Generalizable biomarker development for early prostate cancer detection

Tiffany Tang
University of Notre Dame
Applied and Computational Mathematics and Statistics (ACMS)
ttang4@nd.edu

ENAR

March 24, 2025

UNIVERSITY OF NOTRE DAME

A Team Science

Yuping Zhang*

Ana Kenney

Arul Chinnaiyan

Bin Yu

Cassie Xie (Fred Hutchinson)
Lanbo Xiao (Univ. of Michigan)
Javed Siddiqui (Univ. of Michigan)
Sudhir Srivastava (NIH)
Martin G. Sanda (Emory)
John T. Wei (Univ. of Michigan)
Ziding Feng (Fred Hutchinson)
Jeffrey J. Tosoian (Vanderbilt)
Yingye Zheng (Fred Hutchinson)

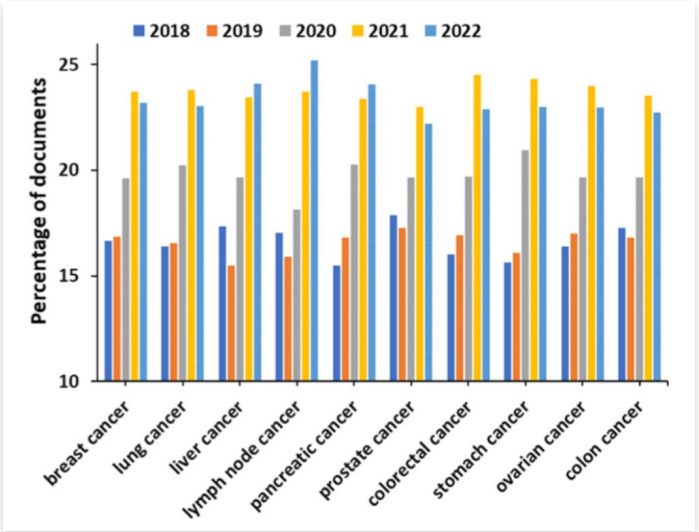
UNIVERSITY OF MICHIGAN

Berkeley
UNIVERSITY OF CALIFORNIA

NIH NATIONAL CANCER INSTITUTE
Early Detection Research Network

Biomarker development for early cancer detection

- **Early detection of cancer** often leads to more effective treatment and a drastically higher chance of survival
  - However, **~50%** of cancers are detected at **late stages** [Crosby et al. (2022)]
- **Biomarkers** have long offered a valuable opportunity to improve early cancer detection
  - Pancreatic cancer: CA19-9
  - Breast cancer: HER2, BRCA1/2
- Ensuring the **reliability** of biomarkers is the goal, but very challenging ("the valley of death")
  - >500 biomarkers fail rigorous EDRN validation [Srivastava (2023)]



Growth of biomarker-related papers over time [Tenchov et al. (2024)]

Biomarker development for early prostate cancer detection

Prostate cancer: a leading cause of cancer death in the developed world

~ 1 in 8 men are diagnosed with prostate cancer in their lifetime

~ 1 in 44 men will die of prostate cancer

- **Unclear benefits of current screening** procedures via prostate-specific antigen (PSA)
  - High rate of invasive biopsies and false positives (i.e., overdiagnosis and overtreatment of indolent cancers)
- Can we create a reliable **non-invasive urine-based biomarker test** to detect prostate cancer with greater accuracy than PSA?

Source: American Cancer Society (2025)

Prior Work: MyProstateScore2.0 (MPS2)

Research

JAMA Oncology | Original Investigation

Development and Validation of an 18-Gene Urine Test for High-Grade Prostate Cancer

Jeffrey J. Tosoian, MD, MPH; Yuping Zhang, PhD; Lanbo Xiao, PhD; Cassie Xie, MS; Nathan L. Samora, MD; Yashar S. Niknafs, PhD; Zoey Chopra, MA; Javed Siddiqui, MS; Heng Zheng, MD; Grace Herron, BA; Neil Vaishampayan, BS; Hunter S. Robinson, MD; Kumaran Arivoli, BS; Bruce J. Trock, PhD; Ashley E. Ross, MD, PhD; Todd M. Morgan, MD; Ganesh S. Palapattu, MD; Simpa S. Salami, MD, MPH; Lakshmi P. Kunju, MD; Scott A. Tomlins, MD, PhD; Lori J. Sokoll, PhD; Daniel W. Chan, PhD; Sudhir Srivastava, PhD; Ziding Feng, PhD; Martin G. Sanda, MD; Yingye Zheng, PhD; John T. Wei, MD; Arul M. Chinnaiyan, MD, PhD; for the EDRN-PCA3 Study Group

