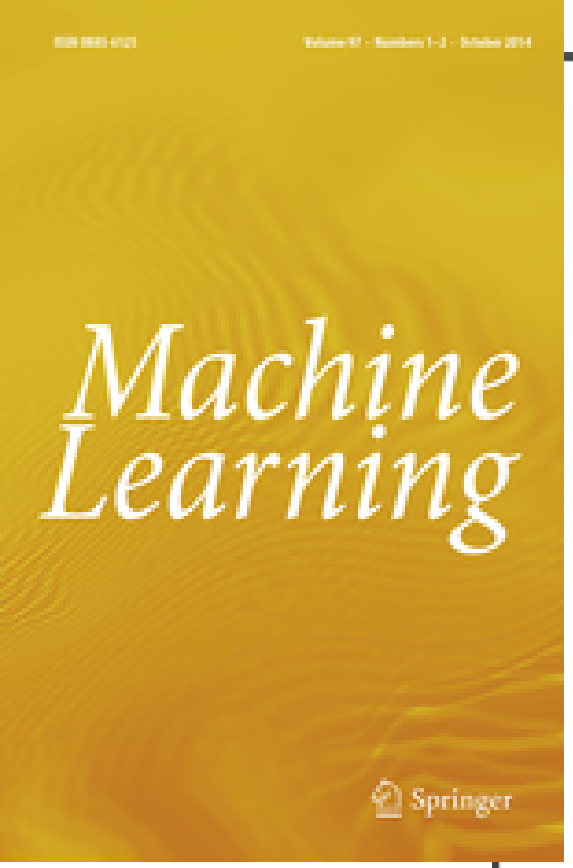
Method	Post-hoc vs Built-in	Coverage
Bayesian neural nets (Ritter et al., 2018)	Built-in	No guarantees
Probabilistic backprop. (Blundell et al., 2015)	Built-in	No guarantees
Monte Carlo dropout (Gal & Ghahramani, 2016)	Built-in	No guarantees
Deep Ensembles (Lakshminarayanan et al., 2017)	Built-in	No guarantees
Discriminative Jackknife (Alaa and vdS,ICML2020)	**Post-hoc**	**1-α**

Alaa and vdS,ICML2020 - Frequentist Uncertainty in Recurrent Neural Networks via *Blockwise* Influence Functions

Responding to COVID-19 with AI and ML (03/2020)

1) Managing limited resources
2) Developing personalized & effective treatment courses for each patient
3) Informing policies and improving collaboration
   - Clinical + Public Health
4) Clinical trials in Covid-19 era
5) Managing uncertainty

https://www.vanderschaar-lab.com/covid-19/



Machine Learning
https://doi.org/10.1007/s10994-020-05928-x

Check for updates

How artificial intelligence and machine learning can help healthcare systems respond to COVID-19

