HEOR: Advancing Evidence to Action
Virtual ISPOR 2020
May 18-20, 2020 | #ISPORAnnual

Breakout Session 6
Modeling & Simulation Studies
Tuesday, May 19, 2020, 2:30 PM - 3:30 PM

Predicting Optimal Treatment Regimens for HR+/HER2- Breast Cancer Based on Electronic Health Records Using Random Forest

Zhanglin Lin Cui, Zbigniew Kadziola, Douglas E Faries,
Ilya Lipkovich, Bohdana Ratitch, Xiaohong Ivy Li,
Kristin M Sheffield, Gebra Cuyun Carter

Eli Lilly and Company, Indianapolis, IN 46295, USA

Conflict of Interest Statement

This study was sponsored by Eli Lilly and Company.

The presenter and coauthors are employees of Eli Lilly and Company.

Introduction

- **Breast cancer** is the **most common cancer** among women, estimated 276,480 new cases of invasive disease and 42,170 deaths in 2020 in the US.
- Breast cancer is a **heterogeneous disease** with several molecular/clinical subtypes based on factors such as hormone receptor (HR) status and human epidermal growth factor receptor 2 (HER2) protein expression.
- **HR+/HER2-** breast cancer accounts for over **70%** of all breast cancers.
- Assessing the cancer's HR status and HER2 status provides information to determine prognosis and predict the response to targeted therapies.
- **Endocrine**-based therapy, **Chemotherapy** and **CDK4&6 inhibitor**-based therapy have been the mainstay for HR+/HER2- metastatic breast cancer (mBC).

Hypothesis and Research Question

- Oncology physicians prescribe treatment based on patients' characteristics, cancer molecular/clinical subtypes, FDA-approved drug labels, treatment guidelines (e.g., NCCN guidelines), patient preferences, and physician experiences.
- The optimal treatment for a given patient (depending on patient and disease characteristics) to achieve the best survival outcome is not always well understood.
- The research question is: Can the optimal treatment for each patient, defined as improved outcomes be reliably estimated using machine learning algorithms?

Primary Objective

- Use random forest (RF) to **predict the treatment resulting in longest overall survival or time to treatment discontinuation (optimal treatment)** for patients initiating first or second line of therapy for HR+/HER2- metastatic breast cancer (mBC) to build understanding of how machine learning may help inform clinical decision-making.

