Predictive Analytics for Optimal Detection of Metastatic Prostate Cancer

Selin Merdan[1], Christine Barnett[1], Brian Denton[1,2], David Miller[2], James Montie[2]

[1]Department of Industrial and Operations Engineering

[2]Department of Urology

April 2, 2021

Prostate cancer is the most common cancer among men

- It is estimated that 248,000 new cases of PCa would be diagnosed and 34,000 men would die of the disease in the United States in 2021
- Clinical staging determines the extent of the disease
- Accurate clinical staging is necessary to determine the optimal course of treatment

https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html

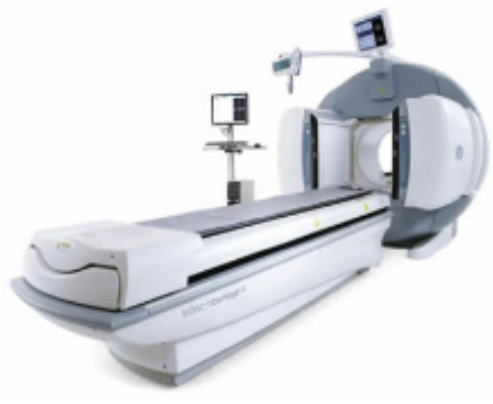
Noninvasive imaging methods can detect bone and lymph node metastasis

Bone Scan (BS)

- Time-consuming procedure ($3 - 4$ hours)
- Costs $\$600 - \$1,000$

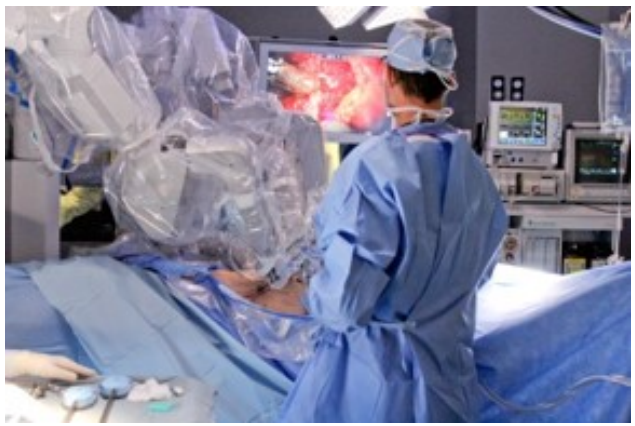
Computed Tomography (CT)

- Exposes patient to 60 times more radiation than an x-ray
- Costs $\$300 - \$1,500$



Significant harms associated with missing a case of metastatic cancer

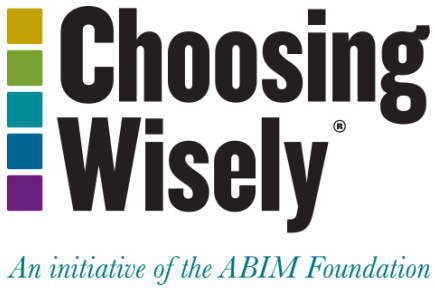
- Negative health outcomes due to delays in chemotherapy



- Missed diagnoses subject patients to unnecessary treatments (e.g., radical prostatectomy) that cause serious side effects

Avoidance of imaging in low-risk PCa top priority for AUA "Choosing Wisely" campaign

- Potentially harmful radiation exposure



- Incidental findings that require painful and risky follow-up procedures (e.g., bone biopsy)
- Blocks access to imaging resources for other patients and unnecessarily increases healthcare costs

Conflicting imaging guidelines for PCa staging

Bone scan (BS)		CT scan	
Clinical guidelines	**Recommend imaging if any of these:**	**Clinical guidelines**	**Recommend imaging if any of these:**
EAU	GS ≥ 8 cT3/T4 disease PSA > 10 ng/ml Symptomatic	**EAU**	GS ≥ 8 cT3/T4 disease PSA > 10 ng/ml Symptomatic
AUA	GS ≥ 8 PSA > 10 ng/ml Symptomatic	**AUA**	GS ≥ 8 PSA > 20 ng/ml cT3/T4 disease Symptomatic
NCCN	cT1 disease & PSA > 20 ng/ml cT2 disease & PSA > 10 ng/ml GS ≥ 8 cT3/T4 disease Symptomatic		
Briganti's CART	GS ≥ 8 $\geq$ cT2 disease & PSA > 10 ng/ml Symptomatic		

EAU: European Urological Association; AUA: American Urological Association; NCCN: National Comprehensive Cancer Network; CART: classification and regression tree.

Michigan Urological Surgery Improvement Collaborative

- Physician-led, statewide collaborative sponsored by the Blue Cross Blue Shield of Michigan
- 43 urology practices from throughout Michigan ($> 90\%$ of urologists in the state)
- Complete preoperative data for men with newly-diagnosed PCa were retrospectively reviewed

Project objectives

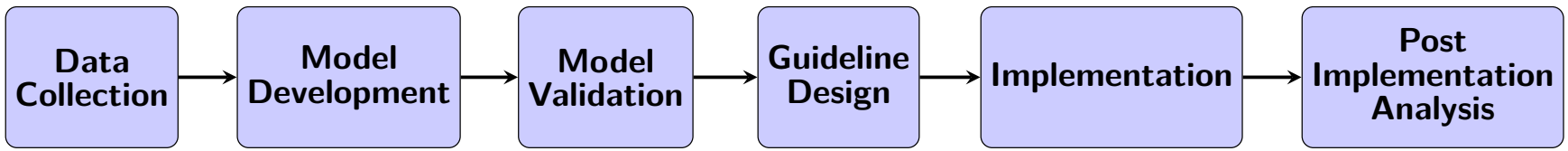
1. To create data-analytic methods to provide clinical recommendations that weigh the benefits and harms of radiological imaging

2. To determine which patients should receive a BS and/or a CT scan and which patients can safely avoid imaging on the basis of individual risk factors

Methodology

The methodological approach consists of

- Multidisciplinary problem assessment and framing
- Development and validation of descriptive, predictive, and prescriptive models for the detection of metastasis
- Model validation and post-implementation measurement of impact



Project Activities

Predictors of metastatic cancer risk

- Descriptive statistical analyses were conducted to assess the relation between imaging test outcomes and clinical risk factors including:
  - Patient demographics
  - Prostate-specific antigen (PSA) (ng/ml)
  - Clinical tumor stage (e.g., T1a/b/c, T2a/b/c and T3/4)
  - Pathology reports

