Causal Interaction Trees: Finding Subgroups with Heterogeneous Treatment Effects in Observational Data

Jon Steingrimsson

Department of Biostatistics
Brown University

CFAR Symposium on Statistics and Data Science

Collaborators

- **Jiabei Yang**
- Joseph W. Hogan
- Ann W. Mwangi
- Rami Kantor
- Issa J. Dahabreh
- Allison Delong
- Monicah Nyambura

Introduction

- Concept of heterogeneous treatment effects is at the centre of comparative effectiveness research and precision medicine
  - ⋆ Treatment effects vary across patients
  - ⋆ Aim to determine which treatment works best, for whom and under what circumstances

Dahabreh et al. (2016), Wager and Athey (2018)

Introduction

- Heterogeneous treatment effect estimation usually done using pre-specified subgroup estimation:
  - ⋆ People belong to multiple subgroups.
  - ⋆ Type I and type II errors.

Dahabreh et al. (2016), Lagakos (2006)

Introduction

- Heterogeneous treatment effect estimation usually done using pre-specified subgroup estimation:
  - ⋆ People belong to multiple subgroups.
  - ⋆ Type I and type II errors.
- Develop Data-driven tree-based algorithms for identification of subgroups with differential treatment effects.

Dahabreh et al. (2016), Lagakos (2006)

Motivating problem

Identify HIV patients that are more or less likely to gain weight while on dolutegravir based ARTs using data from electronic health records on HIV patients in Western Kenya.

- **Some statistical challanges:**
  - repeated measures over time, not temporally aligned, number and timing of the measurements often being informative about outcomes
  - time-varying confounding
  - potentially informative dropout

Dahabreh et al. (2016, 2017)

Subset of related work

- Tree-based algorithms for subgroup identification in observational data include:
  - ⋆ Su et al. (2012) and Kang et al. (2014) proposed likelihood-based splitting statistic and AIC type pruning criteria in tree-based algorithms

Su et al. (2012), Kang et al. (2014), Athey and Imbens (2016)

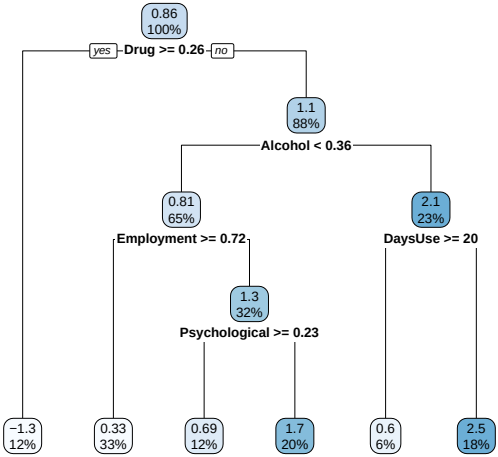
Subset of related work

- Tree-based algorithms for subgroup identification in observational data include:
  - ⋆ Su et al. (2012) and Kang et al. (2014) proposed likelihood-based splitting statistic and AIC type pruning criteria in tree-based algorithms
  - ⋆ Causal Tree algorithm of Athey and Imbens (2016) minimizes an estimator of the mean squared error of subgroup-specific treatment effect in splitting statistic

Su et al. (2012), Kang et al. (2014), Athey and Imbens (2016)

CART Algorithm for Outcome Prediction

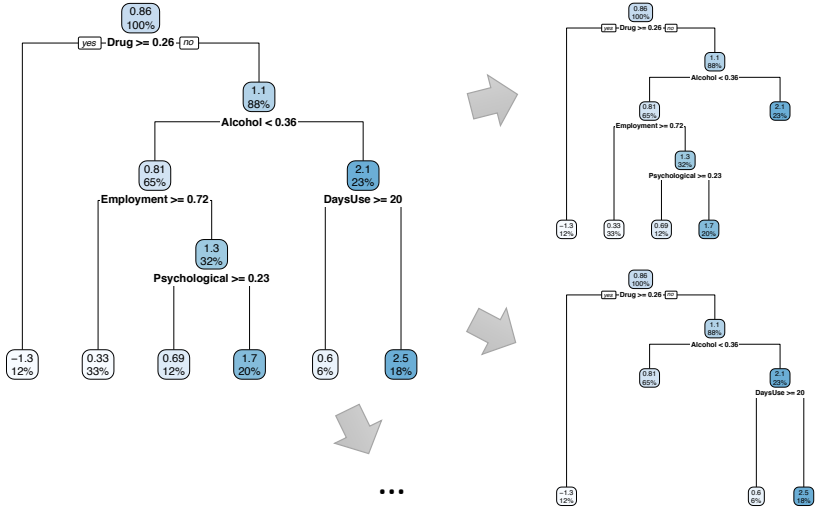
1. Grow a maximum sized tree until pre-determined stopping criterion is met



Steingrimsson and Yang (2019)

