



# The Rise of Simple Models - Explainable AI (xAI) in Medicine

Samuel Demharter, PhD  
Head of Bioinformatics, Target & Biomarker Discovery



## THE TEAM




























# About Abzu

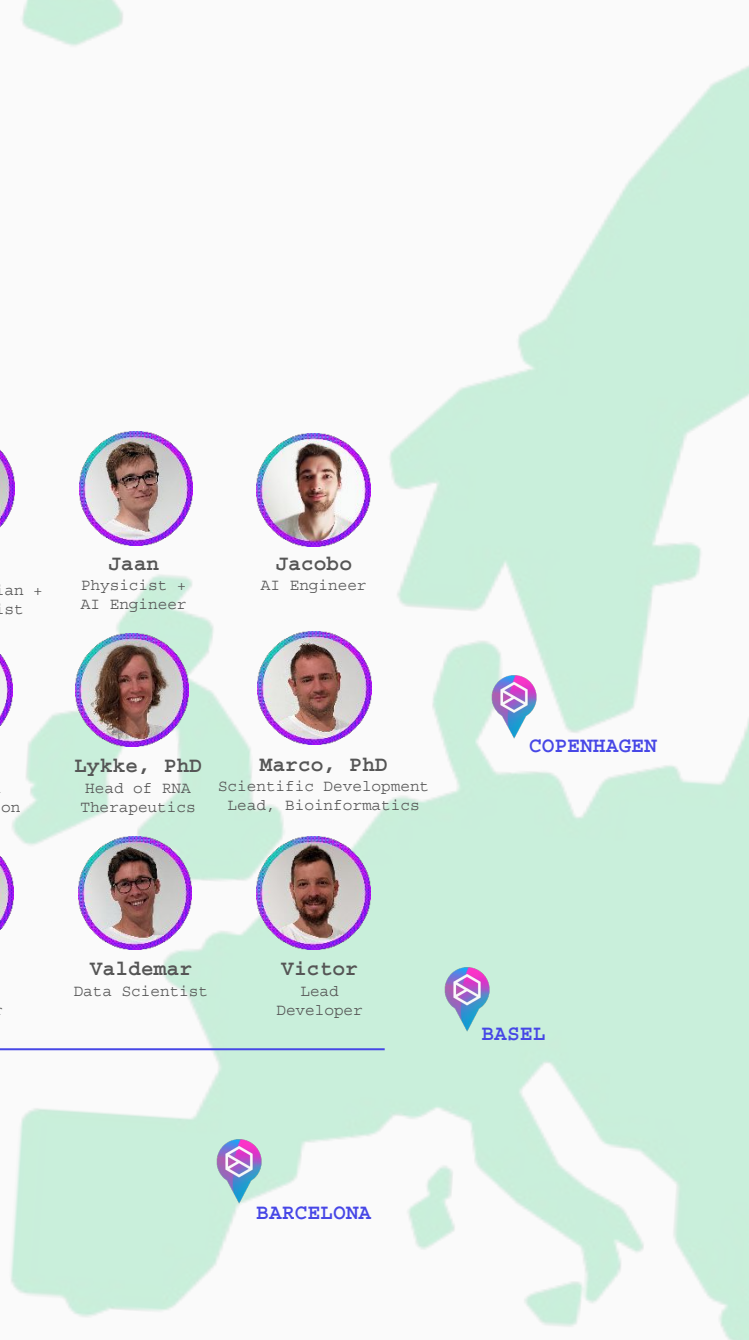
**European**  
Leading deep tech innovation  
in AI & Life Science