Statistical validation

Internal validation

- Validating a predictive model using the development sample will introduce bias known as optimism
- Bootstrapping is an efficient technique that addresses this bias

External validation

- Independent datasets are used to validate the predictive models

Classification modeling using observational data

- Two important challenges:
  - Learning from unlabeled data
    - In practice not all patients receive imaging at diagnosis
  - Learning from imbalanced data
    - A minority of patients has metastatic cancer
- To address these challenges, two machine learning paradigms are studied:
  - Semi-supervised learning
  - Cost-sensitive learning

Proposed classification model: Cos-LapKLR

- Cost-sensitive Laplacian Kernel Logistic Regression (Cos-LapKLR) is a spectral clustering based semi-supervised learning approach
- A weighted graph G is created with vertices corresponding to all patients
  - If patient $\mathbf{x}_i$ is among the k-nearest neighbors of $\mathbf{x}_j$, they are connected by an edge with a nonnegative weight w_{ij}
  - The Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D} = \mathrm{diag}\,(d_1, \ldots, d_{l+u})$ and $d_i = \sum_{j=1}^{l+u} w_{ij}$

Belkin M, Niyogi P, Sindhwani V "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples", *The Journal of Machine Learning Research*, 2006.

Proposed classification model: Cos-LapKLR

- We introduce cost-sensitive logistic loss:

$$f^* = \underset{f\in\mathcal{H}}{\operatorname{argmin}} \frac{1}{l}\sum_{i=1}^{l} \delta \mathbb{1}_{\{y_i=1\}} \log\left(1+e^{-f(\mathbf{x}_i)}\right) + (1-\delta)\mathbb{1}_{\{y_i=-1\}} \log\left(1+e^{f(\mathbf{x}_i)}\right)$$

$$+ \quad \gamma_{\mathcal{H}}\|f\|_{\mathcal{H}}^2 \quad + \quad \gamma_{\mathcal{M}}\mathbf{f^T L f}$$

f: the decision function, $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(\mathbf{x}_i, \mathbf{x})$, u: the number of unimaged patients, $\mathbf{K}$: the positive definite kernel function, $\mathbf{L}$: the Laplacian matrix.

Proposed classification model: Cos-LapKLR

- We introduce cost-sensitive logistic loss:

$$f^* = \underset{f\in\mathcal{H}}{\operatorname{argmin}} \frac{1}{l}\sum_{i=1}^{l} \delta \mathbb{1}_{\{y_i=1\}} \log\left(1+e^{-f(\mathbf{x}_i)}\right) + (1-\delta)\mathbb{1}_{\{y_i=-1\}} \log\left(1+e^{f(\mathbf{x}_i)}\right)$$

$$+ \quad \gamma_{\mathcal{H}}\|f\|_{\mathcal{H}}^2 \quad + \quad \gamma_{\mathcal{M}}\mathbf{f^T L f}$$

Higher cost on missing metastatic cancers

f: the decision function, $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(\mathbf{x}_i, \mathbf{x})$, u: the number of unimaged patients, $\mathbf{K}$: the positive definite kernel function, $\mathbf{L}$: the Laplacian matrix.

Proposed classification model: Cos-LapKLR

- We introduce cost-sensitive logistic loss:

$$f^* = \underset{f\in\mathcal{H}}{\operatorname{argmin}} \frac{1}{l}\sum_{i=1}^{l} \delta \mathbb{1}_{\{y_i=1\}} \log\left(1+e^{-f(\mathbf{x}_i)}\right) + (1-\delta)\mathbb{1}_{\{y_i=-1\}} \log\left(1+e^{f(\mathbf{x}_i)}\right) + \gamma_{\mathcal{H}}\|f\|_{\mathcal{H}}^2 + \gamma_{\mathcal{M}}\mathbf{f^T L f}$$

Higher cost on missing metastatic cancers

Avoid overfitting

f: the decision function, $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(\mathbf{x}_i, \mathbf{x})$, u: the number of unimaged patients, $\mathbf{K}$: the positive definite kernel function, $\mathbf{L}$: the Laplacian matrix.

Proposed classification model: Cos-LapKLR

- We introduce cost-sensitive logistic loss:

$$f^* = \underset{f\in\mathcal{H}}{\operatorname{argmin}} \frac{1}{l}\sum_{i=1}^{l} \delta \mathbb{1}_{\{y_i=1\}} \log\left(1+e^{-f(\mathbf{x}_i)}\right) + (1-\delta)\mathbb{1}_{\{y_i=-1\}} \log\left(1+e^{f(\mathbf{x}_i)}\right) + \gamma_{\mathcal{H}}\|f\|_{\mathcal{H}}^2 + \gamma_{\mathcal{M}}\mathbf{f^T L f}$$

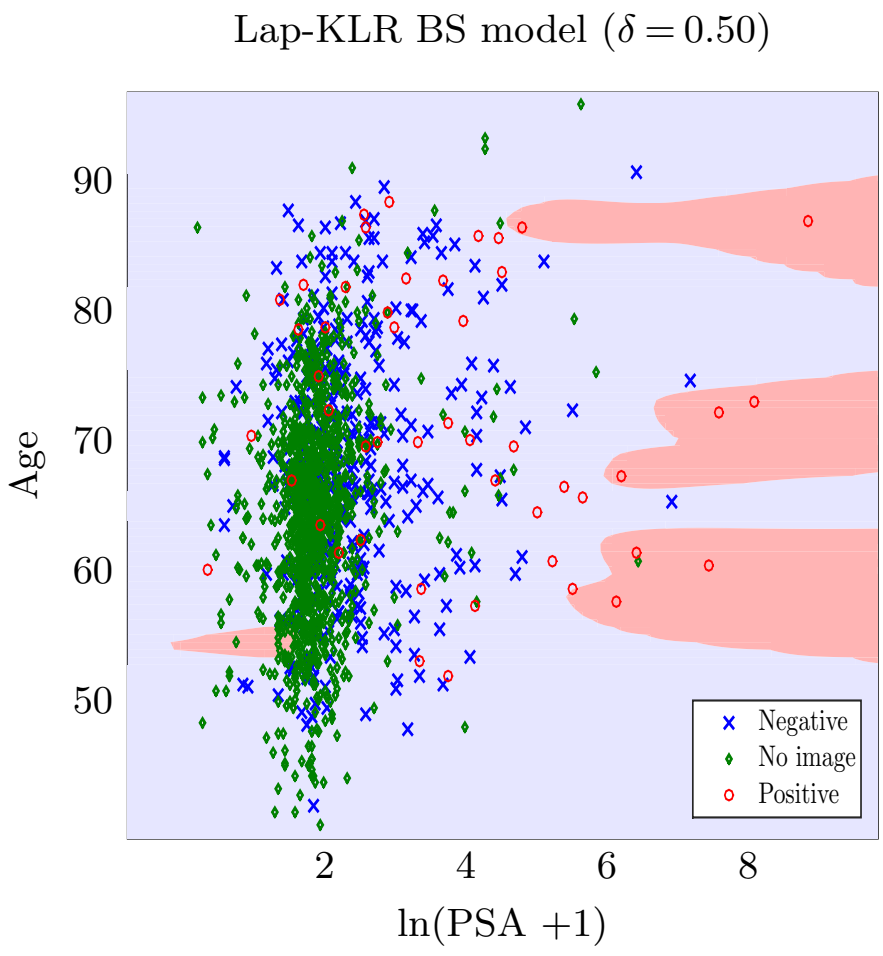
Higher cost on missing metastatic cancers