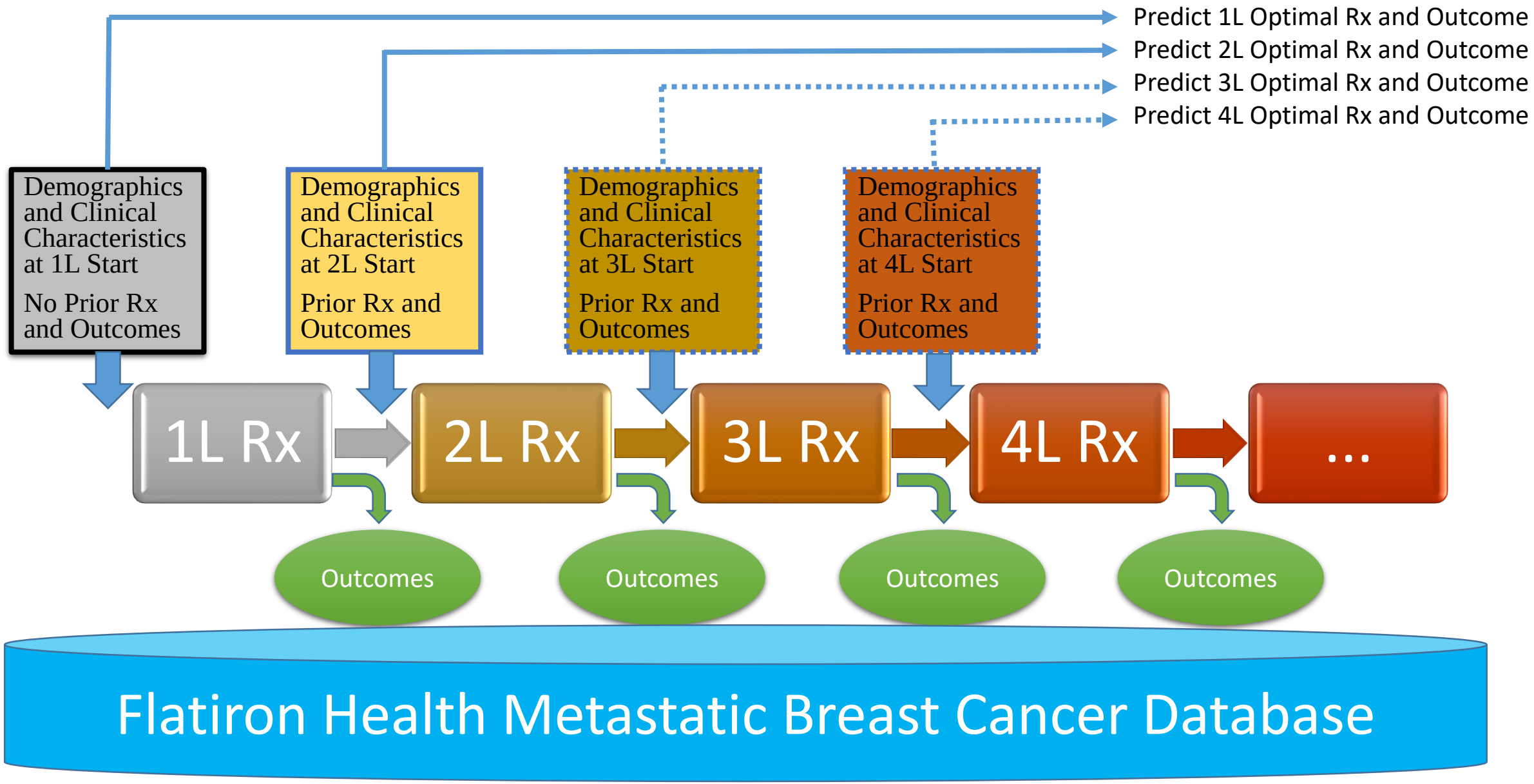
1. Compare ‘Optimal Treatment Patterns’ to ‘Observed Treatment Patterns’
2. Estimate the improvement (or gain) in outcomes from using the Optimal Treatment assignments for each line of therapy separately

Data We Used

- **Flatiron Health Metastatic Breast Cancer (mBC) Database**
  - Adult female patients **diagnosed with a mBC** between 1Jan2011 and 31Jan2019
  - **Received at least one line of systemic therapy** for mBC between 15Feb2015 and 31Jan2019
  - Regimen started between 15Feb2015 and **1 year prior to the end of follow up**
  - Exclude patients who had >90 days period between mBC diagnosis date and start of first-line therapy, as those may have received other therapies outside of the Flatiron network previously
- **Regimen classes**

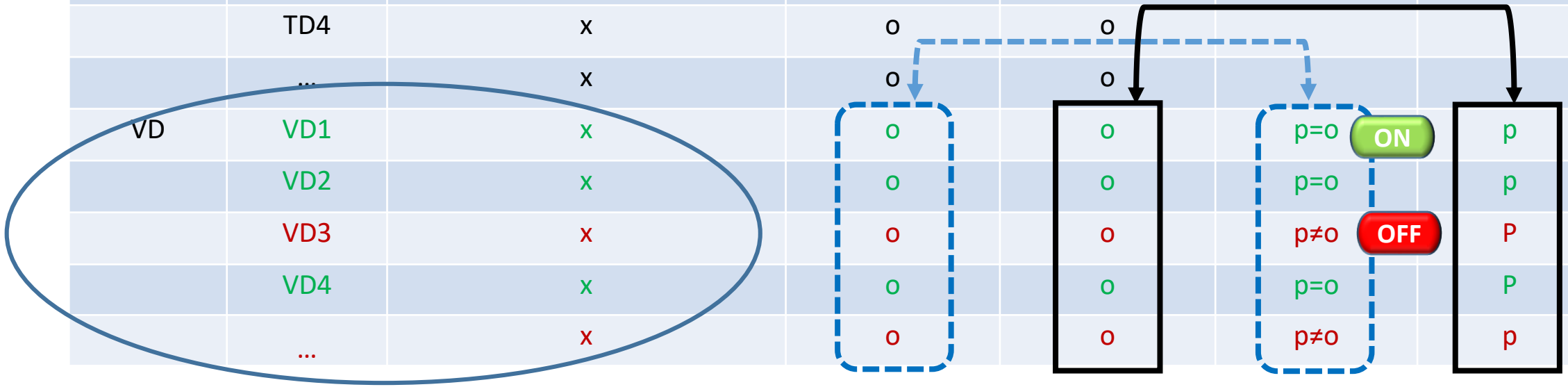
  Individual regimens were grouped into **hierarchy regimen classes** with top three included in this analysis:
  - **CDK4&6** inhibitor-based therapy,
  - **Endo**crine-based therapy, and
  - **Chemo**therapy alone.
- **Outcomes**
  - Overall survival (OS)
  - Time to treatment discontinuation (TTD)
- **Line of Therapy**
  - First line (1L)
  - Second line (2L)