**3 Locations**  
In Barcelona, Copenhagen,  
and Basel

**30 Abzoids**  
Highly-skilled experts in a  
self managing organization

**9 nationalities**  
Brazil, Britain, Denmark,  
Estonia, Germany, Italy,  
Scotland, Spain, + US

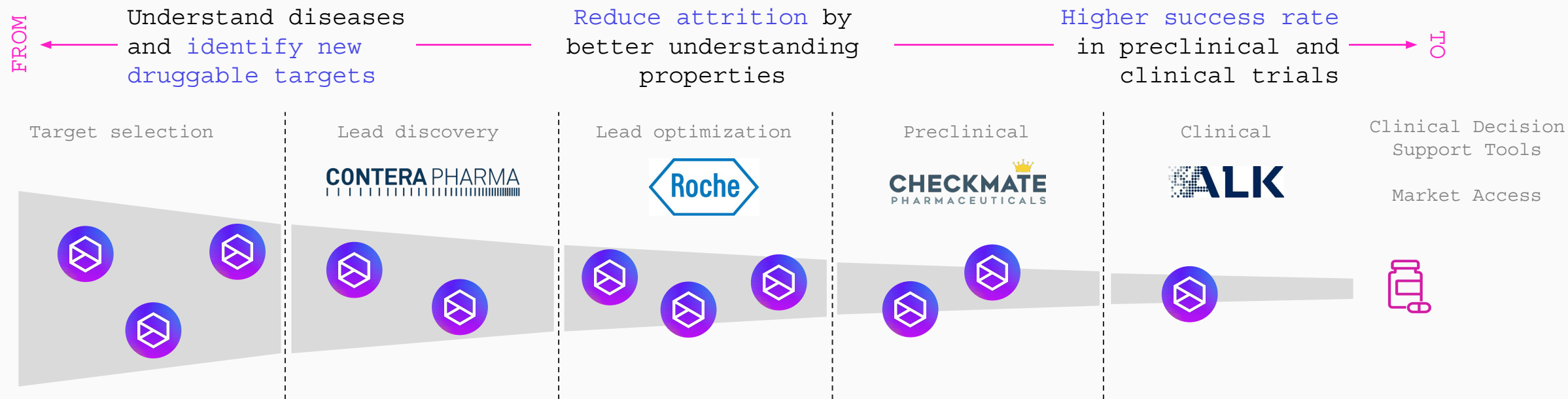
 <b>Caroline</b> Bioinformatician + Data Scientist	 <b>Casper</b> CEO + AI Thought Leader	 <b>Chris, PhD</b> Mathematician + Data Scientist	 <b>Elizabeth</b> Solution + Projects	 <b>Elyse</b> Marketing + Communication	 <b>Emil</b> Developer + Data Scientist	 <b>Helena</b> Bioinformatician + Data Scientist	 <b>Jaan</b> Physicist + AI Engineer	 <b>Jacobo</b> AI Engineer
 <b>James</b> CRO + Sales	 <b>Jonas E.</b> Physicist + Data Scientist	 <b>Jonas N.</b> Commercial + Strategy	 <b>Jonas W.</b> Finance + Operation	 <b>Karin</b> Product Lead	 <b>Kevin</b> Mad Scientist	 <b>Luz</b> UX + Data Visualization	 <b>Lykke, PhD</b> Head of RNA Therapeutics	 <b>Marco, PhD</b> Scientific Development Lead, Bioinformatics
 <b>Maria</b> Office Lead	 <b>Martin</b> Business + Customer	 <b>Meera, PhD</b> Physicist + Data Scientist	 <b>Miquel, PhD</b> Machine Learning Researcher	 <b>Roger</b> Physicist + Developer	 <b>Sam, PhD</b> Head of Bioinformatics	 <b>Tom</b> Lead Developer	 <b>Valdemar</b> Data Scientist	 <b>Victor</b> Lead Developer





OUR TRACK RECORD: DON'T BELIEVE ME, JUST WATCH

We have proven value in multiple cases across Pharma

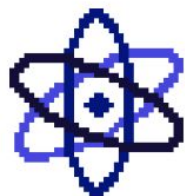


Working with Abzu and using the Qlattice to analyze our *clinical data* has provided us with new insights and helped us generate new hypotheses for exploring the potential for a *biomarker-based enrichment strategy across cancer*.

Art Krieg  
CSO  
**CHECKMATE**  
PHARMACEUTICALS



## xAI-Based Patient Enrichment Strategy



### Data access + data processing

We **enhance your research** with biobank access and prepare your raw data for analysis.