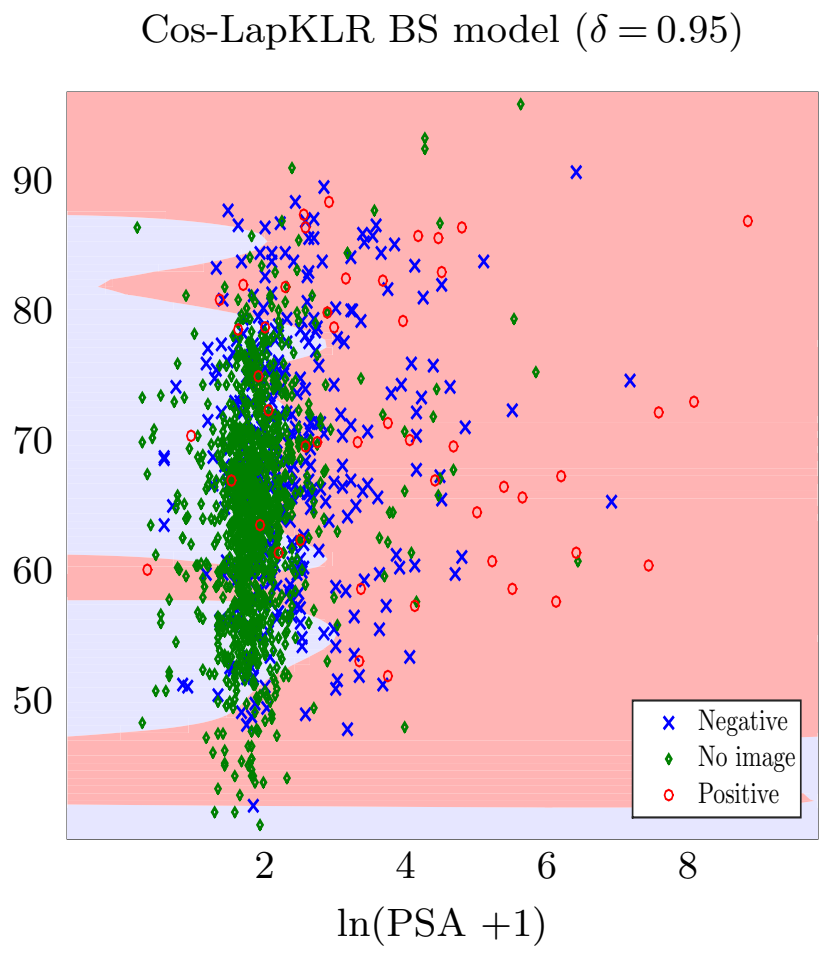
Avoid overfitting

Extract information from unimaged patients

f: the decision function, $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(\mathbf{x}_i, \mathbf{x})$, u: the number of unimaged patients, $\mathbf{K}$: the positive definite kernel function, $\mathbf{L}$: the Laplacian matrix.