Study Design



Predicting Optimal Treatment, Data View

Dataset	Patient ID	Baseline Covariates, Prior Rx and prior TTD	Observed Rx	Observed OS & TTD	Predicted Optimal Rx	Predicted OS & TTD under optimal Rx
TD	TD1	x	o	o		
	TD2	x	o	o		
	TD3	x	o	o		
	TD4	x	o	o		
	…	x	o	o		
VD	VD1	x	o	o	p=o	p
	VD2	x	o	o	p=o	p
	VD3	x	o	o	p≠o	P
	VD4	x	o	o	p=o	P
	…	x	o	o	p≠o	p

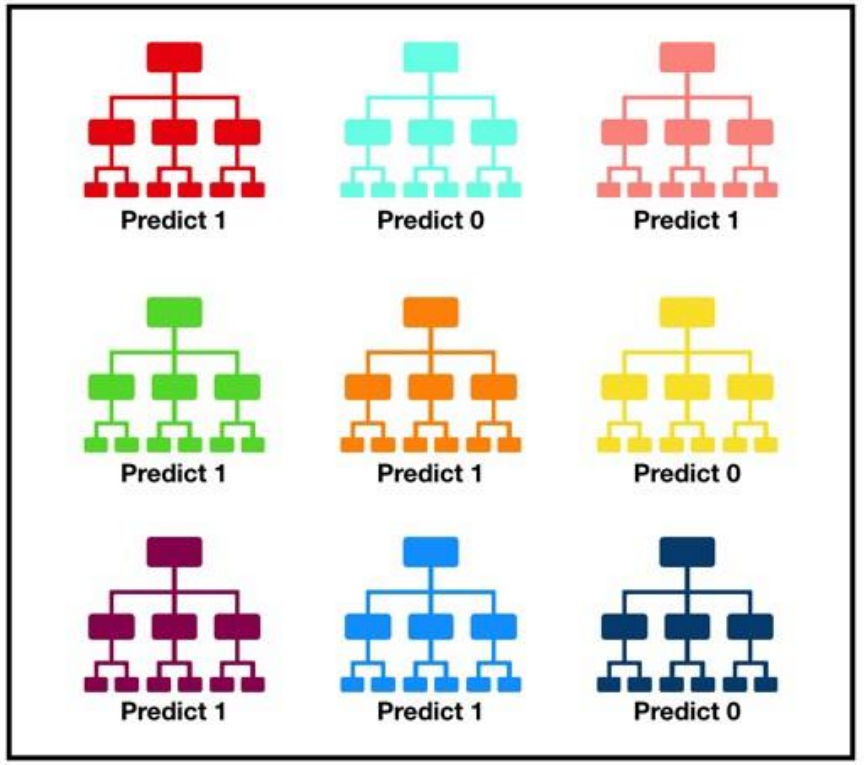


TD = training dataset; VD = validation dataset; Rx = treatment; OS = overall survival; TTD = time to treatment discontinuation; o = observed; p = predicted

Green=Patients On Optimal, Optimal Rx based on a machine-learning model is the same as Observed Rx.
Red=Patients Off Optimal, Optimal Rx based on a machine-learning model is different from Observed Rx.