### High-performing + interpretable results

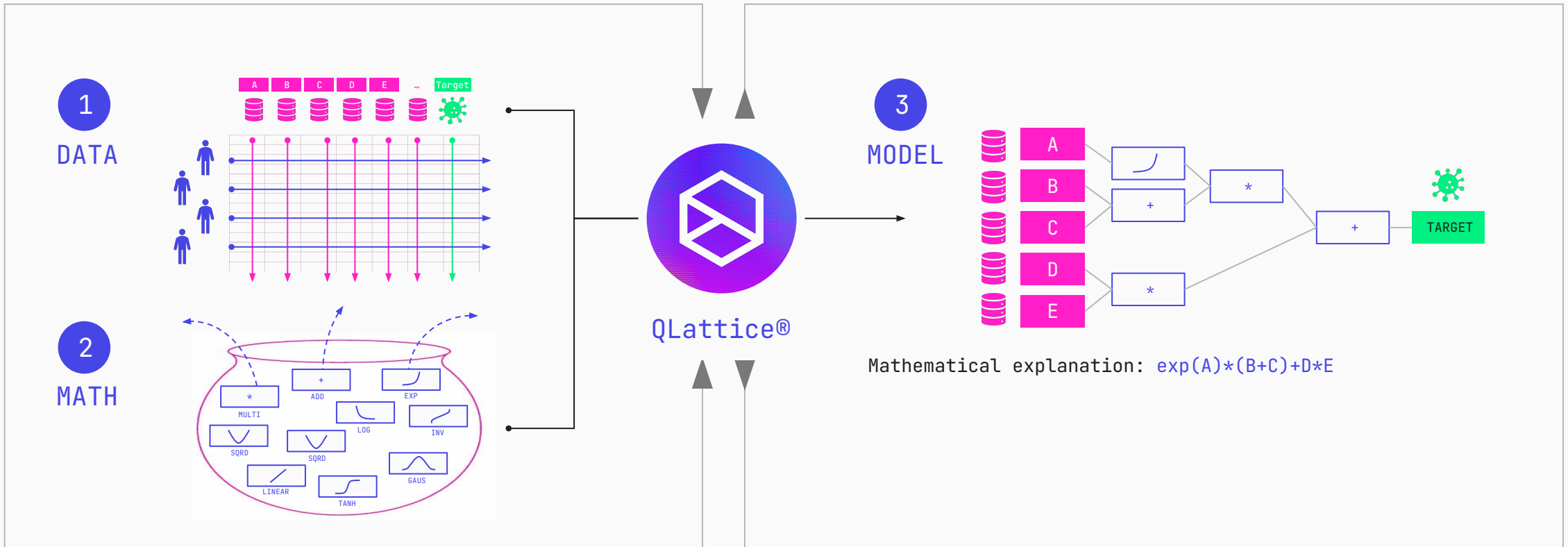
Review a set of composite biomarker signatures with **clear and explainable model interpretations**.



### Regulatory compliance + approval

Confidently **nominate your signature** with our comprehensive report and support.

# We call it the QLattice®



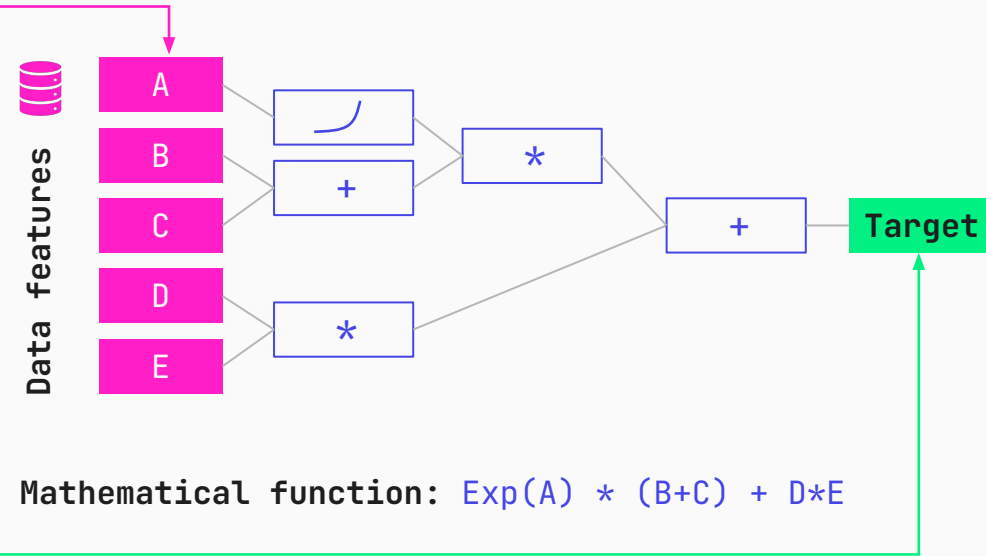
# Enabling smart biomarkers with explainable AI

What we do:

Smart biomarker built with xAI:



Abzu uses Explainable AI (xAI) to find the optimal number of **biomarkers** required to find the highest **quality signature** for predicting a **disease outcome**





CASE: BREAST CANCER

# Locating smart biomarkers targeting breast cancer



7 8 September 2022

© Abzu AI 2020. All rights reserved. Proprietary and confidential.