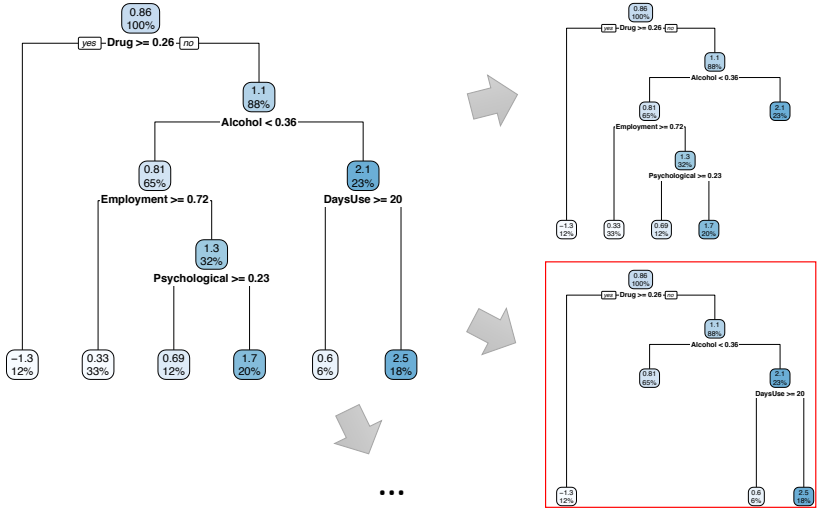
CART Algorithm for Outcome Prediction

2. Prune to create a set of candidate trees for final prediction



CART Algorithm for Outcome Prediction

3. Cross-validate to select final tree from candidate trees generated in pruning step



Generalized Interaction Trees (GIT): Splitting

CART: 1) Grow a maximum sized tree until pre-determined stopping criterion is met

- Grow a maximum sized tree using splitting statistic that maximizes standardized treatment effect differences between child nodes

$$\left(\frac{(\hat{\mu}_1(l) - \hat{\mu}_0(l)) - (\hat{\mu}_1(r) - \hat{\mu}_0(r))}{\sqrt{\widehat{\text{Var}}\left[(\hat{\mu}_1(l) - \hat{\mu}_0(l)) - (\hat{\mu}_1(r) - \hat{\mu}_0(r))\right]}} \right)^2$$

where

- $\hat{\mu}_a(w)$: Subgroup-specific treatment effect estimator
- l, r: left and right subgroups

Generalized Interaction Trees (GIT): Pruning

CART: 2) Prune to create a set of candidate trees for final prediction

- For a penalization parameter $\lambda \in \mathbb{R}_+$, define the **split complexity** for a tree ψ as

$$G^{(\lambda)}(\psi) = \underbrace{\sum_{i \in I_\psi} G_i(\psi)}_{\text{Sum of splitting statistics}} - \underbrace{\lambda |I_\psi|}_{\text{Size penalty}}$$

where

- I_ψ: The set of internal nodes of ψ
- $|I_\psi|$: Number of internal nodes
- $G_i(\psi)$: Splitting statistic for internal node i

Su et al. (2009)

Generalized Interaction Trees (GIT): Pruning

CART: 2) Prune to create a set of candidate trees for final prediction

- For a penalization parameter $\lambda \in \mathbb{R}_+$, define the **split complexity** for a tree ψ as

$$G^{(\lambda)}(\psi) = \underbrace{\sum_{i \in I_\psi} G_i(\psi)}_{\text{Sum of splitting statistics}} - \underbrace{\lambda |I_\psi|}_{\text{Size penalty}}$$

  where

  - I_ψ: The set of internal nodes of ψ
  - $|I_\psi|$: Number of internal nodes
  - $G_i(\psi)$: Splitting statistic for internal node i

- Varying λ creates a sequence of candidate trees that maximize the split complexity for different values of λ
  - $\lambda = 0$: maximum sized tree
  - $\lambda \to \infty$: root-only tree

Su et al. (2009)

Generalized Interaction Trees (GIT): Final Tree Selection

CART: 3) Select final tree from candidate trees generated in pruning step

- **Complication:** True treatment effects are not observed on individual participants

Generalized Interaction Trees (GIT): Final Tree Selection

CART: 3) Select final tree from candidate trees generated in pruning step

- **Complication:** True treatment effects are not observed on individual participants
- Loosely speaking, select the final tree that **maximizes the cross-validated split complexity** for a fixed penalization parameter λ
  - Select λ at quantile of asymptotic distribution of splitting statistic

Inference

- Likely overestimation of treatment effect differences.
- Data splitting: Fit tree on one part of data and estimate treatment effects using the other part.