Considerations for choosing the machine learning (ML) method for evaluating individualized treatment recommendation (ITR)

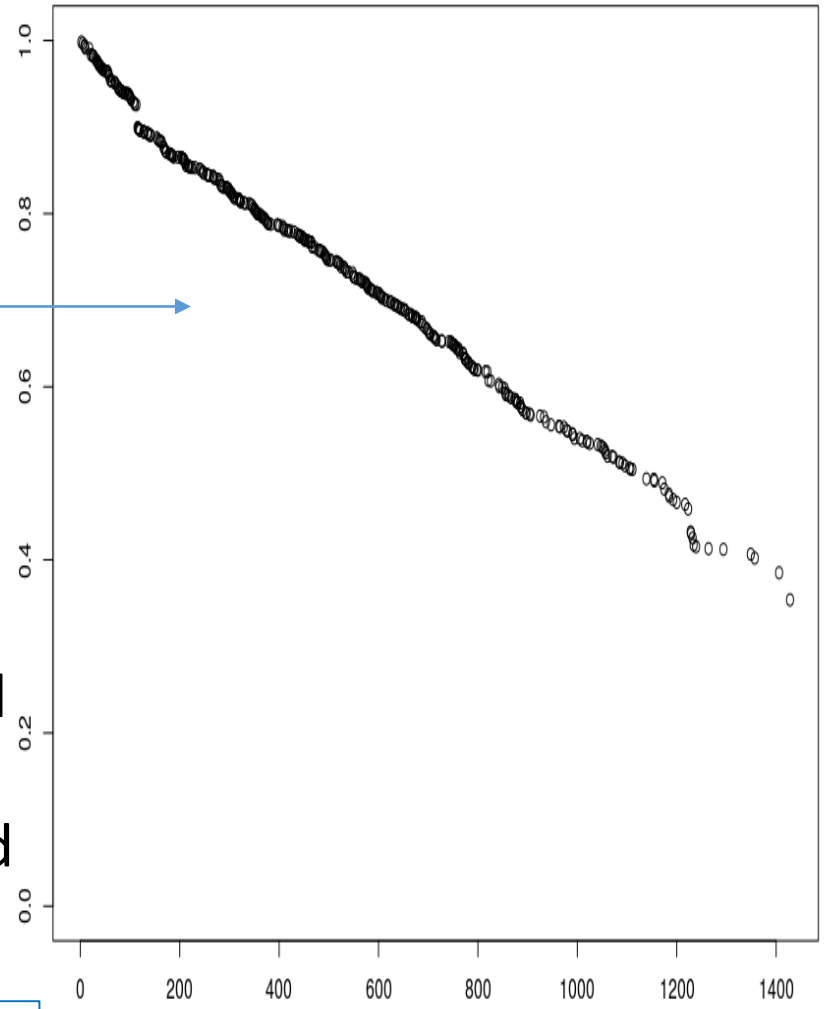
- **Limited number of machine learning methods** are available for survival outcomes with censoring.
- **Random forest (RF)** is a benchmark off-the-shelf and commonly-used method for predicting potential survival curves.
  - RF is already implemented as an R package, although also presents challenges in evaluating potential outcomes and optimal treatment which we successfully overcame.
- Direct ITR methods for survival are still under construction …
  - Outcome-weighted learning (OWL) based on Zhang et al. (2017) Multicategory Outcome Weighted Margin-based Learning for Estimating Individualized Treatment Rules.



RF consists of a large number of individual decision trees. Each tree predicts a treatment and the treatment with the most votes becomes the model's prediction. For survival outcomes, a forest of survival curves are generated to estimate the optimal Rx.

IPW weighted HR for On vs. Off optimal treatment and IPW weighted RMST (restricted mean survival time)

- Randomly split the data into:
  - 80% training dataset
    - Build the RF survival models: separately for each treatment
  - 20% validation dataset
    1) For each patient estimate her optimal treatment
       - Predict counterfactual survival curve for each treatment
       - Optimal treatment is defined as the treatment which gives the highest AUC
    2) Build the RF propensity score model for On vs. Off optimal treatment
    3) Use IPW to calculate weighted HR for On vs. Off optimal treatment
    4) Use IPW to calculate weighted RMST (5 & 10 years) for On and Off optimal treatment
    5) Record wHR , wRMST5_On, wRMST5_Off, wRMST10_On, wRMST10_Off
- Repeat the random split 1000 times and calculate 2.5,50, and 97.5 percentile for wHR, wRMST5, and wRMST10.
- Conduct the above for the given line of therapy (L1 or L2) and given outcome (OS or TTD) separately