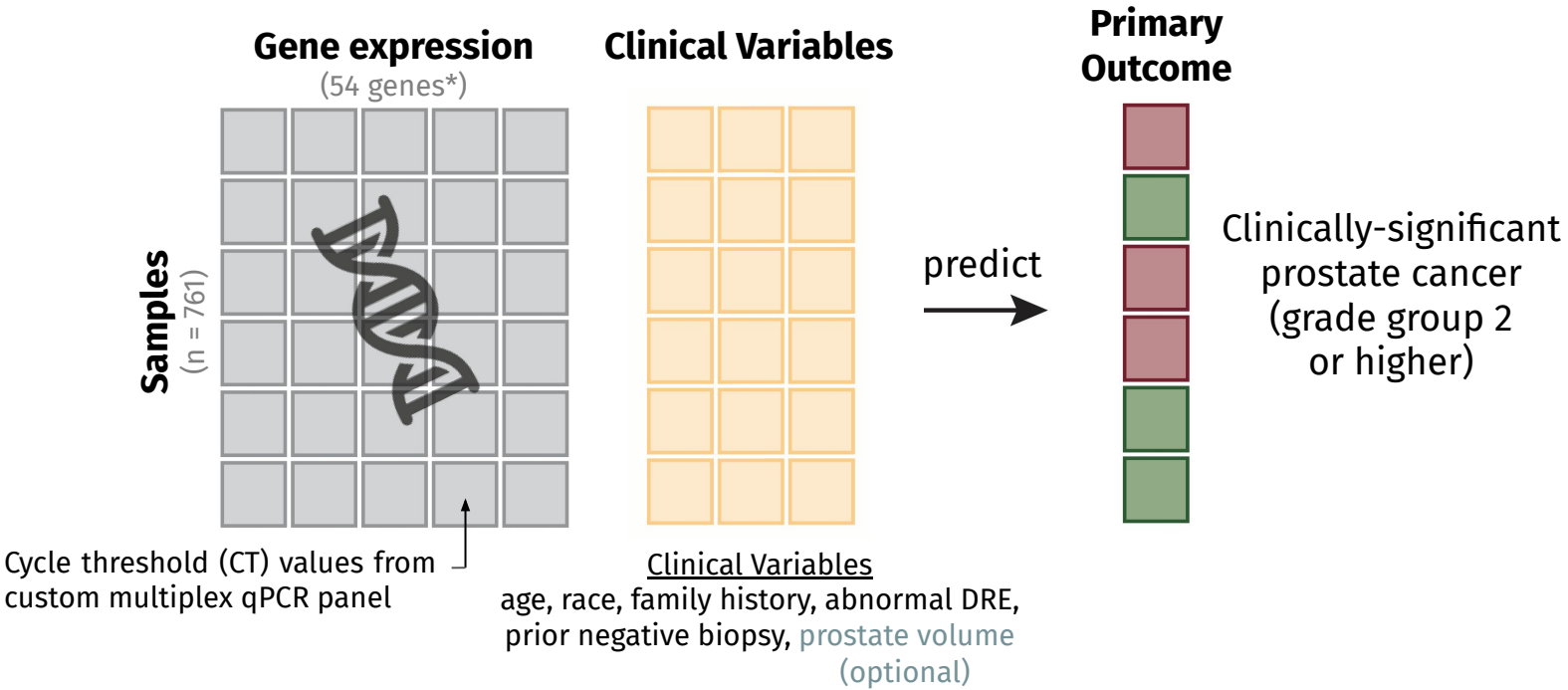
IMPORTANCE Benefits of prostate cancer (PCa) screening with prostate-specific antigen (PSA) alone are largely offset by excess negative biopsies and overdetection of indolent cancers resulting from the poor specificity of PSA for high-grade PCa (ie, grade group [GG] 2 or greater).

OBJECTIVE To develop a multiplex urinary panel for high-grade PCa and validate its external performance relative to current guideline-endorsed biomarkers.

Supplemental content

Prior Work: MyProstateScore2.0 (MPS2)

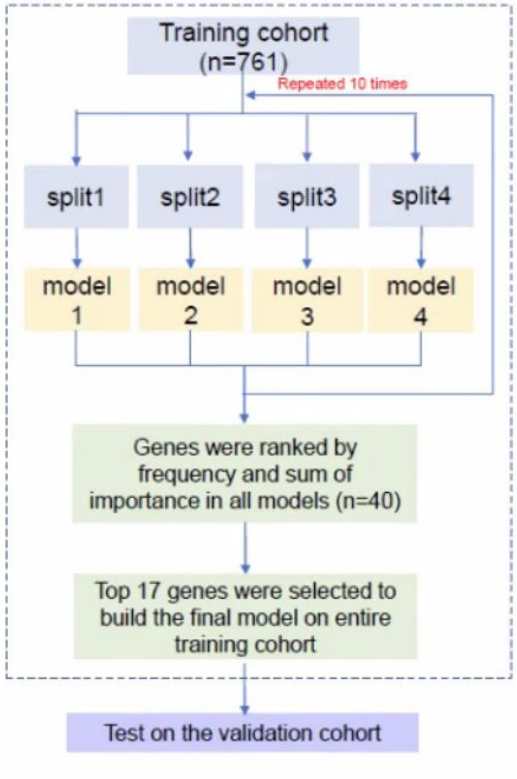
Training/Development Cohort Data



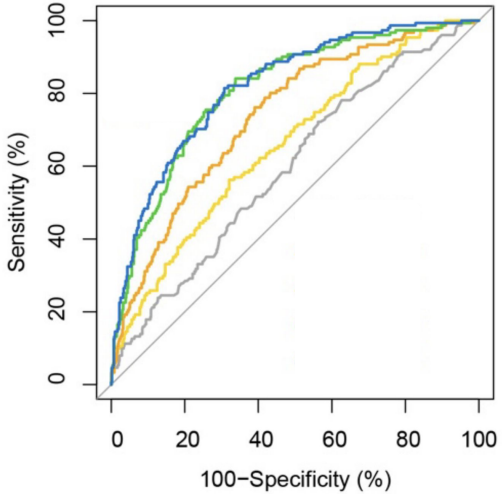
* Carefully selected out of 58,724 genes based on differential expression and predefined nomination criteria