Back to EHR data: Data structure

Assume we have longitudinal data

$$\boldsymbol{O} = (\boldsymbol{L}_0, A_0, C_0, \boldsymbol{L}_1, A_1, C_1, \ldots, \boldsymbol{L}_K, A_K, C_K, Y_{K+1}).$$

Where

- $\boldsymbol{L}_k$ are covariates
- C_k censoring indicator
- A_k treatment indicator
- Y_{K+1} outcome

Back to EHR data: Target parameter and estimator

Want to estimate a contrast between always on DTG-containing ART vs never on DTG-containing ART if censoring abolished.

$$\delta(w) = \mathrm{E}[Y_{K+1}^{\mathbf{1},\boldsymbol{c}=\mathbf{0}} | \boldsymbol{L}_0 \in w] - \mathrm{E}[Y_{K+1}^{\mathbf{0},\boldsymbol{c}=\mathbf{0}} | \boldsymbol{L}_0 \in w]$$

We propose to use longitutional TMLE to estimate $\delta(w)$

Subgroup Discovery for Longitudtional Data Algorithm

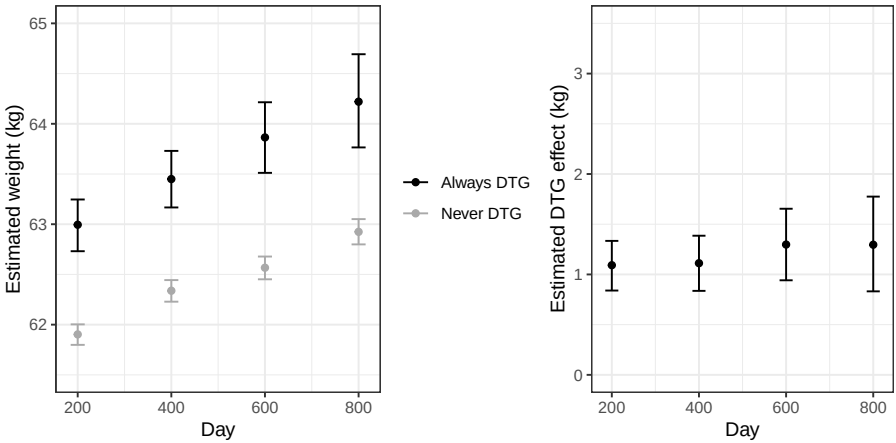
Use longitutional TMLE in combintion with the causal interaction tree algorithm for subgroup discovery with longitudional data.

Analysis of EHRs

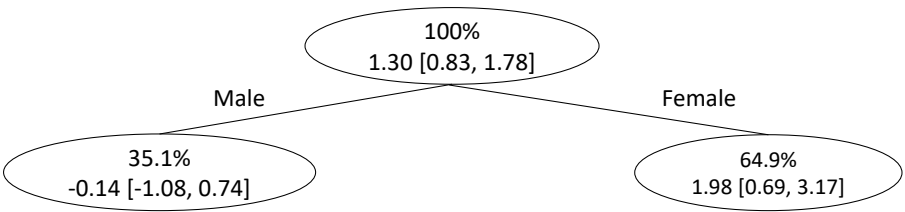
Data from Academic Model Providing Access to Healthcare that focuses on prevention and treatment of HIV in western Kenya. Data split into 200 day intervals.

- Time 0 defined as maximum time of first database entry or July 1, 2016.
- Outcome (Y_{K+1}): weight (kg).
- Treatment (A_k): Whether on DTG-containing ART
- Ten time-dependent covariates and five additional baseline covariates.
- Sample size $62,073$, always on DTG 424, never on DTG $30,307$.

Estimated average weight gain



Final tree



References

- Yang, Jiabei, Ann W. Mwangi, Rami Kantor, Issa J. Dahabreh, Monicah Nyambura, Allison Delong, Joseph W. Hogan, and Jon A. Steingrimsson. Tree-based Subgroup Discovery In Electronic Health Records: Heterogeneity of Treatment Effects for DTG-containing Therapies. arXiv preprint arXiv:2208.14329 (2023).
- Yang, J., Dahabreh, I. J., and Steingrimsson, J. A. (2022). Causal interaction trees: Finding subgroups with heterogeneous treatment effects in observational data. Biometrics, 78(2), 624-635.
- Steingrimsson, J. A., and Yang, J. (2019). Subgroup identification using covariate adjusted interaction trees. Statistics in medicine, 38(21), 3974-3984.

Funding

This work was supported in part by Patient-Centered Outcomes Research Institute (PCORI) awards ME-2019C3-17875, ME-2021C2-22365; National Library of Medicine (NLM) grant R01LM013616 and National Institute of Allergy and Infectious Diseases (NIAID) awards P30AI042853 and R01AI167694 and K24AI134359. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NLM, NIAID, PCORI, PCORI's Board of Governors, or PCORI's Methodology Committee.

NextGen scholars: Advertisement

Scholarship that fully funds participation in the masters program at Brown University that is restricted to graduated from HBCUs.

References I

Susan Athey and Guido Imbens.

Recursive partitioning for heterogeneous causal effects.
Proceedings of the National Academy of Sciences, 113(27):7353–7360, 2016.