AUC = area under the survival curve. IPW = inverse probability weighting. HR = hazard ratio. RMST = restricted mean survival time

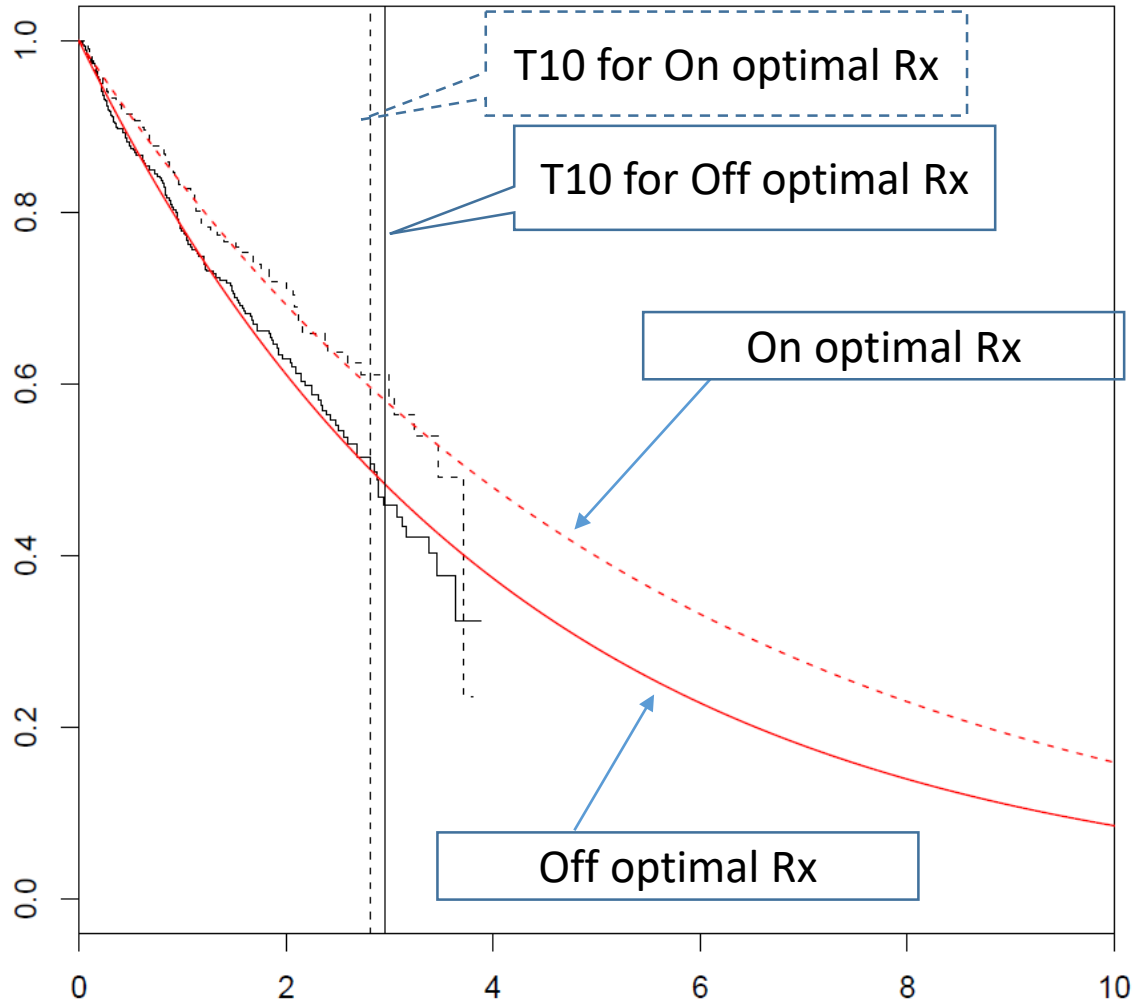
Calculation of weighted RMST on 20% data

- For patients On optimal treatment (dashed line)
  1) Find time T10 with 10% patients at risk
  2) Find the best (minimum AIC) parametric fit to M-M curve: consider exponential, Weibull, Gompertz, loglogistic, lognormal, generalized gamma distribution
  3) Calculate wRMST as

     AUC for observed survival for time from 0 to t10

     plus

     AUC for parametric curve for time from t10 to 5 (or 10) years
- Similarly for patients Off optimal treatment (solid line)



AIC = Akaike's Information Criterion.
K-M curve = Kaplan-Meier curve.
wRMST = inverse probability weighted restricted mean survival time.