# Locating smart biomarkers targeting breast cancer



Collect samples

Data quantification

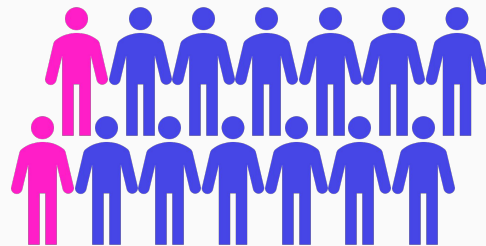
Data preparation

Explainable AI

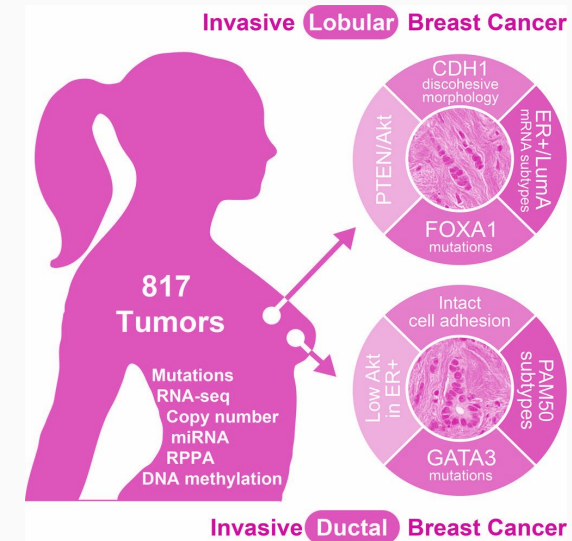
Smart biomarkers



705 Samples, breast tumor



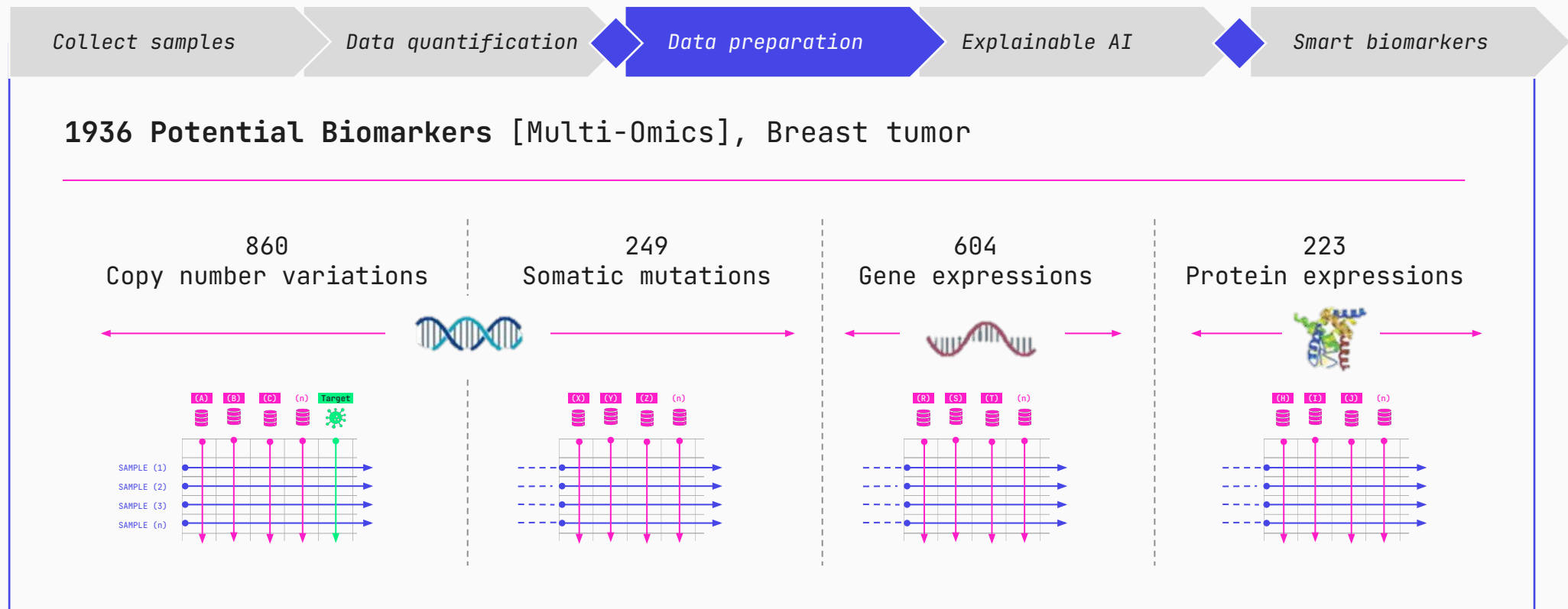
♥ 611 survived  
+ 94 died



Ciriello, Giovanni, et al. "Comprehensive molecular portraits of invasive lobular breast cancer." *Cell* 163.2 (2015): 506-519.



# Locating smart biomarkers targeting breast cancer



# Locating smart biomarkers targeting breast cancer



Collect samples

Data quantification

Data preparation

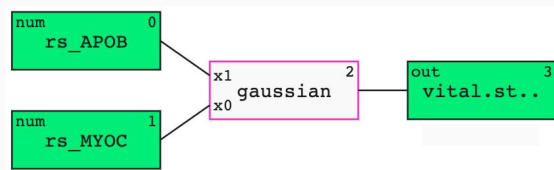
Explainable AI

Smart biomarker



## A simple explanation: Two biomarkers only

Biomarkers:  
[DATA FEATURES]



Quality:  
[PREDICTION ACC. & BENCHMARK]