Prior Work: MyProstateScore2.0 (MPS2)

Model Development



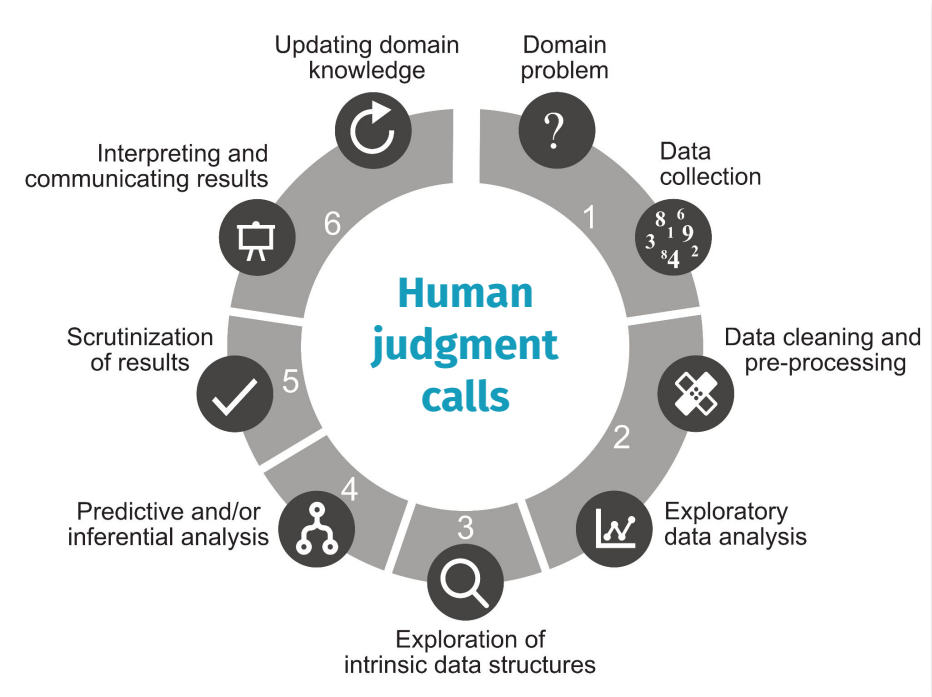
- **Final (locked) model:** logistic regression + elastic net using 18 genes and clinical variables as predictors
  - Why 18? Chosen to fit in OpenArray™ platform
- Evaluated on **external validation cohort from EDRN**



AUROC	Method	
81.8%	MPS2+	18 genes with prostate volume
80.7%	MPS2	18 genes without prostate volume
73.7%	MPS	3 genes + clinical
65.9%	PCPT	clinical only
59.7%	PSA	current standard

What's the problem? Unaccounted uncertainty

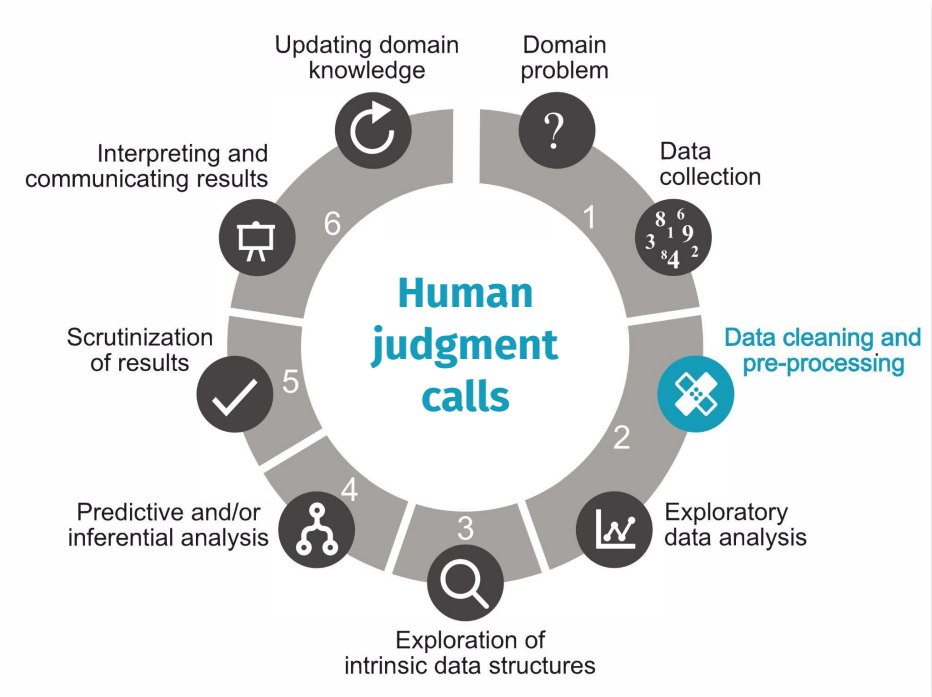
There are numerous **human judgment calls** (or choices) throughout MPS2 development