More accurate nonlinear classification rules can be obtained via Cos-LapKLR

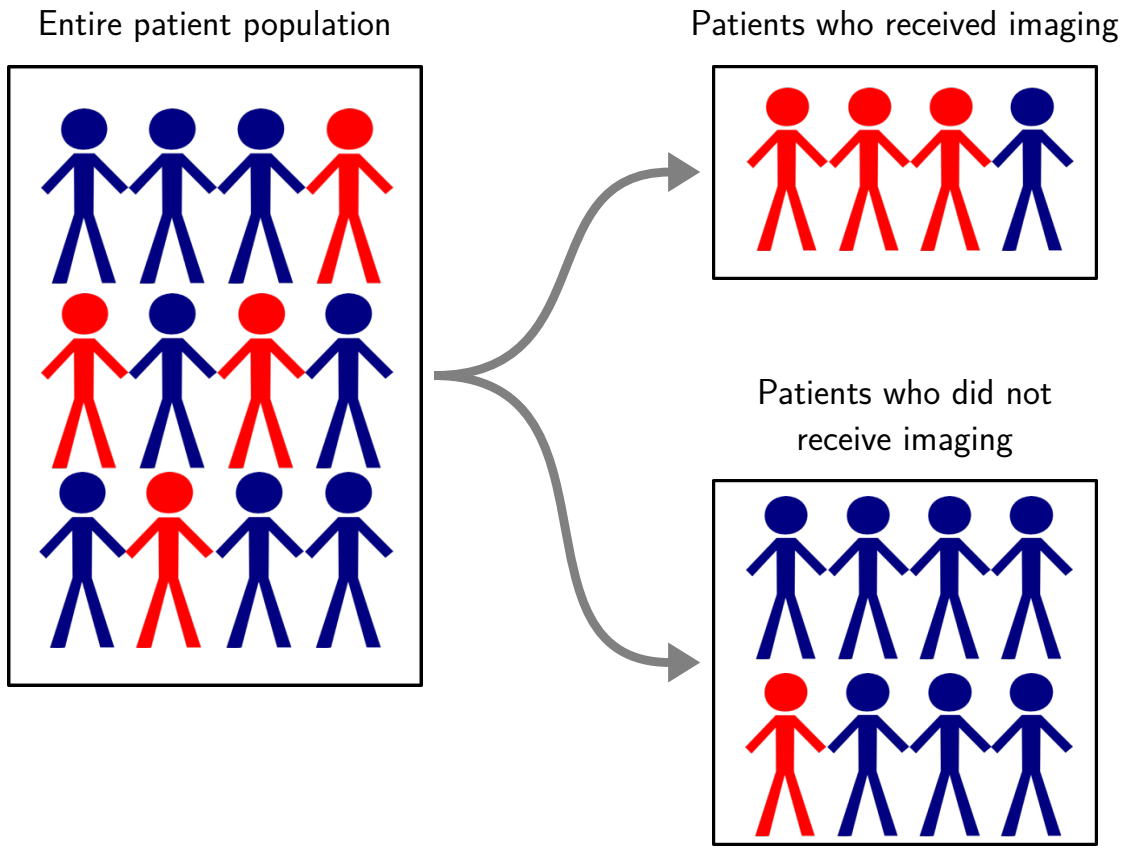




Alternative classification models for comparison

- Several other classification models adapted for imbalance data learning are implemented:
  - Cost-sensitive logistic regression and support vector machines
  - Random forests and AdaBoost combined with advanced resampling techniques

Verification bias



Verification bias correction

- We develop a logistic regression model to estimate the adjusted sensitivity and specificity based on the entire patient population:

$$\mathbb{P}(\mathsf{G}+ \mid \text{Disease present}) = \frac{\mathbb{P}(\text{Disease present} \mid \mathsf{G}+)\mathbb{P}(\mathsf{G}+)}{\mathbb{P}(\text{Disease present})}$$

$$\mathbb{P}(\mathsf{G}- \mid \text{Disease not present}) = \frac{\mathbb{P}(\text{Disease not present} \mid \mathsf{G}-)\mathbb{P}(\mathsf{G}-)}{\mathbb{P}(\text{Disease not present})}$$

- Main assumption: risk factors considered by the guideline are the only factors that influence imaging decisions

Begg, C. B., Greenes, R. A. "Assessment of diagnostic tests when disease verification is subject to selection bias" *Biometrics*, 1983.

Verification bias greatly impacts the apparent performance of imaging guidelines

	Uncorrected		Bias-corrected	
Clinical guidelines	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
Bone scan				
EAU	97.9	33.4	84.5	75.7
AUA	97.9	43.5	81.2	82.0
NCCN	97.9	40.8	82.3	80.9
Briganti's CART	89.6	45.4	79.3	83.3
CT scan				
EAU	98.4	36.5	89.9	74.4
AUA	96.8	49.2	87.2	82.5

The numbers are the percentages. EAU: European Urological Association; AUA: American Urological Association; NCCN: National Comprehensive Cancer Network; CART: classification and regression tree.

Verification bias greatly impacts the apparent performance of imaging guidelines

	Uncorrected		Bias-corrected	
Clinical guidelines	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
Bone scan				
EAU	97.9	33.4	84.5	75.7
AUA	**97.9**	43.5	**81.2**	82.0
NCCN	97.9	40.8	82.3	80.9
Briganti's CART	89.6	45.4	79.3	83.3
CT scan				
EAU	98.4	36.5	89.9	74.4
AUA	96.8	49.2	87.2	82.5

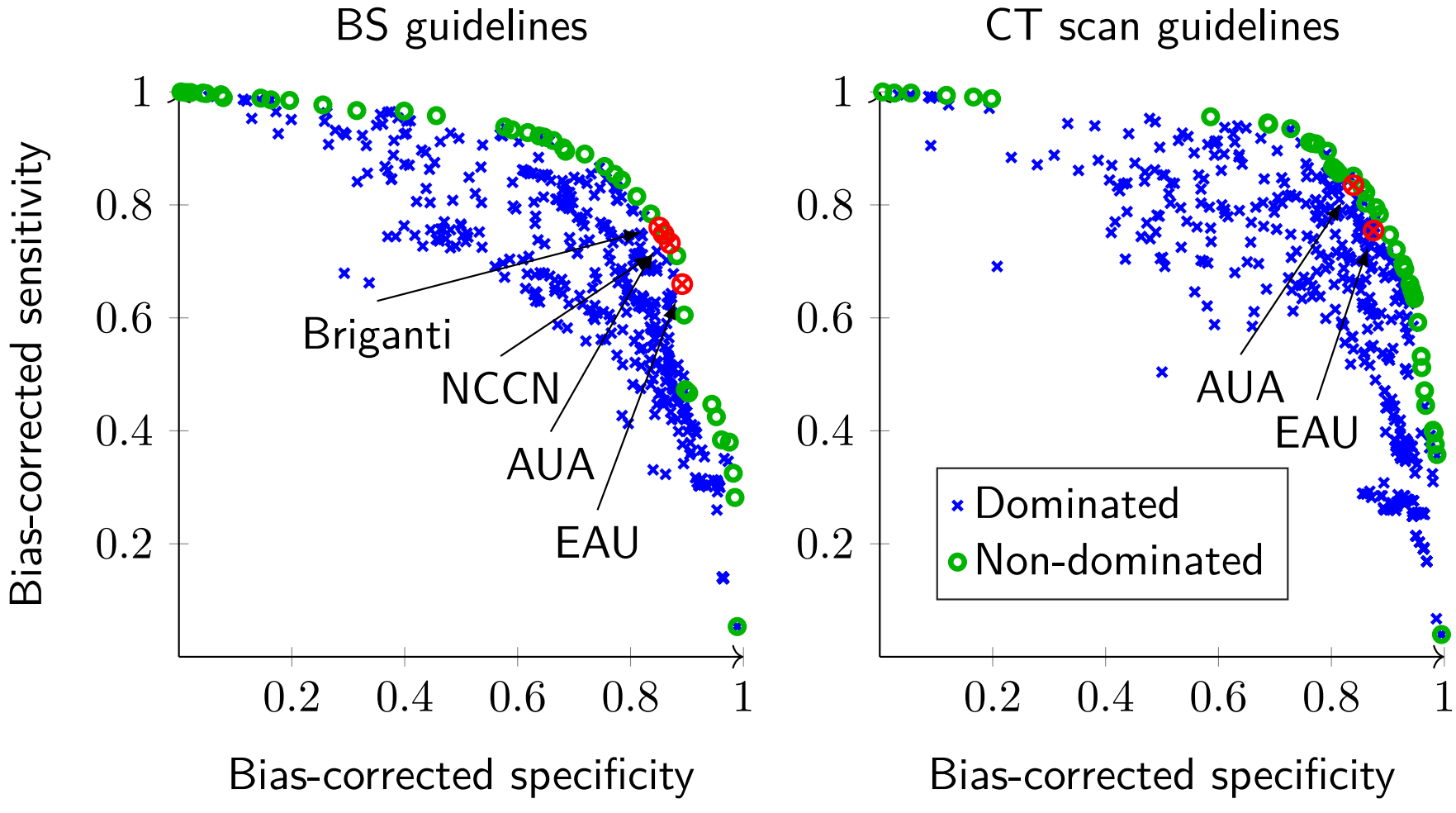
The numbers are the percentages. EAU: European Urological Association; AUA: American Urological Association; NCCN: National Comprehensive Cancer Network; CART: classification and regression tree.