QLattice®	0.65
Random Forest	0.68

- A 2-feature **non-linear model** that stratifies a subpopulation of cancer patients was found
- Two gene-expression features were found: **APOB** and **MYOC**

# Locating smart biomarkers targeting breast cancer



Collect samples

Data quantification

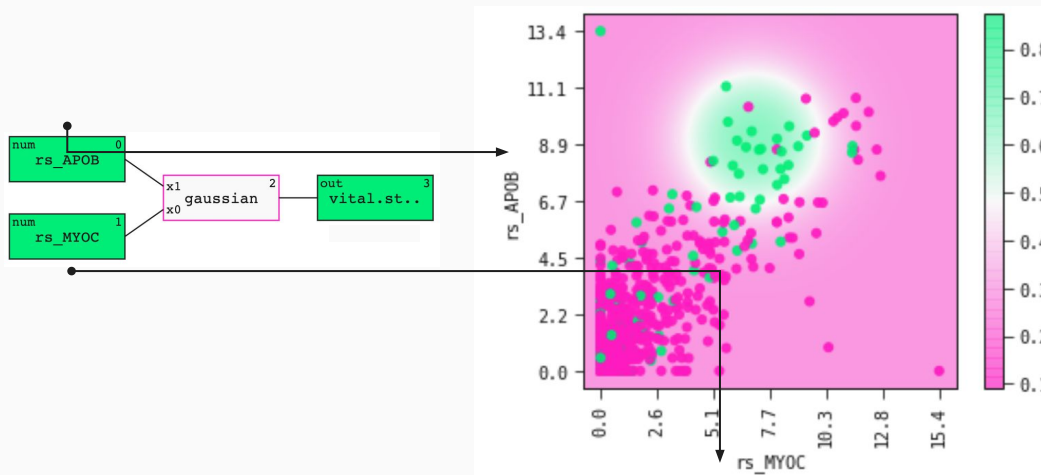
Data preparation

Explainable AI

Smart biomarker



A simple explanation: Two biomarkers only



- A “hotspot” is revealed at:  
**APOB** expression levels of ~9  
**MYOC** expression levels of ~7  
 individuals are at increased risk of death
- This subpopulation was significantly **enriched in ductal cancers**

# Locating smart biomarkers targeting breast cancer



Collect samples

Data quantification

Data preparation

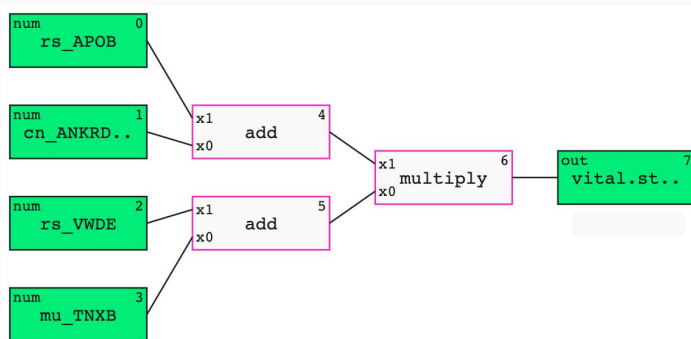
Explainable AI

Smart biomarker



## Adding complexity [still simple!]: exploring multi-omics

Biomarkers:  
[DATA FEATURES]



Quality:  
[PREDICTION ACC. & BENCHMARK]



- A 4-feature *multi-omic* model, now outperforming Random Forrest
- Again, high-lighting *APOB*, while also including a *mutation* and *copy number* feature

## Locating smart biomarkers targeting breast cancer



Collect samples

Data quantification

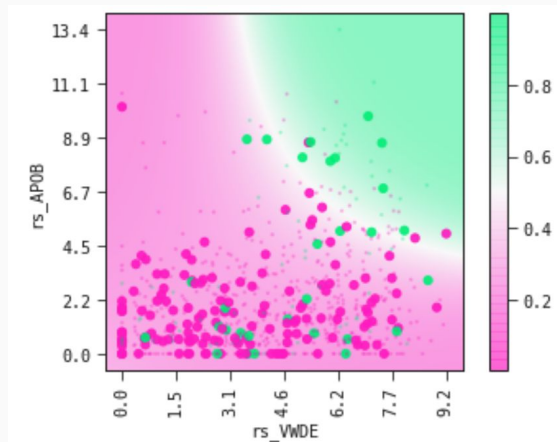
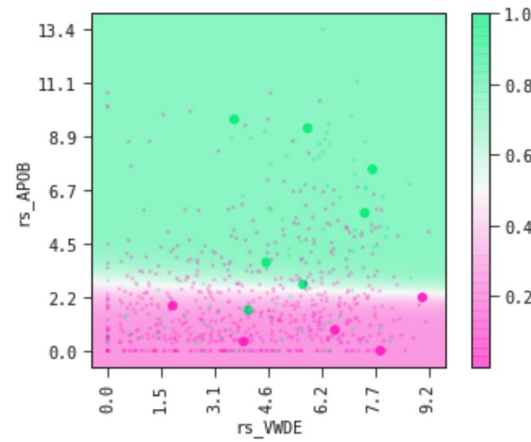
Data preparation