The Data Science Life Cycle [Yu and Barter (2024)]

What's the problem? Unaccounted uncertainty

There are numerous **human judgment calls** (or choices) throughout MPS2 development



The Data Science Life Cycle [Yu and Barter (2024)]

Data Cleaning Choices

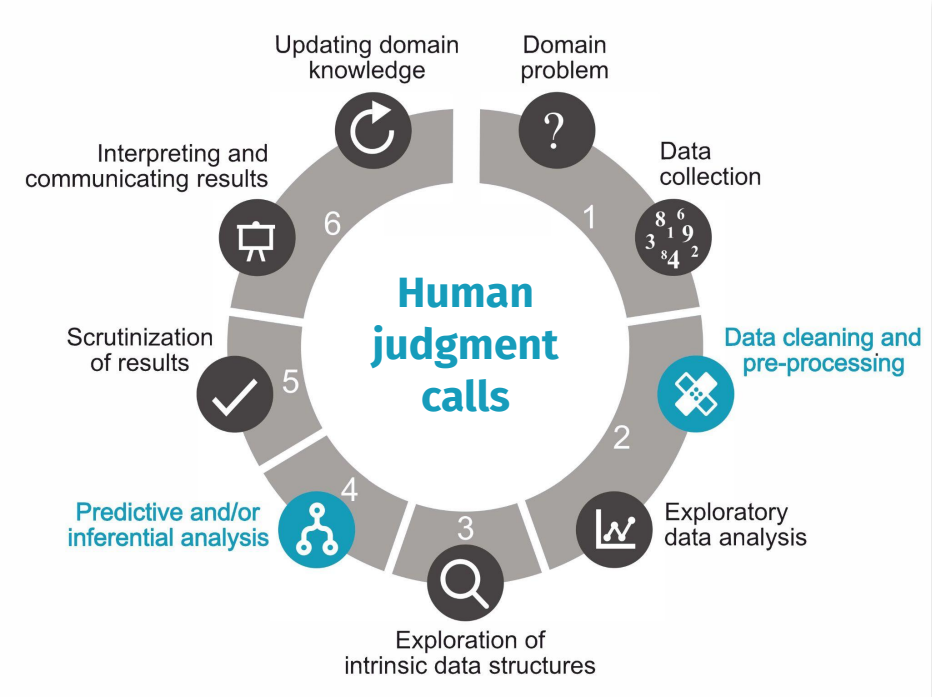
- How to threshold CT values if undetermined or no amplification
- Sample filtering/quality control choices
- ...

What's the problem? Unaccounted uncertainty

There are numerous **human judgment calls** (or choices) throughout MPS2 development

Modeling Choices

- Which model?
- Which clinical variables?
- Which genes?
- ...



The Data Science Life Cycle [Yu and Barter (2024)]

Data Cleaning Choices

- How to threshold CT values if undetermined or no amplification
- Sample filtering/quality control choices
- ...

At a minimum, our scientific conclusions should be **stable** across **reasonable data and modeling perturbations**

[Ioannidis (2005); Kraft et al. (2009); Donoho (2010); Casadevall and Fang (2011); Nosek et al. (2012); Yu (2013), Gelman and Loken (2014), ...]

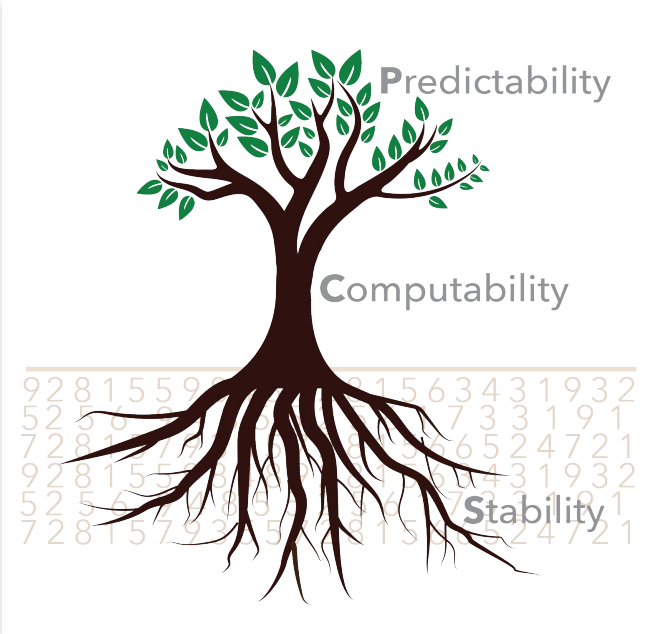
PCS Framework for Veridical Data Science [Yu and Kumbier (2020)]

Three principles for **veridical (trustworthy) data science**

Predictability: is my model a good representation of reality (as measured by prediction accuracy)?

Computability: is my pipeline computable?

Stability: are my model/findings stable across reasonable perturbations of the data science life cycle?