Issa J Dahabreh, Rodney Hayward, and David M Kent.

Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence.
International journal of epidemiology, 45(6):2184–2193, 2016.

Issa J Dahabreh, Thomas A Trikalinos, David M Kent, and Christopher H Schmid.

Heterogeneity of treatment effects.
In *Methods in comparative effectiveness research*, pages 247–292. Chapman and Hall/CRC, 2017.

Joseph Kang, Xiaogang Su, Lei Liu, and Martha L Daviglus.

Causal inference of interaction effects with inverse propensity weighting, g-computation and tree-based standardization.
Statistical Analysis and Data Mining: The ASA Data Science Journal, 7(5):323–336, 2014.

Stephen W Lagakos et al.

The challenge of subgroup analyses-reporting without distorting.
New England Journal of Medicine, 354(16):1667, 2006.

Jared K Lunceford and Marie Davidian.

Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study.
Statistics in medicine, 23(19):2937–2960, 2004.

Andrew Nobel et al.

Histogram regression estimation using data-dependent partitions.
The Annals of Statistics, 24(3):1084–1105, 1996.

References II

Leonard A Stefanski and Dennis D Boos.

The calculus of m-estimation.
The American Statistician, 56(1):29–38, 2002.

Jon Arni Steingrimsson and Jiabei Yang.

Subgroup identification using covariate-adjusted interaction trees.
Statistics in medicine, 2019.

Xiaogang Su, Chih-Ling Tsai, Hansheng Wang, David M Nickerson, and Bogong Li.

Subgroup analysis via recursive partitioning.
Journal of Machine Learning Research, 10(Feb):141–158, 2009.

Xiaogang Su, Joseph Kang, Juanjuan Fan, Richard A Levine, and Xin Yan.

Facilitating score and causal inference trees for large observational studies.
Journal of Machine Learning Research, 13(Oct):2955–2994, 2012.

Aad W Van Der Vaart and Jon A Wellner.

Weak convergence.
In *Weak convergence and empirical processes*, pages 16–28. Springer, 1996.

Stefan Wager and Susan Athey.

Estimation and inference of heterogeneous treatment effects using random forests.
Journal of the American Statistical Association, 113(523):1228–1242, 2018.

Jiabei Yang, Issa J Dahabreh, and Jon A Steingrimsson.

Causal interaction trees: Finding subgroups with heterogeneous treatment effects in observational data.
Biometrics, 78(2):624–635, 2022.

References III

Jiabei Yang, Ann W Mwangi, Rami Kantor, Issa J Dahabreh, Monicah Nyambura, Allison Delong, Joseph W Hogan, and Jon A Steingrimsson.
Tree-based subgroup discovery in electronic health records: Heterogeneity of treatment effects for dtg-containing therapies.
arXiv preprint arXiv:2208.14329, 2022.

Questions

Thank you!

Subgroup Identification Using Tree-Based Algorithms

- Advantages of tree-based algorithms:
  - Data driven
  - Partition data directly into identifiable subgroups

Subgroup Identification Using Tree-Based Algorithms

- Advantages of tree-based algorithms:
  - Data driven
  - Partition data directly into identifiable subgroups
- Randomized controlled trials
  - Jon A. Steingrimsson and Jiabei Yang. “Subgroup identification using covariate-adjusted interaction trees.” Statistics in medicine 38, no. 21 (2019): 3974-3984.

Subgroup Identification Using Tree-Based Algorithms

- Advantages of tree-based algorithms:
  - Data driven
  - Partition data directly into identifiable subgroups
- Randomized controlled trials
  - Jon A. Steingrimsson and Jiabei Yang. “Subgroup identification using covariate-adjusted interaction trees.” Statistics in medicine 38, no. 21 (2019): 3974-3984.
- Observational studies
  - Treatment not randomized
    - Jiabei Yang, Issa J. Dahabreh, and Jon A. Steingrimsson. “Causal interaction trees: Finding subgroups with heterogeneous treatment effects in observational data.” Biometrics (2021).

Subgroup Identification Using Tree-Based Algorithms

- Advantages of tree-based algorithms:
  - Data driven
  - Partition data directly into identifiable subgroups
- Randomized controlled trials
  - Jon A. Steingrimsson and Jiabei Yang. “Subgroup identification using covariate-adjusted interaction trees.” Statistics in medicine 38, no. 21 (2019): 3974-3984.
- Observational studies
  - Treatment not randomized
    - Jiabei Yang, Issa J. Dahabreh, and Jon A. Steingrimsson. “Causal interaction trees: Finding subgroups with heterogeneous treatment effects in observational data.” Biometrics (2021).
- Longitudinal studies
  - Intermittent missing and dropout
  - Time-varying treatment
  - Large electronic health record database
    - Jiabei Yang, Ann W. Mwangi, Rami Kantor, et al. ”Tree-based Subgroup Discovery In Electronic Health Records: Heterogeneity of Treatment Effects for DTG-containing Therapies.” arXiv preprint arXiv:2208.14329 (2022).