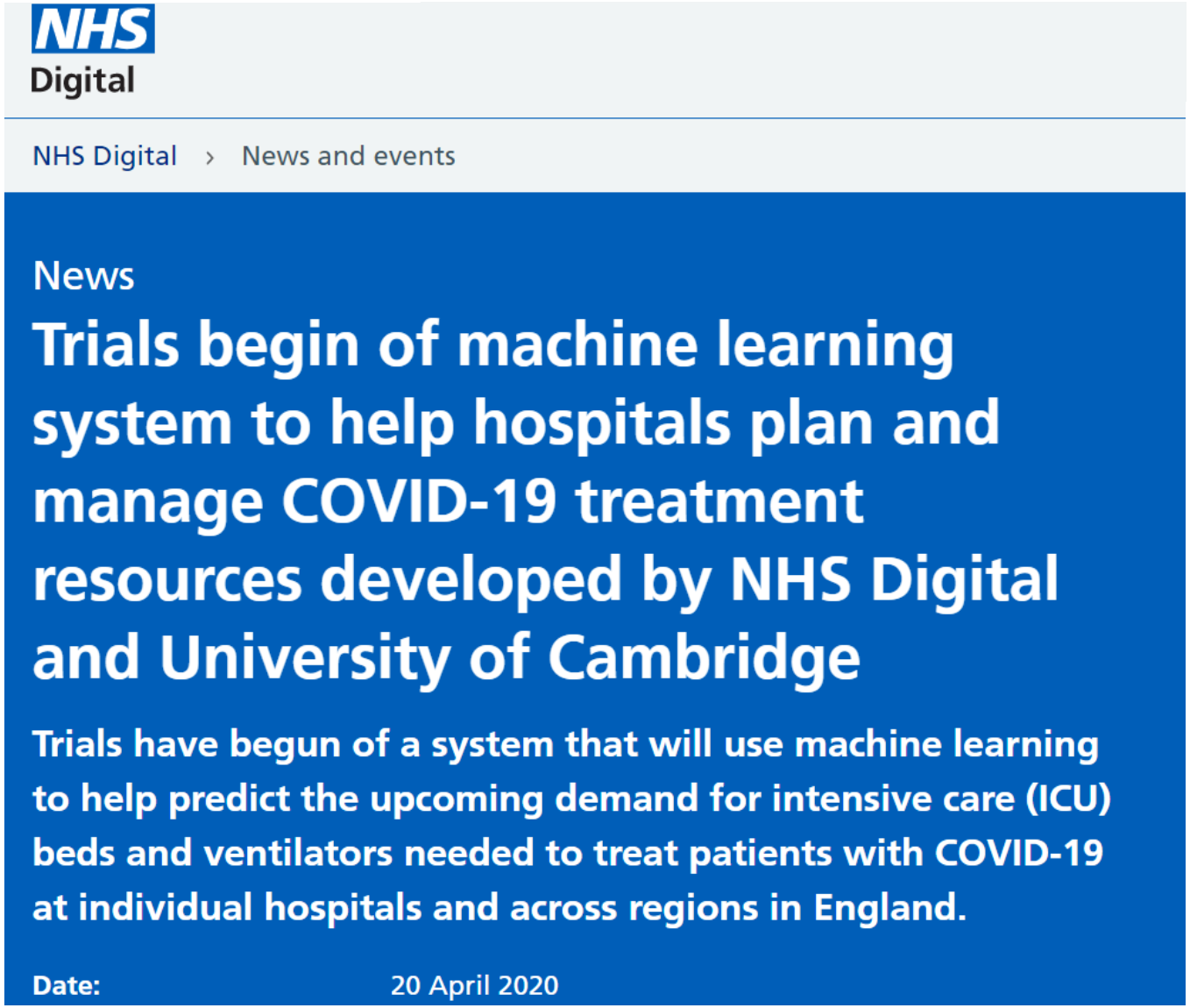
Mihaela van der Schaar[1,2] · Ahmed M. Alaa[2] · Andres Floto[1] · Alexander Gimson[3] · Stefan Scholtes[1] · Angela Wood[1] · Eoin McKinney[1] · Daniel Jarrett[1] · Pietro Lio[1] · Ari Ercole[1,3]

Received: 19 July 2020 / Revised: 18 October 2020 / Accepted: 21 October 2020

Abstract

The COVID-19 global pandemic is a threat not only to the health of millions of individuals, but also to the stability of infrastructure and economies around the world. The disease will inevitably place an overwhelming burden on healthcare systems that cannot be effectively dealt with by existing facilities or responses based on conventional approaches. We believe that a rigorous clinical and societal response can only be mounted by using intelligence derived from a variety of data sources to better utilize scarce healthcare resources

Covid-19 at the National Level - Resource Planning



NHS Digital

NHS Digital › News and events

News

Trials begin of machine learning system to help hospitals plan and manage COVID-19 treatment resources developed by NHS Digital and University of Cambridge