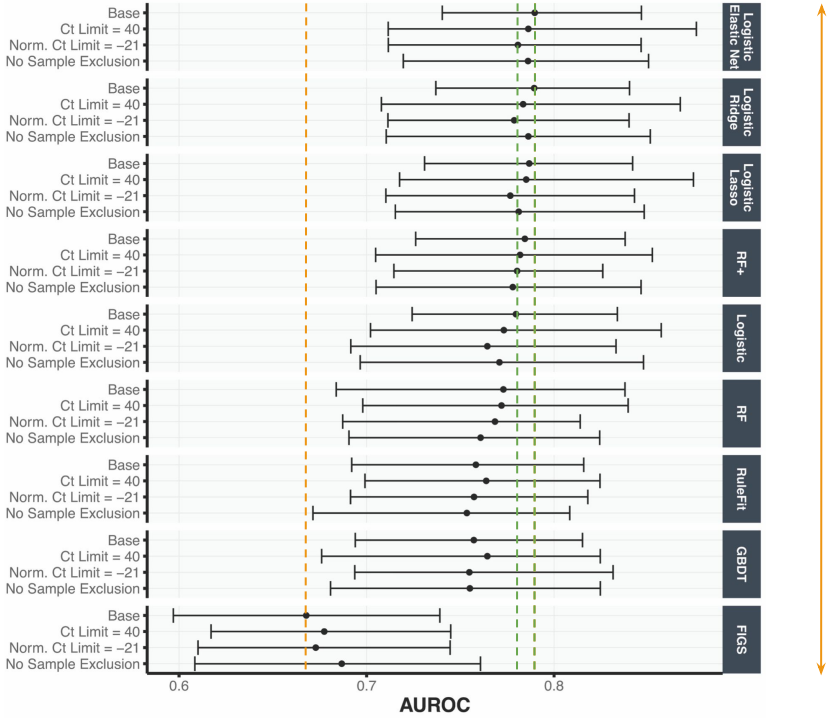
Credits: Rebecca Barter

Stress-testing MPS2 under the PCS framework

Data preprocessing pipelines

<2% difference in mean AUROC across data preprocessing pipelines

Inner 95% quantile range of AUROCs from repeated CV



Methods

Up to >10% difference in mean AUROC across prediction methods

Difference between mean AUROC across data preprocessing pipelines << across methods

Improving MPS2 using the PCS framework

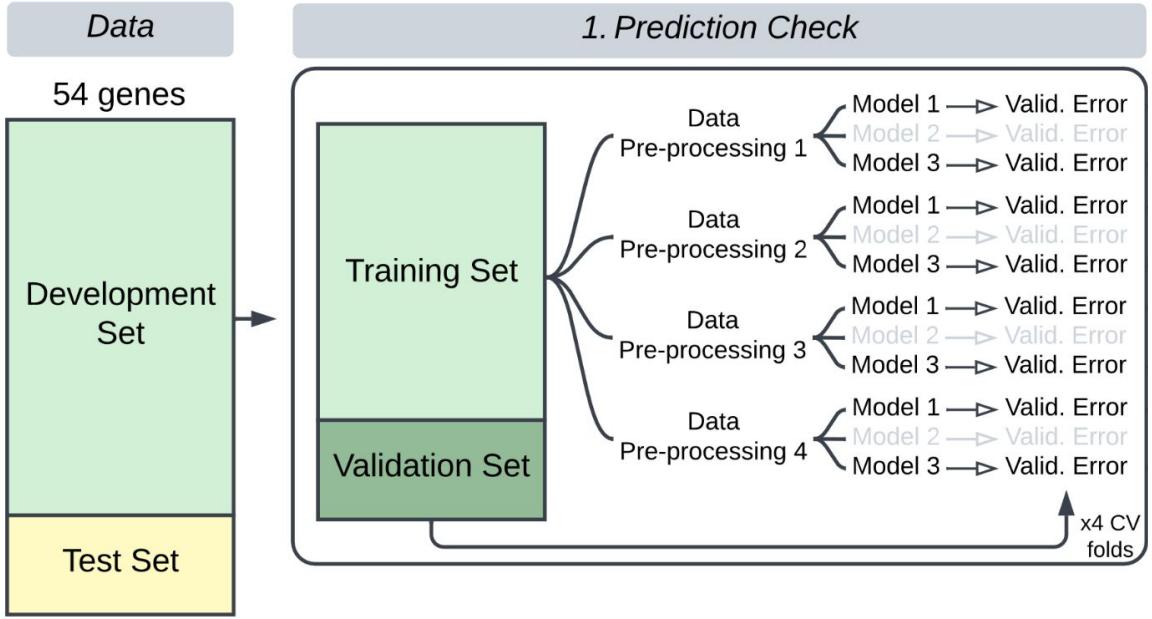
Beyond stress-testing, the PCS framework can also be used to improve the model development process.

Example: Do we need all 18 genes or can we develop a simpler, cheaper gene panel?