Verification bias greatly impacts the apparent performance of imaging guidelines

	Uncorrected		Bias-corrected	
Clinical guidelines	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
Bone scan				
EAU	97.9	33.4	84.5	75.7
AUA	**97.9**	**43.5**	**81.2**	**82.0**
NCCN	97.9	40.8	82.3	80.9
Briganti's CART	89.6	45.4	79.3	83.3
CT scan				
EAU	98.4	36.5	89.9	74.4
AUA	96.8	49.2	87.2	82.5

The numbers are the percentages. EAU: European Urological Association; AUA: American Urological Association; NCCN: National Comprehensive Cancer Network; CART: classification and regression tree.

Published guidelines are near-Pareto optimal



Predicted impact of recommendations if implemented

Bone scan PSA > 20 or GS > 7	CT scan PSA > 20 or GS > 7 or clinical T stage $\geq$ T3
■ **20.7%** (prev **27%**) of patients would be scanned – Of those, **17.0%** (prev **12%**) would be positive ■ Estimated **0.8%** of patients have missed metastatic disease ■ **38%** negative scans would be avoided	■ **22.6%** (prev **27%**) of patients would be scanned – Of those, **14.3%** (prev **10%**) would be positive ■ Estimated **0.4%** of patients have missed metastatic disease ■ **44%** negative scans would be avoided

MUSIC Imaging Appropriateness Criteria instituted across Michigan

- Statewide goal of performing imaging in $\geq \mathbf{95\%}$ of patients that meet the criteria and in $< \mathbf{10\%}$ of those that do not
- MUSIC members were provided with a toolkit including placards with the criteria and explanations for patients

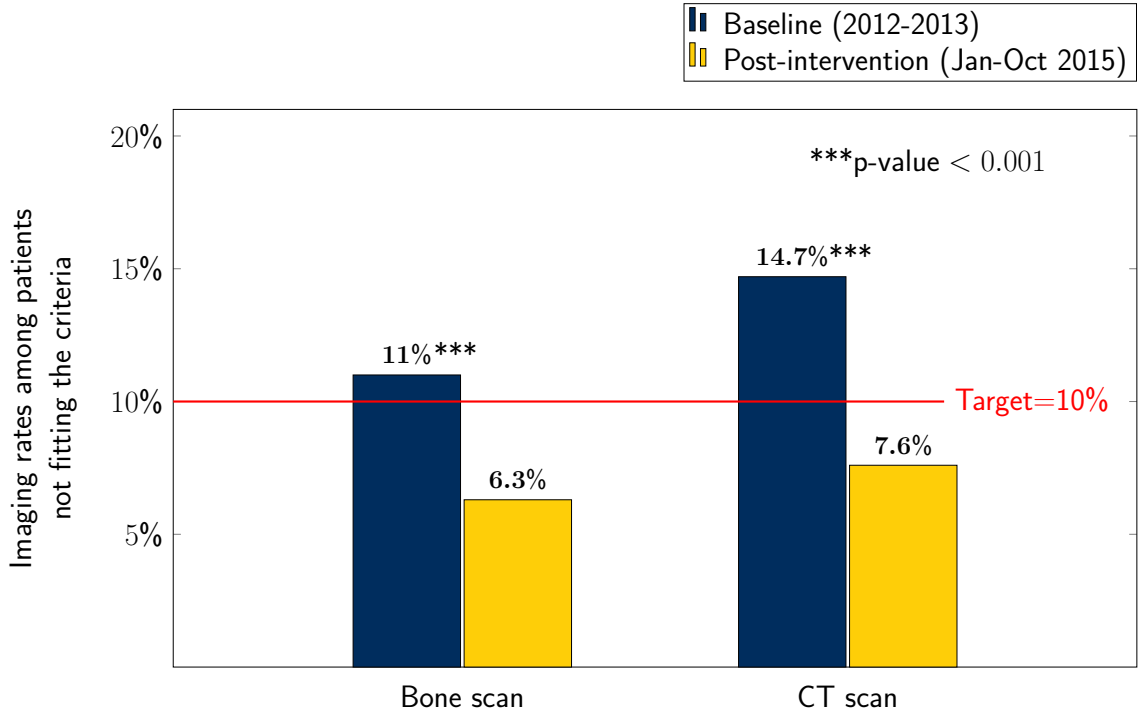
MUSIC Imaging Appropriateness Criteria

	Bone Scan	CT Scan
	Order Bone Scan If:	Order Bone Scan If:
PSA	> 20	> 20
	OR	*OR*
Gleason	≥ 8	≥ 8
		OR
Clinical T Stage		≥ cT3

Imaging Goals

- Perform imaging in ≥ 95% of patients meeting criteria
- Perform imaging in < 10% of patient NOT meeting criteria

MUSIC achieved state-wide improvement in the utilization of imaging tests



Hurley, P. et al. "A State-wide Intervention Improves Appropriate Imaging in Localized Prostate Cancer," *Urology*, 197(5):1222-1228, 2016.

Main Takeaway Messages

- New guidelines based on a combination of descriptive, predictive, and prescriptive analytics resulted in:
  - Sustained multi-year quality of life improvement among 5,000 patients/year diagnosed with prostate cancer in Michigan State
  - Cost savings of more than $262,000 annually in direct savings by reducing unnecessary imaging
  - Publications in the journal *Urology* and *Operations Research* make the approach available for other applications

References

- Selin Merdan, Christine Barnett, David C. Miller, James E. Montie, Brian T. Denton. Data Analytics for Optimal Detection of Metastatic Prostate Cancer, *Operations Research*, Published Online, 2021.
- Selin Merdan, Paul R. Womble, David C. Miller, Christine Barnett, Zhu Ye, Susan M. Linsell, James E. Montie, Brian T. Denton. Toward better use of bone scans among men with early-stage prostate cancer. *Urology*.
- Rachel Risko, Selin Merdan, Paul R. Womble, Christine Barnett, Zhu Ye, Susan M. Linsell, James E. Montie, David C. Miller, Brian T. Denton. Clinical predictors and recommendations for staging computed tomography scan among men with prostate cancer. *Urology*.