Subgroup Identification Using Tree-Based Algorithms

- Advantages of tree-based algorithms:
  - Data driven
  - Partition data directly into identifiable subgroups
- Randomized controlled trials
  - Jon A. Steingrimsson and Jiabei Yang. “Subgroup identification using covariate-adjusted interaction trees.” Statistics in medicine 38, no. 21 (2019): 3974-3984.
- Observational studies
  - Treatment not randomized
    - Jiabei Yang, Issa J. Dahabreh, and Jon A. Steingrimsson. “Causal interaction trees: Finding subgroups with heterogeneous treatment effects in observational data.” Biometrics (2021).
- Longitudinal studies
  - Intermittent missing and dropout
  - Time-varying treatment
  - Large electronic health record database
    - Jiabei Yang, Ann W. Mwangi, Rami Kantor, et al. “Tree-based Subgroup Discovery In Electronic Health Records: Heterogeneity of Treatment Effects for DTG-containing Therapies.” arXiv preprint arXiv:2208.14329 (2022).

Identifiability of $\mu_a(w)$

- Consistency of potential outcomes: For binary A,

$$Y = Y^1 A + Y^0(1 - A)$$

- Mean exchangeability:

$$\mathrm{E}[Y^a | A = a, \boldsymbol{X} = \boldsymbol{x}] = \mathrm{E}[Y^a | \boldsymbol{X} = \boldsymbol{x}]$$

  for all covariate patterns $\boldsymbol{x} \in w$ that have a positive density.

- Positivity:

$$0 < \mathrm{P}(A = 1 | \boldsymbol{X} = \boldsymbol{x}) < 1$$

  for all covariate patterns $\boldsymbol{x} \in w$ that have a positive density.

Inference from CIT Algorithm

- Data-splitting for consistent terminal node estimator
  - First part of dataset for fitting CIT algorithm
  - Second part of dataset for treatment effect estimation
- Under certain conditions, data adaptive CITs are consistent

Fithian et al. (2014), Nobel et al. (1996)

Simulation

- 1000 observations in training, and 1000 in test data:
  - $\boldsymbol{X}$ 6-dimensional multivariate normal, $\mathrm{E}(X^{(j)}) = 0$, $\mathrm{Var}(X^{(j)}) = 1$, $\mathrm{cov}(X^{(j)}, X^{(k)}) = 0.3$ when $j \neq k$

Simulation

- 1000 observations in training, and 1000 in test data:
  - $\boldsymbol{X}$ 6-dimensional multivariate normal, $\mathrm{E}(X^{(j)}) = 0$, $\mathrm{Var}(X^{(j)}) = 1$, $\mathrm{cov}(X^{(j)}, X^{(k)}) = 0.3$ when $j \neq k$
  - A with the true probability of treatment

$$e(\boldsymbol{X}, \beta) = \text{expit}\left(0.6X^{(1)} - 0.6X^{(2)} + 0.6X^{(3)}\right)$$

Simulation

- 1000 observations in training, and 1000 in test data:
  - $\boldsymbol{X}$ 6-dimensional multivariate normal, $\mathrm{E}(X^{(j)}) = 0$, $\mathrm{Var}(X^{(j)}) = 1$, $\mathrm{cov}(X^{(j)}, X^{(k)}) = 0.3$ when $j \neq k$
  - A with the true probability of treatment

$$e(\boldsymbol{X}, \beta) = \text{expit}\left(0.6X^{(1)} - 0.6X^{(2)} + 0.6X^{(3)}\right)$$

  - Outcome with heterogeneous treatment effect:

$$Y = 2 + 2A + 2I(X^{(1)} < 0) + e^{X^{(2)}} + 3A \cdot I(X^{(4)} > 0) + (X^{(5)})^3 + \epsilon$$

  - Outcome with homogeneous treatment effect:

$$Y = 2 + 2A + 2I(X^{(1)} < 0) + e^{X^{(2)}} + 3I(X^{(4)} > 0) + (X^{(5)})^3 + \epsilon$$

  where $\epsilon \sim N(0, 1)$.

Implementation

- Causal Interaction Trees:
  - ⋆ Generalized Interaction Trees implemented with IPW, g-formula, DR estimators (IPW-CIT, g-CIT, DR-CIT)
  - ⋆ Use `rpart`'s ability to incorporate user written splitting functions

Athey and Imbens (2016)

Implementation

- Causal Interaction Trees:
  - ⋆ Generalized Interaction Trees implemented with IPW, g-formula, DR estimators (IPW-CIT, g-CIT, DR-CIT)
  - ⋆ Use `rpart`'s ability to incorporate user written splitting functions
- Evaluated performance against Causal Trees (`causalTree`):
  - ⋆ Original CT: default splitting rule and cross-validation method
  - ⋆ Optimized CT: oracle version with optimal splitting rule and cross-validation method