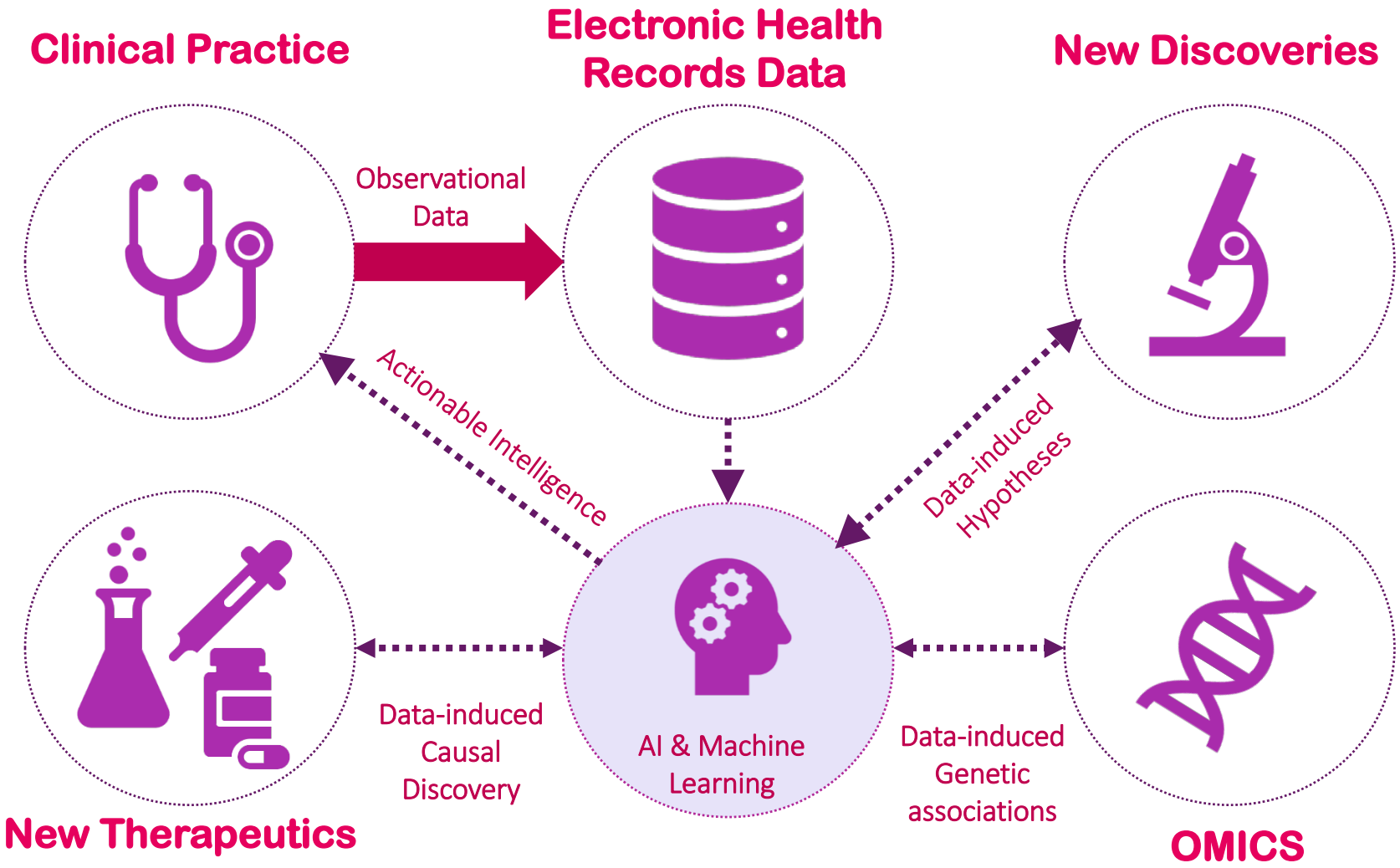
Trials have begun of a system that will use machine learning to help predict the upcoming demand for intensive care (ICU) beds and ventilators needed to treat patients with COVID-19 at individual hospitals and across regions in England.

Date: 20 April 2020

Demonstrator

Z. Qian, A. Alaa, M. van der Schaar, "CPAS: the UK's National Machine Learning-based Hospital Capacity Planning System for COVID-19," *Machine Learning*, 2020.

Catalyzing a Revolution in Healthcare using AI & ML



Cambridge Centre for AI in Medicine's **Inaugural Event**

FREE WEBINAR **22 JANUARY 2021**
2pm - 6pm GMT (9am - 1pm EST)

Speakers include...

Prof. Mihaela van der Schaar
Director
Cambridge Centre for AI in Medicine

Prof. Andres Floto
Co-Director
Cambridge Centre for AI in Medicine

Dr Sarah Teichmann
Head of Cellular Genetics
Wellcome Sanger Institute

Dr Jim Weatherall
Vice President, Data Science & AI, R&D
AstraZeneca

Dr Tony Wood
Senior VP, Medicinal Science & Technology
GSK

New Horizons in Machine Learning for Medicine

Join us at the Cambridge Centre for AI in Medicine as we reveal the transformations we are driving at the critical interface of machine learning, medicine and bioscience.

Our speakers are drawn from the frontiers of machine learning, science, clinical research, pharmaceutical R&D and the NHS.

PLUS: the brightest stars among the CCAIM faculty's PhD students offer bite-sized presentations on their latest high-impact research.

- **10+ high-profile speakers**
- **Two audience Q&A sessions**
- **PhD student showcases**
- **22 January, from 2pm**

To **REGISTER** your interest and receive an agenda & Zoom link ahead of the event, visit Eventbrite at **bit.ly/CCAIMEVENT**

UNIVERSITY OF CAMBRIDGE

Cambridge Centre for AI in Medicine

We are hiring brilliant students and post-docs

Inspiration Exchange

Themed discussion sessions for machine learning students (particularly masters, Ph.D., and post-docs).

Subscribe!

vanderschaar-lab.com

@MihaelaVDS

mv472@cam.ac.uk

linkedin.com/in/mihaela-van-der-schaar/