Thank you

Sensitivity and specificity of imaging tests

- Bone scan
  - Sensitivity of 86%
  - Specificity of 81%
- CT scan
  - Sensitivity of 42%
  - Specificity of 82%

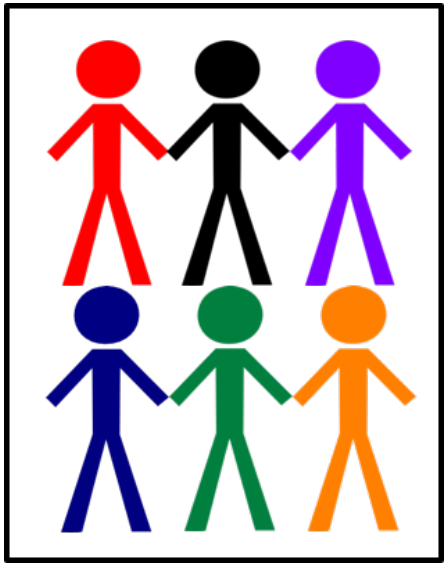
State-wide increase in the utilization of imaging tests for high-risk patients

- An increase in the use of imaging tests in patients that meet the criteria from
  - 82% to 84% for BS
  - 74% to 77% for CT scan
- The MUSIC consortium has made measurable improvements in a short period of time and additional increases are anticipated

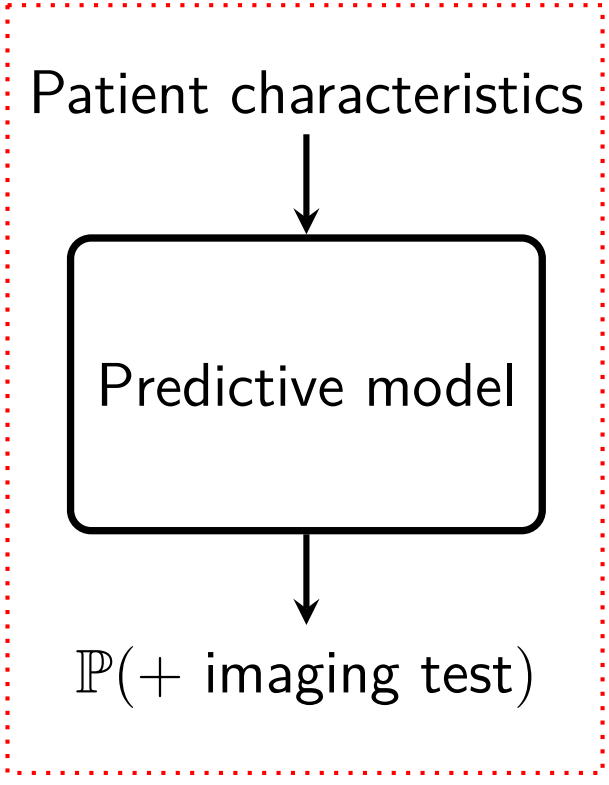
Compliance with MUSIC Imaging Criteria

- Our results were presented at collaborative-wide meetings with clinical champions who returned to their practices to present the results to their own practice group
- MUSIC members were provided with a toolkit including placards with the criteria and explanations for patients
- Members received comparative performance feedback that detailed how well their practice patterns correlated with the MUSIC Imaging Criteria

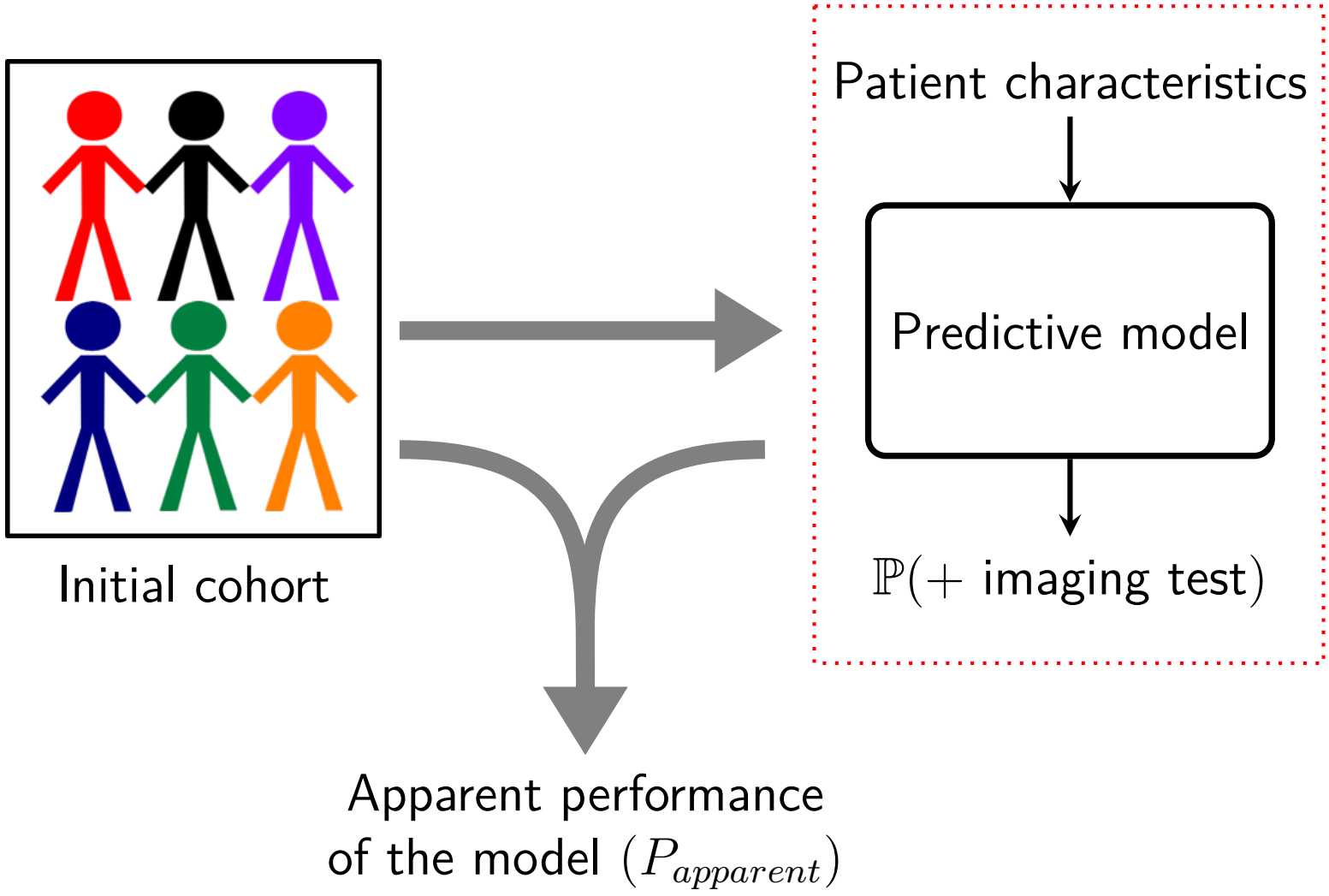
Optimism bias associated with internal validation