Athey and Imbens (2016)

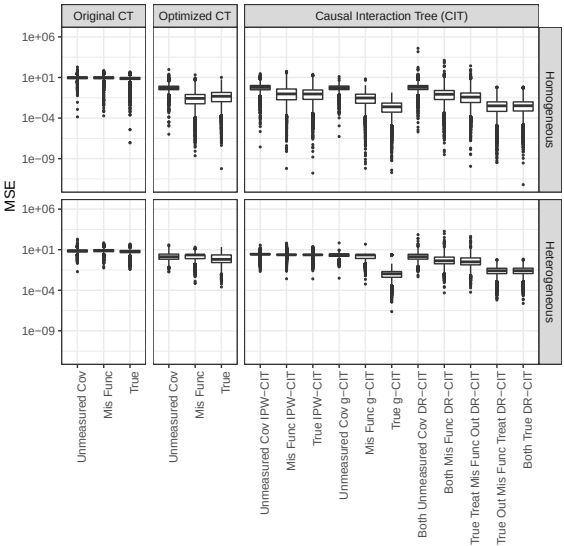
Implementation

- Terms in GLM for estimating nuisance parameters:

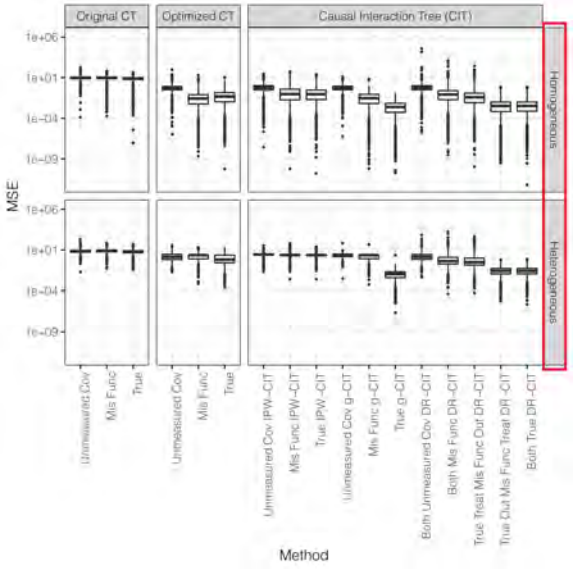
Model	Probability of treatment	Outcome
Correct model specification	$X^{(1)}, X^{(2)}, X^{(3)}$	A, $I(X^{(1)} < 0)$, $e^{X^{(2)}}$ $AI(X^{(4)} > 0)$ or $I(X^{(4)} > 0)$, $(X^{(5)})^3$
Functional form misspecification	$e^{X^{(1)}}, \ldots, e^{X^{(6)}}$	Main effects and all two-way interactions between $\boldsymbol{X}$ and A
Unmeasured covariate*	$X^{(1)}, X^{(3)}$ $X^{(4)}, X^{(5)}, X^{(6)}$	Main effects and all two-way interactions between $\boldsymbol{X}$ and A except for terms involving $X^{(2)}$