Explainable AI

Smart biomarker



## Adding complexity [still simple!]: exploring multi-omics

Patients without  
the TNXB mutationPatients with the  
TNXB mutation

A “genetic switch” is identified:

- The left side shows the decision boundary for **non-TNXB mutation** carriers: individuals with high **APOB** and **VWDE** levels are less likely to survive
- The right side shows carriers of the **TNXB-mutation** carriers: Here, lower levels of **APOB** lead to higher risk of death, regardless of the levels of **VWDE**

MULTIPLE CASES: SAME RESULT

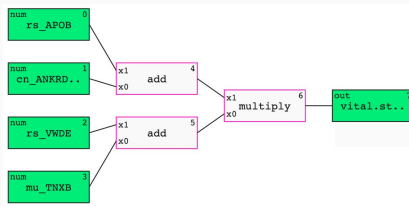
# Uncovering smart biomarkers with explainable AI

## Breast cancer

**Target:** Identify mechanism driving breast cancer severity

**Data:** Multi-Omics  
705 subjects, 1936 features  
(mutations, copy numbers, gene expression, protein levels)

**Smart Biomarkers:**



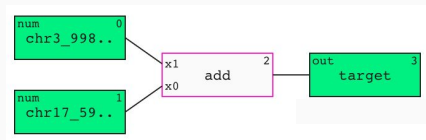
<b>Quality:</b> [PRED. ACC. & BENCHMARK]	<b>QLattice®</b>	0.75
	<b>Random Forest</b>	0.68

Ciriello et al. 2015,  
[https://www.cell.com/cell/fulltext/S0092-8674\(15\)01195-2](https://www.cell.com/cell/fulltext/S0092-8674(15)01195-2)

## Liver cancer

Classification of hepatocellular carcinoma patients using methylation biomarkers

Methylated cfDNA  
91 subjects, 1712 cfDNAs  
(30,000 before variance filter was applied)



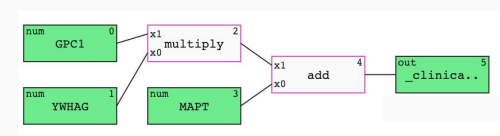
<b>QLattice®</b>	0.95
<b>Random Forest</b>	0.91

Wen, Li, Guo, Liu et al. 2015,  
<https://pubmed.ncbi.nlm.nih.gov/26516143/>

## Alzheimer's disease

Identify potential biomarkers for Alzheimer's disease

Proteomics (mass spec.)  
137 subjects, 1166 proteins

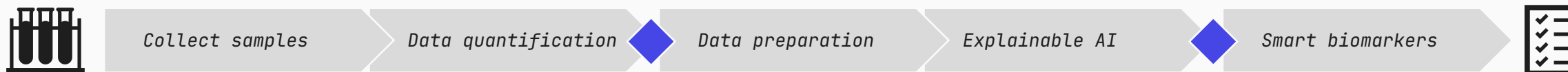


<b>QLattice®</b>	0.96
<b>Random Forest</b>	0.89

Bader & Geyer et al. 2021,  
<https://www.embopress.org/doi/full/10.15252/sb.20199356>