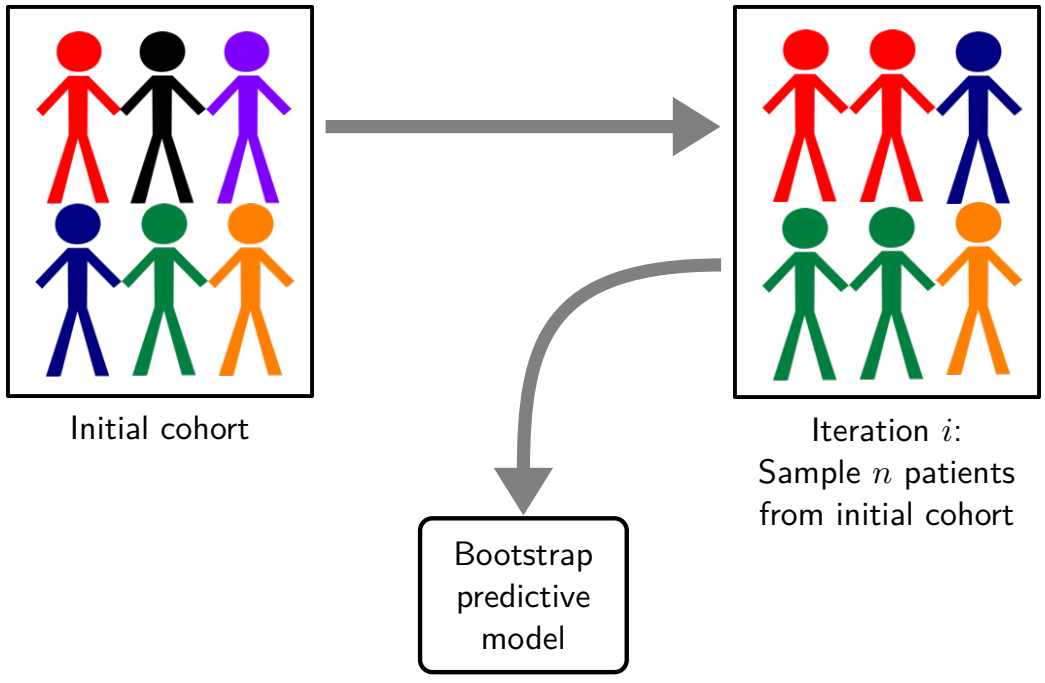


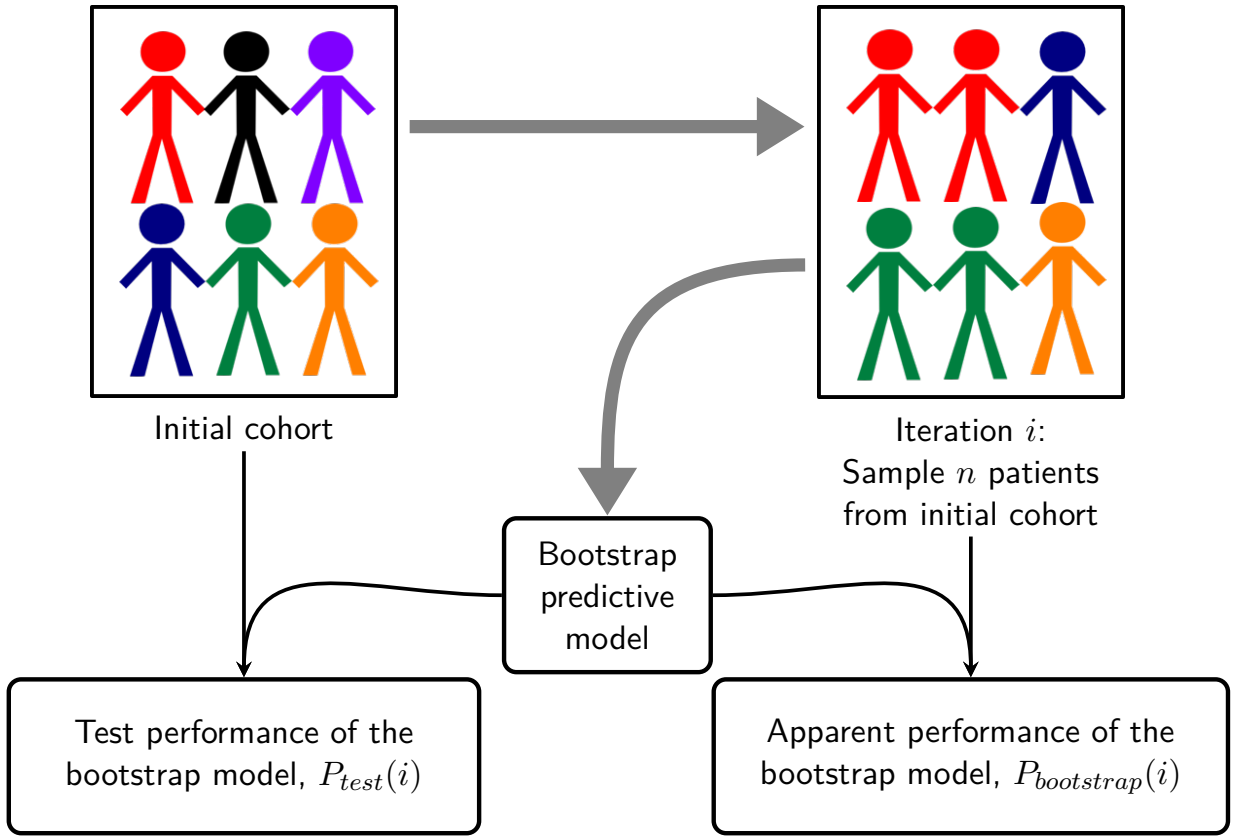
Initial cohort



Optimism bias associated with internal validation




Initial cohort
Iteration i:
Sample n patients
from initial cohort
Bootstrap
predictive
model