Results from Flatiron data analysis: Demographic summaries by regimen classes

1. Good sample sizes with regimen classes for machine learning although small sample sizes with individual regimen

LOT	Outcome	CDK4&6	Chemo	Endo	Total
1L	OS	1674	791	1500	**3965**
1L	TTD	1666	705	1420	3791
2L	OS	993	678	784	**2455**
2L	TTD	981	614	734	2329

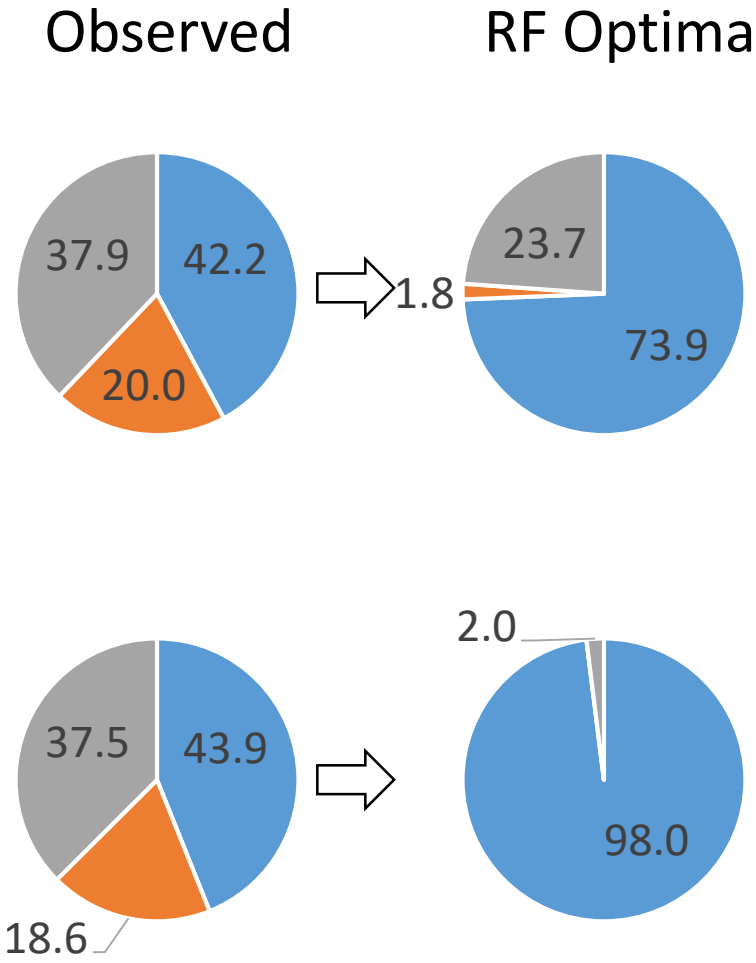
LOT = line of therapy. OS = overall survival. TTD = time to treatment discontinuation. ER = estrogen receptor. PR = progesterone receptor. BRCA = breast cancer gene name. S = South. NE = Northeast. MidW = Midwest. W = West. ECOG = Eastern Cooperative Oncology Group.