- We developed a **simplified MPS2 (sMPS2)** model, which uses only **7 genes** and achieves similar accuracy as the 18-gene MPS2 model.
- Key Steps:
  - Use **prediction performance as a reality check** and exclude models that don't fit the data well
  - Select **features** that were **stably important** across **data and model perturbations**
  - Very careful **data splitting**

sMPS2: Model Development

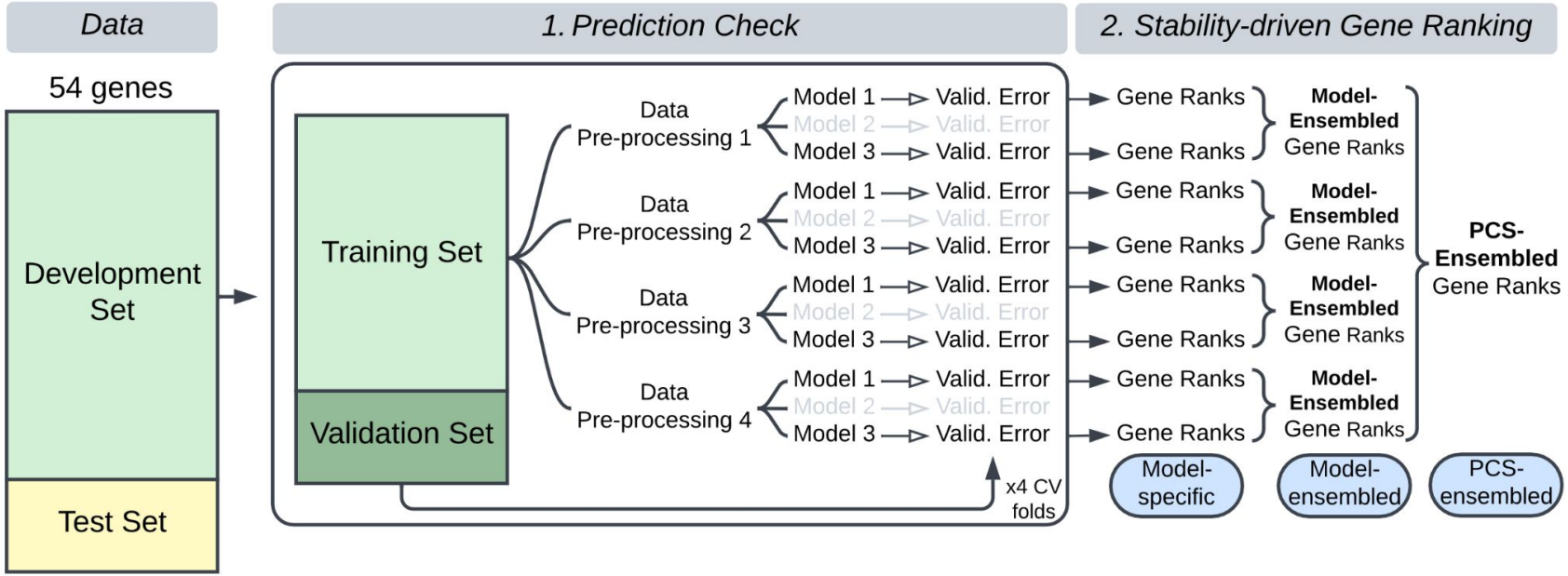
sMPS2: Model Development



Use prediction performance as a reality check

(Exclude pipelines (e.g., FIGS) with poor fits)