THE ANALYSIS CHALLENGE

# The challenge in data constraints vs. algorithms for analysis

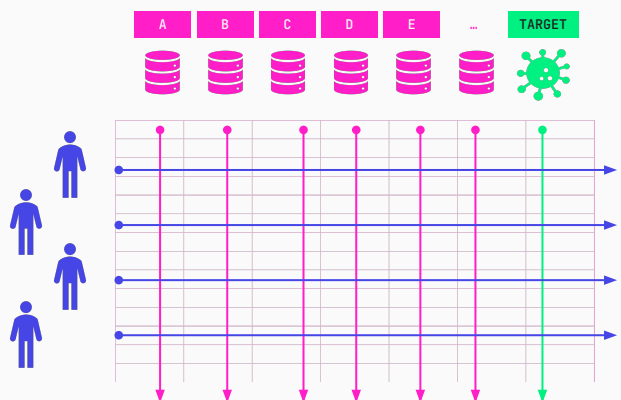


Data for high-throughput assays

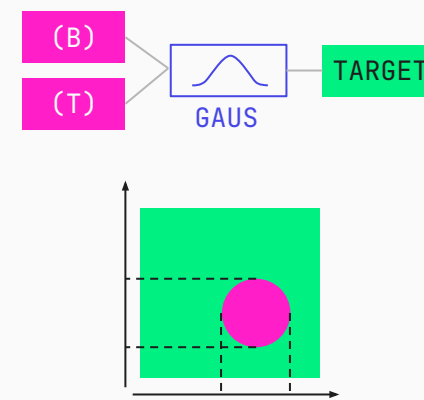


Explainable AI modelling

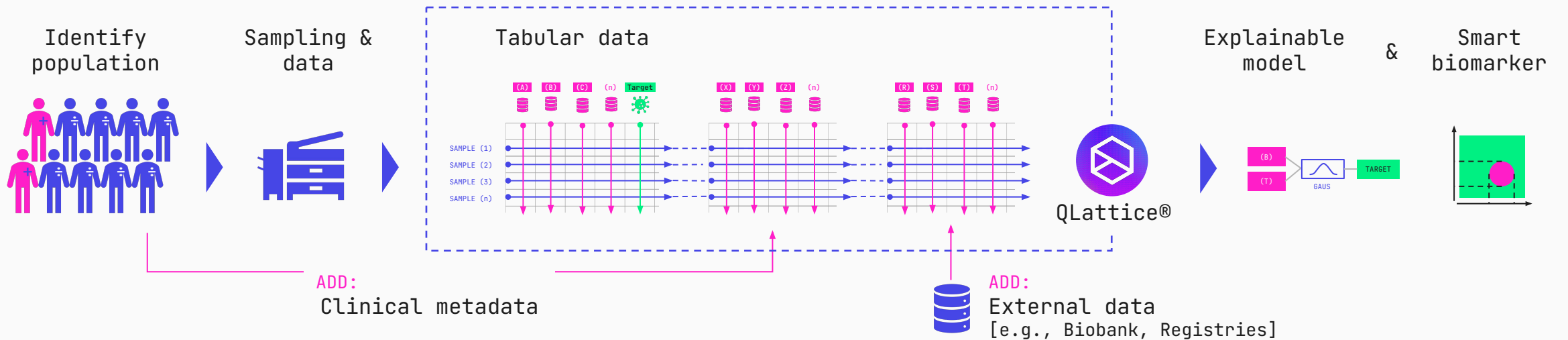
### Tabular data



### Smart biomarkers



# Locating smart biomarkers using xAI - the QLattice®







IF YOU WANT TO KNOW MORE

Contact us and discuss opportunities

Gartner "Cool Vendor"



“ Abzu recognized as a Cool Vendor in artificial intelligence ”

Gartner, Jun. 2022



Academic Network and Biobank Access

Combining symbolic regression with the Cox proportional hazards model improves prediction of heart failure deaths [READ MORE >](#)  
 Wilstrup, C., Cave, C. *BMC Medical Informatics Decision Making* 22, 196 (2022).

Symbolic regression analysis of interactions between first trimester maternal serum adipokines in pregnancies which develop pre-eclampsia [READ MORE >](#)  
 Casper Wilstrup, Paula L. Hedley, Line Rode, Sophie Placing, Karen R. Wejdenann, Anne-Cathrine Shalmi, Karin Sundberg, Michael Christiansen

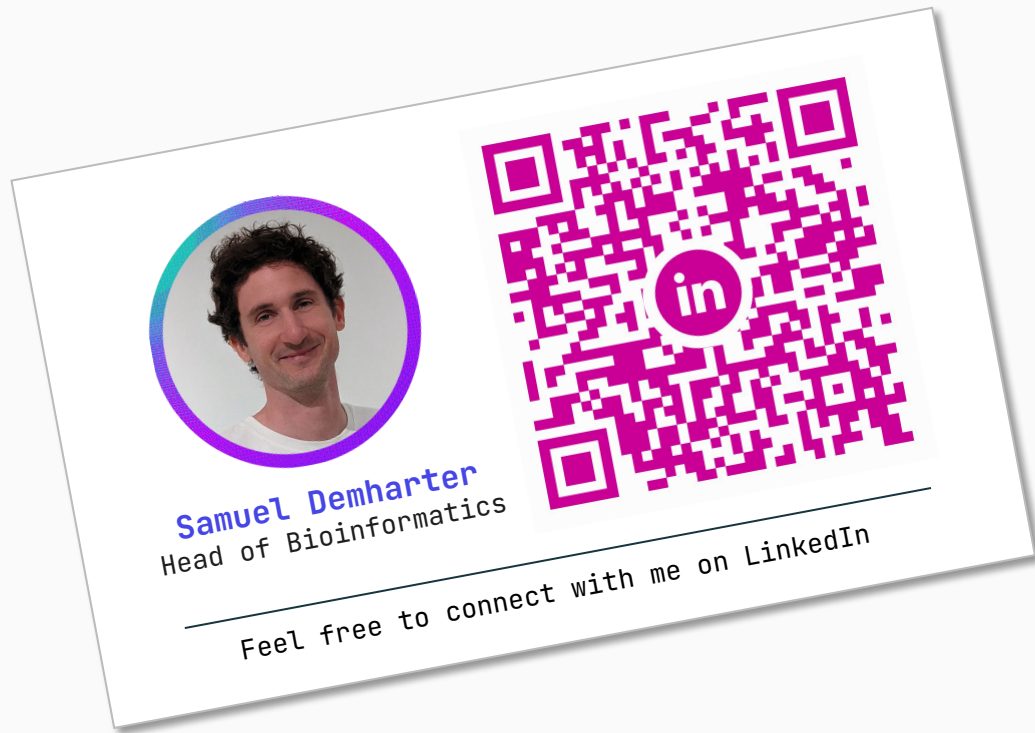
Explainable "White-box" machine learning is the way forward in pre-eclampsia screening [READ MORE >](#)  
 Michael Christiansen, MD, FRCPath, Casper Wilstrup, Paula L. Hedley, PhD, MPH, American Journal of Obstetrics and Gynecology, S0002-9378 (2022).