2. Differences exist in Demographics and Clinical Characteristics among CDK4&6, endocrine and chemo cohort

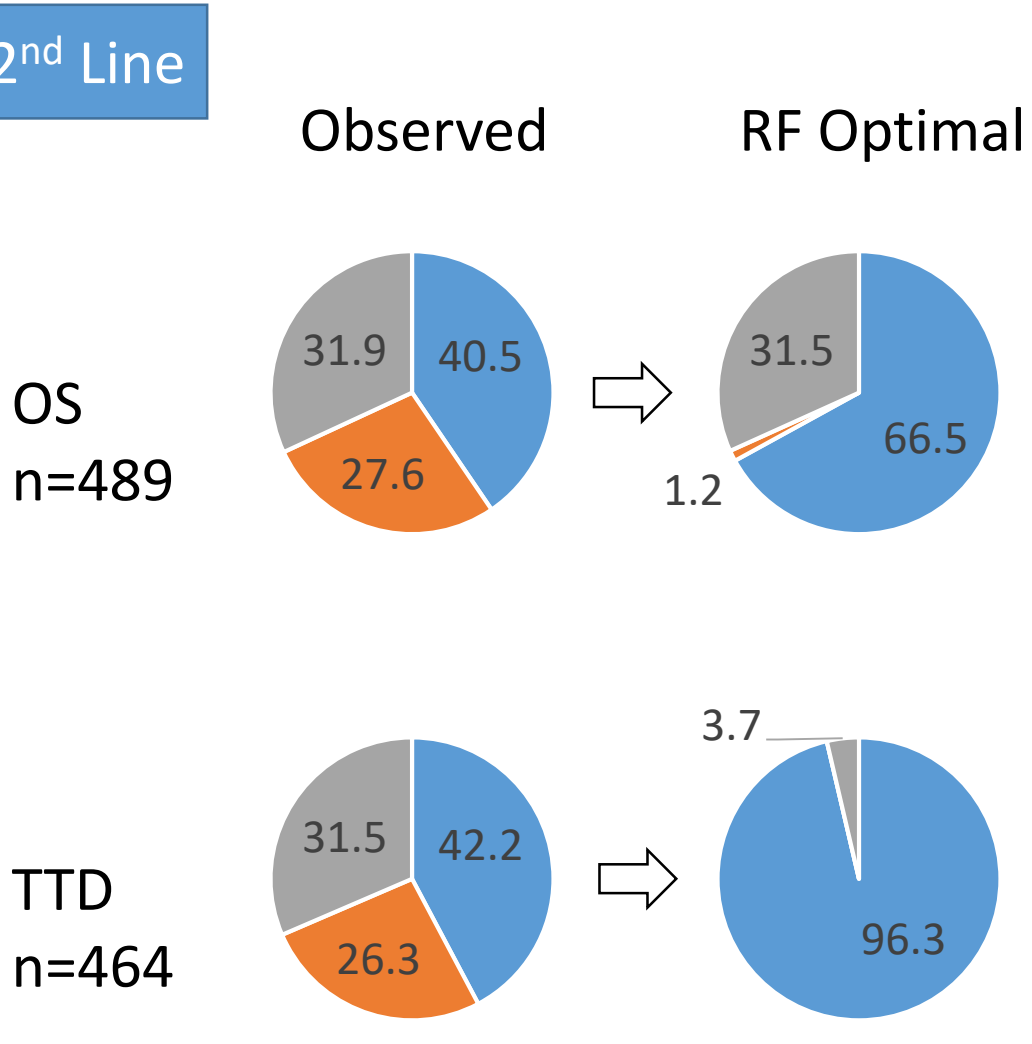
Baseline variable	P-value comparing regimen class cohort at 1st line start	P-value comparing regimen class cohort at 2nd line start
1) Age	<0.01	<0.01
2) Body weight	0.307	0.598
3) Time from metastasis diagnosis	<0.01	<0.01
4) ER status (-, +, unknown)	<0.01	<0.01
5) PR status (-, +, unknown)	<0.01	<0.01
6) BRCA status (-, +, unknown)	0.220	0.419
7) Practice region (S, NE, MidW, W, unknown)	<0.01	0.068
8) Practice type (community, academic)	<0.00	0.772
9) Race (White, Black, Asian, Other Race)	0.178	0.186
10) Group Stage (I, II, III, IV, not documented)	<0.01	0.169
11) ECOG Performance Stage (0, 1, 2, 3, 4, unknown)	<0.01	0.075
12) Prior TTD	N/A	<0.01
13) Prior regimen class	N/A	<0.01

Summary of Optimal Treatment Recommendation for HR+/HER2- Breast Cancer using Random Forest (RF)

1st Line



2nd Line



CDK4&6 Chemo Endocrine

Gains in Observed Survival Outcomes for Patients On vs Off Optimal Treatment for HR+/HER2- Breast Cancer using RF

The reported summaries are medians of predictive distribution computed from 1000 random samples of 20% of the original data