sMPS2: Model Development

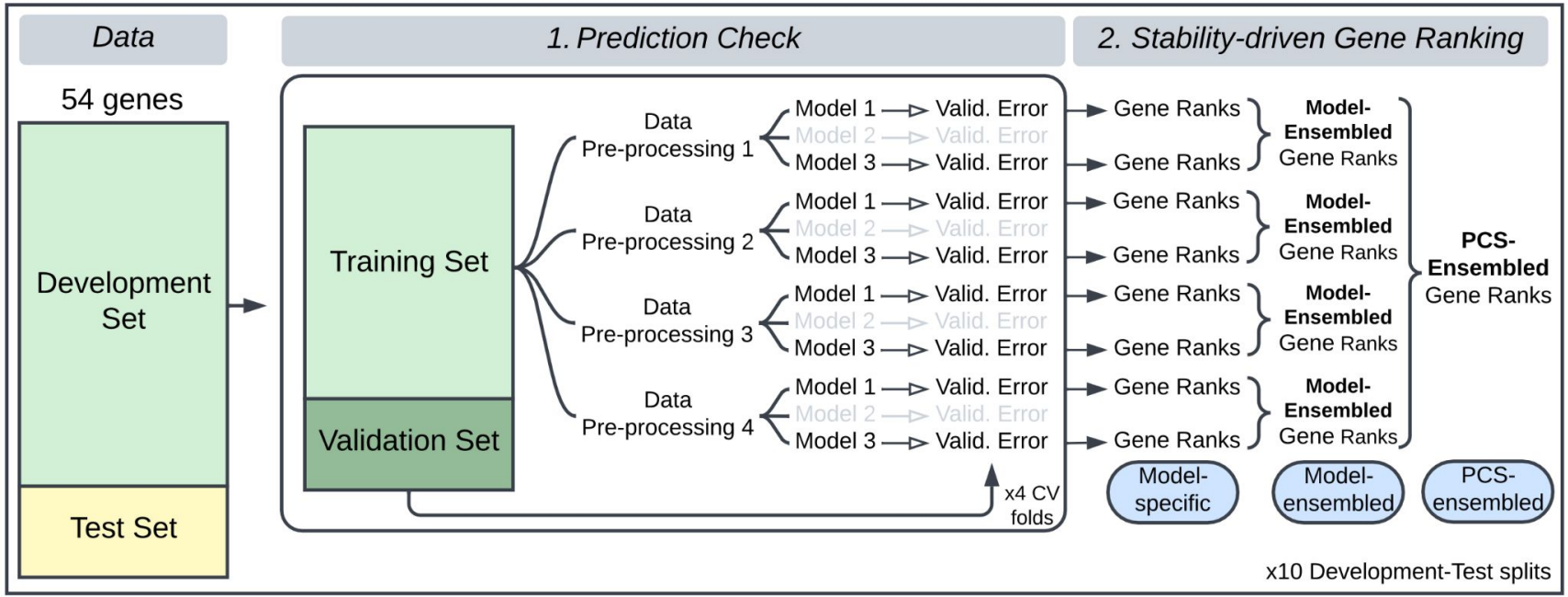


Use prediction performance as a reality check

(Exclude pipelines (e.g., FIGS) with poor fits)

Ensemble feature importances across data & model perturbations

sMPS2: Model Development



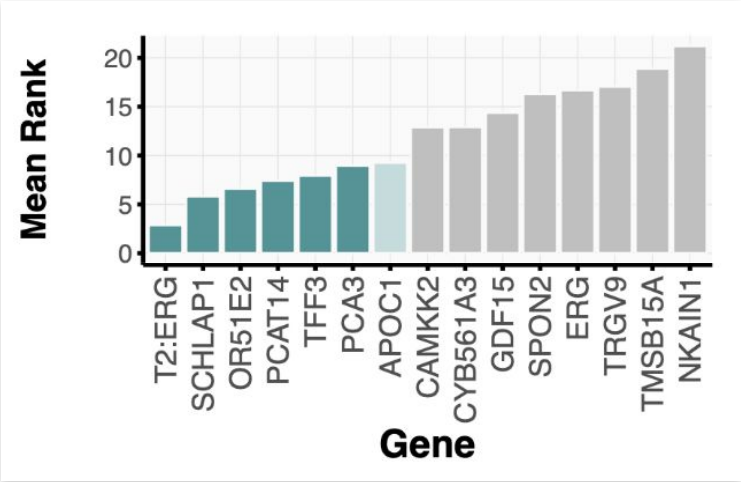
Use prediction performance as a reality check

(Exclude pipelines (e.g., FIGS) with poor fits)

Ensemble feature importances across data & model perturbations

sMPS2: Model Development

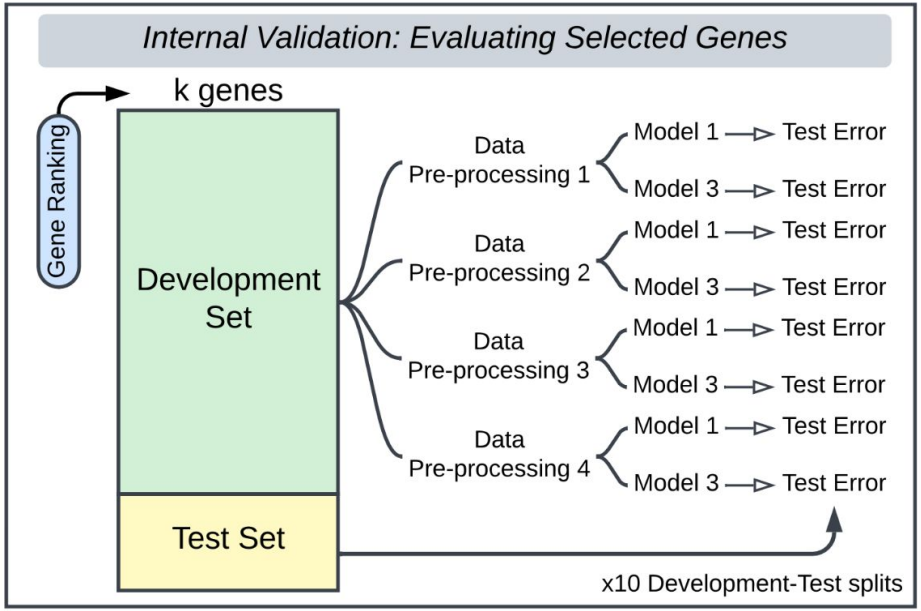
Examining these PCS-ensembled gene rankings reveals 6-7 very **staby important genes**

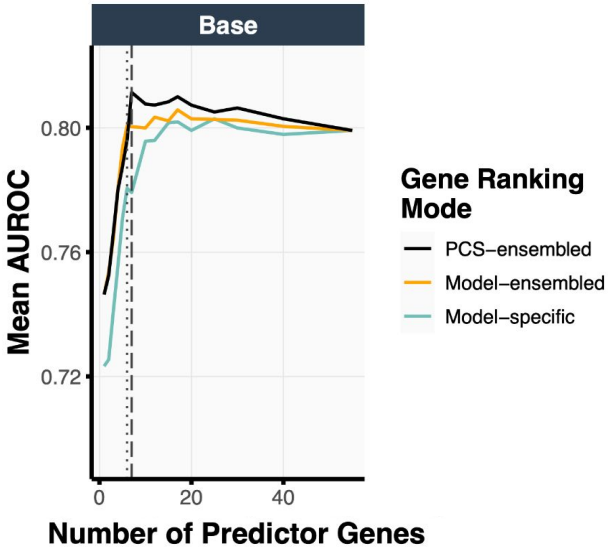


* Mean ranking is only one stability metric
  + There are many others ways to evaluate stability (e.g., % in top *k*)
  + Other stability metrics showed APOC1 is not as stable