Estimate the optimism:

$$o(i) = P_{bootstrap}(i) - P_{test}(i)$$

Using bootstrapping to correct for optimism bias

- After m iterations of bootstrapping, we estimated the expected optimism:

$$\text{Optimism} = \frac{\sum_{i=1}^{m} o(i)}{m}$$

- This optimism estimate was used to update the apparent performance of our model:

$$P_{validated} = P_{apparent} - \text{Optimism}$$

Agreement between internal and external validation estimates

	Internal Validation		**External Validation**	
	Bone scan $(\mathbf{n = 416})$	**CT scan** $(\mathbf{n = 643})$	**Bone scan** $(\mathbf{n = 664})$	**CT scan** $(\mathbf{n = 507})$
ROC area	0.82	0.87	0.81	0.86
Brier score	0.080	0.060	0.068	0.061
Calibration slope	0.86	0.90	0.99	0.94

Performance measures were found by applying the predictive models fit in the development samples to the external validation samples.

Bias correction

$$\mathbb{P}(\mathrm{G}+ \mid \text{Disease present}) = \frac{\mathbb{P}(\text{Disease present} \mid \mathrm{G}+)\mathbb{P}(\mathrm{G}+)}{\mathbb{P}(\text{Disease present})}$$

$$\mathbb{P}(\mathrm{G}- \mid \text{Disease not present}) = \frac{\mathbb{P}(\text{Disease not present} \mid \mathrm{G}-)\mathbb{P}(\mathrm{G}-)}{\mathbb{P}(\text{Disease not present})}$$

Bias correction

$$\mathbb{P}(\text{G}+ \mid \text{Disease present}) = \frac{\mathbb{P}(\text{Disease present} \mid \text{G}+)\mathbb{P}(\text{G}+)}{\mathbb{P}(\text{Disease present})}$$

$$\mathbb{P}(\text{G}- \mid \text{Disease not present}) = \frac{\mathbb{P}(\text{Disease not present} \mid \text{G}-)\mathbb{P}(\text{G}-)}{\mathbb{P}(\text{Disease not present})}$$

- Separate the entire population into G+ and G−
- Develop logistic regression model among patients who received imaging to estimate the probability of metastatic disease for every patient

Bias correction

$$\mathbb{P}(\text{G}+ \mid \text{Disease present}) = \frac{\mathbb{P}(\text{Disease present} \mid \text{G}+)\mathbb{P}(\text{G}+)}{\mathbb{P}(\text{Disease present})}$$

$$\mathbb{P}(\text{G}- \mid \text{Disease not present}) = \frac{\mathbb{P}(\text{Disease not present} \mid \text{G}-)\mathbb{P}(\text{G}-)}{\mathbb{P}(\text{Disease not present})}$$

Bias correction

$$\mathbb{P}(\text{G}+ \mid \text{Disease present}) = \frac{\mathbb{P}(\text{Disease present} \mid \text{G}+)\mathbb{P}(\text{G}+)}{\mathbb{P}(\text{Disease present})}$$

$$\mathbb{P}(\text{G}- \mid \text{Disease not present}) = \frac{\mathbb{P}(\text{Disease not present} \mid \text{G}-)\mathbb{P}(\text{G}-)}{\mathbb{P}(\text{Disease not present})}$$

$$\begin{aligned}\mathbb{P}(\text{Disease present}) = {} & \mathbb{P}(\text{Disease present}|\text{G}+)\mathbb{P}(\text{G}+)+ \\ & \mathbb{P}(\text{Disease present}|\text{G}-)\mathbb{P}(\text{G}-)\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\text{Disease not present}) = {} & \mathbb{P}(\text{Disease not present}|\text{G}+)\mathbb{P}(\text{G}+)+ \\ & \mathbb{P}(\text{Disease not present}|\text{G}-)\mathbb{P}(\text{G}-)\end{aligned}$$

Algorithm 1: Cost-sensitive Laplacian Kernel Logistic Regression (Cos-LapKLR).

Input : l labeled examples $\{(\mathbf{x}_i, y_i)\}_{i=1}^{l}$, u unlabeled examples $\{\mathbf{x}_j\}_{j=l+1}^{l+u}$.

Output: Estimated function $f : \mathbb{R}^{(l+u)} \to \mathbb{R}$.

Step 1: Construct the data adjacency graph with $(l+u)$ nodes and compute the edge weights w_{ij} by k nearest neighbors.

Step 2: Choose a kernel function and compute the kernel matrix $\mathbf{K} \in \mathbb{R}^{(l+u)\times(l+u)}$.

Step 3: Compute the graph Laplacian matrix: $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D} = \mathrm{diag}\,(d_1, \ldots, d_{l+u})$ and $d_i = \sum_{j=1}^{l+u} w_{ij}$.

Step 4: Choose the regularization parameters $\gamma_{\mathcal{H}}$, $\gamma_{\mathcal{M}}$, and the cost parameter δ.

Step 5: Compute $\boldsymbol{\alpha}^*$ using the LM-BFGS algorithm.

Step 6: Output function $f^*(\mathbf{x}) = \sum_{i=1}^{l+u} \alpha_i^* \mathbf{K}(\mathbf{x}_i, \mathbf{x})$.

Proportions of non-dominated classification modeling techniques

Statistical models	**Bone scan** $(\mathbf{n = 40})$	**CT scan** $(\mathbf{n = 42})$
Cos-LapKLR	7.50	30.95
Cos-LR	47.50	0.00
Cos-SVM	27.50	40.48
RF	17.50	9.52
AdaBoost	0.00	19.05

The numbers are the percentages. Cos-SVM: Cost-sensitive support vector machines; Cos-LR: Cost-sensitive logistic regression, RF: Random forests