Line of Therapy	Adjustment	Outcome	N	% On Optimal	Off Optimal, RMST, Years	On Optimal, RMST, Years	HR [On vs Off] (95% PI)
1L	Unadjusted	TTD	758	44.3	0.88	2.17	0.46 (0.39, 0.54)
1L	Unadjusted	OS	792	39.1	3.65	4.36	0.79 (0.62, 1.03)
2L	Unadjusted	TTD	464	42.2	0.63	1.53	0.49 (0.41, 0.60)
2L	Unadjusted	OS	489	38.7	2.97	4.12	0.61 (0.45, 0.83)
1L	IPW	TTD	758		0.88	2.12	0.47 (0.39, 0.55)
1L	IPW	OS	792		3.58	4.21	0.81 (0.64, 1.04)
2L	IPW	TTD	464		0.64	1.48	0.50 (0.41, 0.60)
2L	IPW	OS	489		3.04	4.13	0.62 (0.46, 0.85)

- Less than 50% HR+/HER2- patients received optimal treatment as estimated by RF.
- The benefits of RF-predicted optimal treatment, in terms of overall survival (OS) and time to treatment discontinuation (TTD), are observed across lines of therapy (LOT) in HR+/HER2- subtype.
- IPW adjustment predicts slightly more conservative survival benefits than unadjusted results.

RMST = restricted mean survival time, over a 5-year horizon for TTD and 10-year horizon for OS. HR = hazard ratio. IPW = inverse probability weighting.

Limitations

- **Sample size** was adequate to perform the analysis, but larger samples would have improved the results.
- Key variables such as site of metastases were not available in Flatiron, leading to future sensitivity analysis for **unmeasured confounders**.
  - This violates the assumptions of appropriately using RF and the inclusion of factors such as comorbidities may impact the optimal treatment estimation.
- **Regimen classes** are used currently; future analyses should use **individual regimens** as appropriate sample size becomes available.
- Optimal treatment was estimated for each line of therapy separately. **Treatment sequence** was not taken into account.
- **Other machine learning methods,** such as extreme gradient boosting (XGBoost), outcome-weighted learning (OWL), and penalized regression (PR) may be explored as sensitivity analysis.
- **External validation** was not performed.

Conclusion

- RF was feasible using oncology EHR data.
- Machine learning may be useful to inform physicians of optimal treatment that improve outcomes for patients with HR+/HER2- mBC.
- We estimated that the optimal treatments as identified by the RF algorithm was superior to usual care for HR+/HER2- mBC in 1L and 2L settings.

References

1. Siegel RL, KD Miller, A Jemal. Cancer Statistics, 2020. CA Cancer J Clin 2020; 70:7-30
2. Senkus E, et al. *Ann Oncol* 2015;26 Suppl 5:v8-30.
3. Cardoso F, et al. *Breast* 2018;39:131-8.
4. O'Shaughnessy J. *Oncologist* 2005;10:20-9.
5. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 2015;386:1341-52.
6. Richman J & Dowsett M. *Nat Rev Clin Oncol* 2019;16:296-311.
7. National Cancer Institute. SEER Program. https://seer.cancer.gov/about/. Accessed 11 Nov 2019.
8. Cianfrocca M & Goldstein LJ. *Oncologist* 2004;9:606-16.
9. Daly B & Olopade OI. *JAMA* 2015;313:141-2.
10. Breast Cancer Treatment (PDQ®)—Health Professional Version - National Cancer Institute. General Information about Breast Cancer. https://www.cancer.gov/types/breast/hp/breast-treatment-pdq#section/_1.
11. Santa-Maria CA, Gradishar WJ. Changing Treatment Paradigms in Metastatic Breast Cancer: Lessons Learned. JAMA Oncol; Jul, 2015.
12. Reinert T, Barrios CH. Optimal management of hormone receptor positive metastatic breast cancer in 2016. Ther Adv Med Oncol. 2015;7(6):304-320.
13. Chong Zhang, Jingxiang Chen, Haoda Fu, Xuanyao He, Ying-Qi Zhao, and Yufeng Liu. 2017. Multicategory Outcome Weighted Margin-based Learning for Estimating Individualized Treatment Rules Statistica Sinica. doi:10.5705/ss.202017.0527
14. Advanced Breast Cancer Community. http://www.advancedbreastcancercommunity.org/understanding-abc

Presenter Contact Information

Zhanglin Lin Cui

Eli Lilly and Company, Indianapolis, IN 46285

cui_zhanglin@lilly.com

317-363-9155