*: Remove $X^{(2)}$ from dataset when fitting CITs

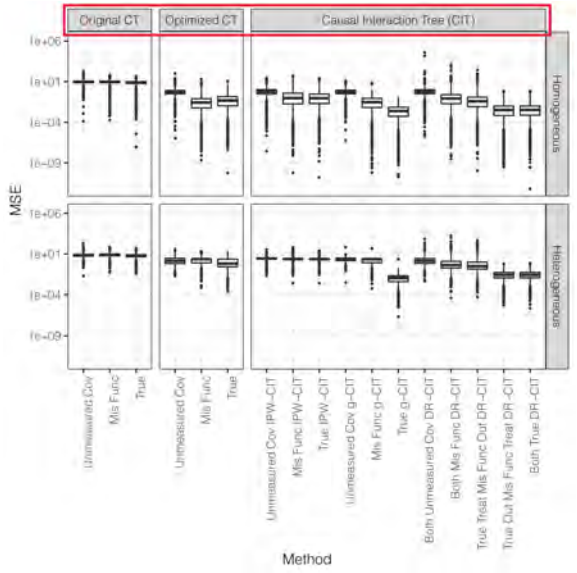
Simulation Result - Test Dataset MSE



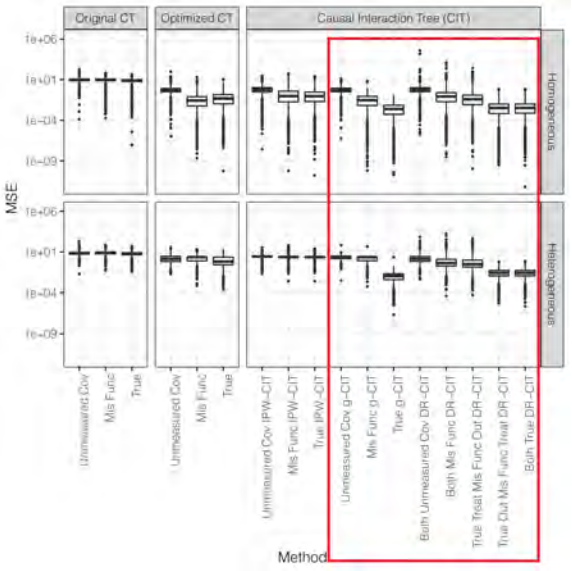
Simulation Result - Test Dataset MSE



Simulation Result - Test Dataset MSE

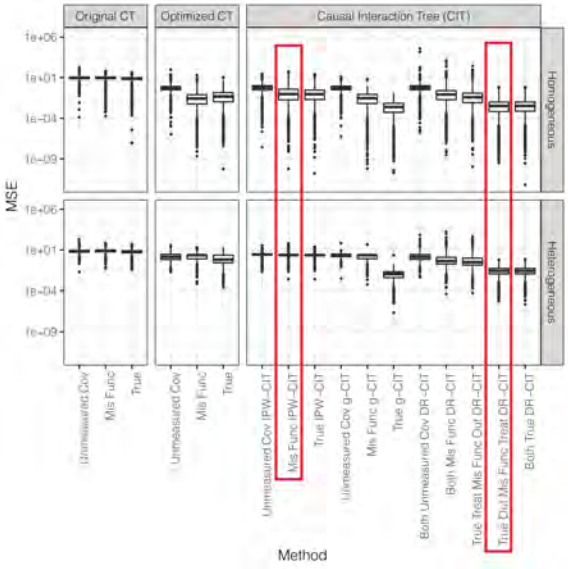


Simulation Result - Test Dataset MSE



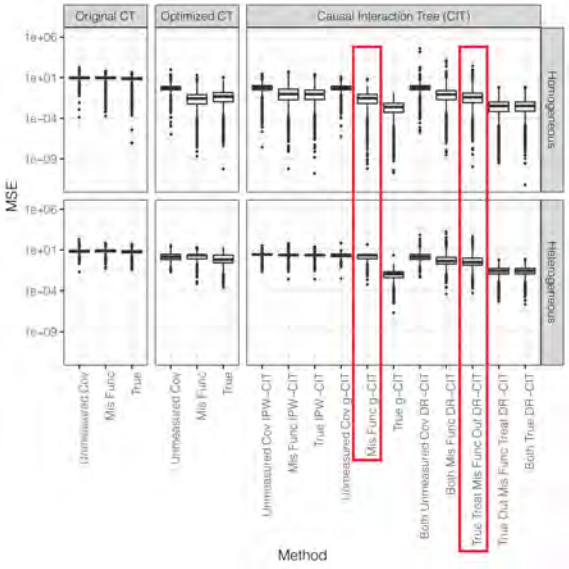
- g-CITs and DR-CITs perform substantially better than Causal Trees

Simulation Result - Test Dataset MSE



- g-CITs and DR-CITs perform substantially better than Causal Trees
- Robustness of DR-CITs compared with IPW-CITs and g-CITs

Simulation Result - Test Dataset MSE



- g-CITs and DR-CITs perform substantially better than Causal Trees
- Robustness of DR-CITs compared with IPW-CITs and g-CITs

Simulation Result

- Performance in terms of other measures also shows similar trend
  - Proportion of simulations identifying correct tree structure
  - Average number of noise variables
  - Proportion of simulations identifying correct first split in heterogeneous setting

Simulation Result

- Performance in terms of other measures also shows similar trend
  - Proportion of simulations identifying correct tree structure
  - Average number of noise variables
  - Proportion of simulations identifying correct first split in heterogeneous setting
- CTs are sensitive to chosen splitting rule and cross-validation method
  - Original CTs overfit by splitting on noise variables
  - Best CTs underfit by not making any splits

Simulation Result

- Performance in terms of other measures also shows similar trend
  - Proportion of simulations identifying correct tree structure
  - Average number of noise variables
  - Proportion of simulations identifying correct first split in heterogeneous setting

		Homogeneous Effect		Heterogeneous Effect		
Algorithm	Model	Correct Trees	Number Noise	Correct Trees	Number Noise	Correct First Split
Original CT	Unmeasured Cov	0.00	28.34	0.00	21.49	0.63
	Mis Func	0.00	25.28	0.00	19.87	0.30
	True	0.02	25.75	0.02	19.31	0.58
Optimized CT	Unmeasured Cov	0.99	0.03	0.52	0.15	0.87
	Mis Func	0.99	0.01	0.27	0.27	0.47
	True	0.99	0.01	0.59	0.19	0.84
DR-CIT	Both Unmeasured Cov	0.89	0.65	0.54	0.60	0.97
	Both Mis Func	0.90	0.56	0.67	0.65	0.99
	True Treat Mis Func Out	0.91	0.58	0.71	0.58	1.00
	True Out Mis Func Treat	0.95	0.14	0.93	0.14	1.00
	Both True	0.95	0.14	0.93	0.12	1.00