Identifying interactions in omics data for clinical biomarker discovery using symbolic regression [READ MORE >](#)  
 Niels Johan Christensen, Samuel Demharter, Meera Machado, Lykke Pedersen, Marco Salvatore, Voldemar Stentoft-Hansen, Miquel Triana Iglesias. *Bioinformatics* (2022).



Let's talk today







KNOCK ON OUR DOOR, CALL US, FOLLOW US

Thank you! And get in touch!


[sam.demharter@abzu.ai](mailto:sam.demharter@abzu.ai)


 [abzu.ai](https://abzu.ai)


 [info@abzu.ai](mailto:info@abzu.ai)

**Denmark** Orient Pl. 1, Mezz., 2150 Copenhagen, Denmark

**Spain** c/o Carrer d'Àlaba, 100, 08018 Barcelona, Spain

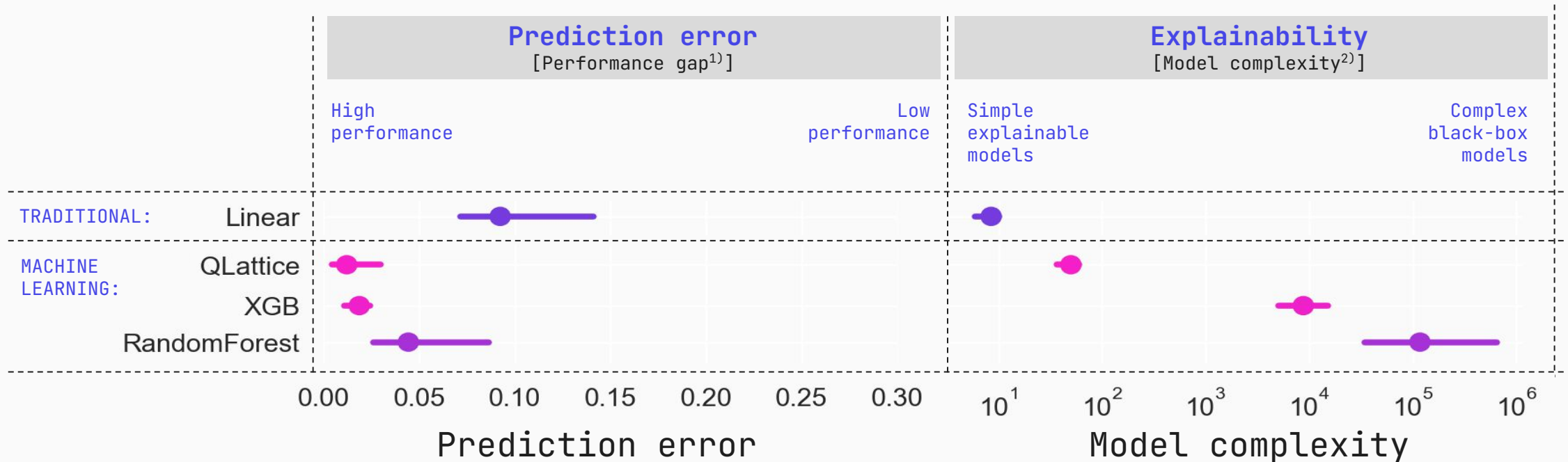
 [@AbzuAI](https://twitter.com/AbzuAI)

 [abzu](https://www.linkedin.com/company/abzu)

 [abzu\\_ai](https://www.instagram.com/abzu_ai)

DEMONSTRATED BENCHMARK

QLattice® has proven 'best of both worlds' in benchmarks



1): The performance gap is defined by the median difference of the given algorithm to the best performing model on each dataset in the real-world datasets of SRBench.  
 2): The model complexity is given the by SRBench definition of model size.