sMPS2: Internal Validation

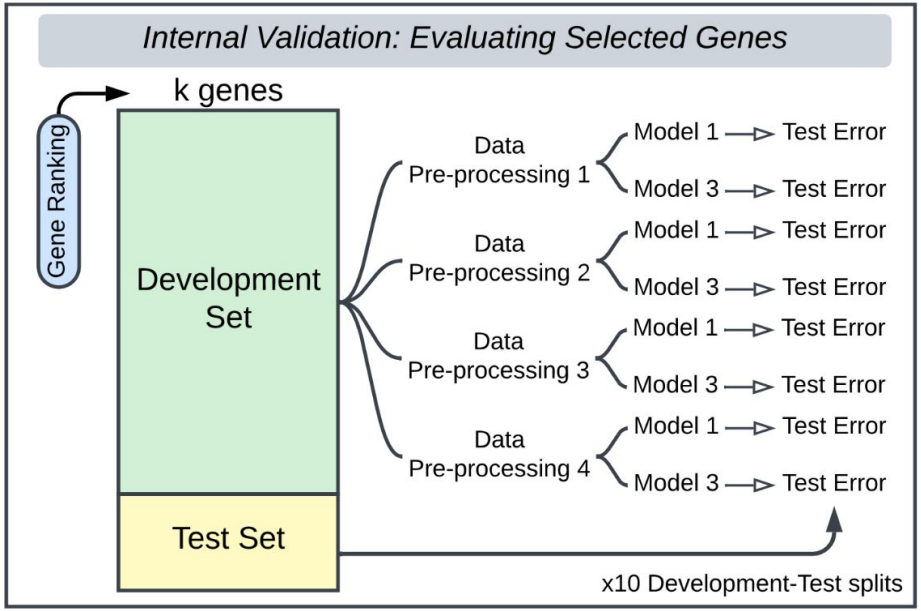
- 6-7 top stably important genes yielded the best (or competitively high) test AUROC compared to using different number of gene predictors
- PCS-ensembled gene rankings > model-ensembled or model-specific

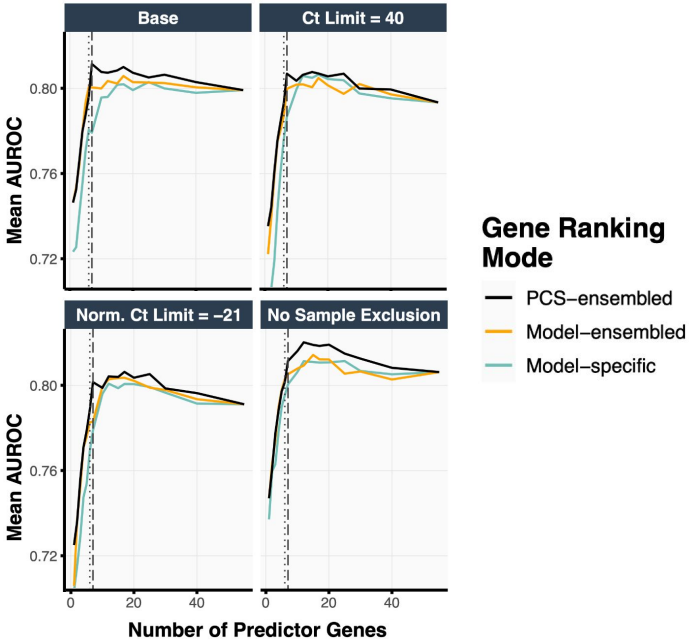




sMPS2: Internal Validation

- 6-7 top stably important genes yielded the best (or competitively high) test AUROC compared to using different number of gene predictors
- PCS-ensembled gene rankings > model-ensembled or model-specific





sMPS2: External Validation (EDRN)

+ **Final (locked) model:** logistic regression + ridge using 7 genes (6 stably important genes + 1 reference gene *KLK3*) and clinical variables

Without prostate volume

AUROC	Method	
80.7%	MPS2	18 genes
78.5%	**s^{8}MPS2**	8 genes
78.4%	**s^{7}MPS2**	7 genes
73.7%	MPS	3 genes
59.7%	PSA	current standard

With prostate volume

AUROC	Method	
81.8%	MPS2+	18 genes
80.9%	**s^{8}MPS2+**	8 genes
80.6%	**s^{7}MPS2+**	7 genes
73.7%	MPS	3 genes
59.7%	PSA	current standard

+ Difference between MPS2 and sMPS2 is smaller than uncertainty due to data preprocessing choices (~2%)

Summary and Discussion

Using the PCS framework, we:

- + Rigorously stress-tested MPS2 for early prostate cancer detection
- + Simplified the 18-gene MPS2 model to a 7-gene sMPS2 model with similar accuracy

Still, there are human judgment calls throughout this analysis

→ needs documentation and justification

Thank you!

Email: ttang4@nd.edu
Website: tiffanymtang.github.io/