Simulation Result

- Performance in terms of other measures also shows similar trend
  - Proportion of simulations identifying correct tree structure
  - Average number of noise variables
  - Proportion of simulations identifying correct first split in heterogeneous setting

		Homogeneous Effect		Heterogeneous Effect		
Algorithm	Model	Correct Trees	Number Noise	Correct Trees	Number Noise	Correct First Split
Original CT	Unmeasured Cov	0.00	28.34	0.00	21.49	0.63
	Mis Func	0.00	25.28	0.00	19.87	0.30
	True	0.02	25.75	0.02	19.31	0.58
Optimized CT	Unmeasured Cov	0.99	0.03	0.52	0.15	0.87
	Mis Func	0.99	0.01	0.27	0.27	0.47
	True	0.99	0.01	0.59	0.19	0.84
DR-CIT	Both Unmeasured Cov	0.89	0.65	0.54	0.60	0.97
	Both Mis Func	0.90	0.56	0.67	0.65	0.99
	True Treat Mis Func Out	0.91	0.58	0.71	0.58	1.00
	True Out Mis Func Treat	0.95	0.14	0.93	0.14	1.00
	Both True	0.95	0.14	0.93	0.12	1.00

- Original CTs overfit by splitting on noise variables

Simulation Result

- Performance in terms of other measures also shows similar trend
  - Proportion of simulations identifying correct tree structure
  - Average number of noise variables
  - Proportion of simulations identifying correct first split in heterogeneous setting

		Homogeneous Effect		Heterogeneous Effect		
Algorithm	Model	Correct Trees	Number Noise	Correct Trees	Number Noise	Correct First Split
Original CT	Unmeasured Cov	0.00	28.34	0.00	21.49	0.63
	Mis Func	0.00	25.28	0.00	19.87	0.30
	True	0.02	25.75	0.02	19.31	0.58
Optimized CT	Unmeasured Cov	0.99	0.03	0.52	0.15	0.87
	Mis Func	0.99	0.01	0.27	0.27	0.47
	True	0.99	0.01	0.59	0.19	0.84
DR-CIT	Both Unmeasured Cov	0.89	0.65	0.54	0.60	0.97
	Both Mis Func	0.90	0.56	0.67	0.65	0.99
	True Treat Mis Func Out	0.91	0.58	0.71	0.58	1.00
	True Out Mis Func Treat	0.95	0.14	0.93	0.14	1.00
	Both True	0.95	0.14	0.93	0.12	1.00

- Original CTs overfit by splitting on noise variables
- Optimized CTs underfit by not making any splits

Analysis of SUPPORT Study - Result

- All CITs found no subgroups
  - Fully grown tree by DR-CIT first splits on if estimate of 2-month survival probability at study entry is greater than or equal to 0.85 or not
  - Data-splitting based treatment effect and confidence interval

P(Survival) ≥ 0.85	P(Survival) < 0.85
-0.007 (-0.-091, 0.081)	0.066 (0.022, 0.108)

  - Previous predetermined subgroup analysis on this covariate not significant
- CTs are sensitive to chosen splitting rule and cross-validation method
  - Original CT produced a tree with 232 terminal nodes
  - Optimized CT found no subgroups
  - Performance consistent with simulations where Original CT build large trees and Optimized CT build small trees
- Important feature of CITs: find no subgroups when there is no signal

Future Work - GIT Algorithm in Longitudinal Data

- Motivating dataset from AMPATH project
  - Collaborative project with the goal of providing HIV care in Kenya
  - Examine heterogeneous treatment effect of Dolutegravir (DTG) based therapy on weight gain and viral load response
  - 88,366 participants with 1,193,340 visits met eligibility criteria from database

Future Work - GIT Algorithm in Longitudinal Data

- Motivating dataset from AMPATH project
  - Collaborative project with the goal of providing HIV care in Kenya
  - Examine heterogeneous treatment effect of Dolutegravir (DTG) based therapy on weight gain and viral load response
  - 88,366 participants with 1,193,340 visits met eligibility criteria from database
- Estimate conditional expectation of potential outcome at end of study had participant received treatment pattern of interest and remained in study
  - Treatment pattern of interest: Always or never on DTG
  - Conditional on baseline covariates including demographic variables, lab measures, baseline clinic visit information

Analysis of EHRs

Covariates

- Time-dependent covariates: HIV viral load, systolic blood pressure, diastolic blood pressure, weight, height, whether the individual had active tuberculosis, was being treated for tuberculosis, was married or living with partner, and whether the individual was covered by the National Health Insurance Fund.
- Additional baseline covariates: Gender, age when starting ART, age at time 0, time on ART at time 0, whether the individual was enrolled on or after July 1, 2016 and whether the individual was newly initiating ART or had been on ART 200 days before time 0