統計数理(2025)
第 73 巻 第 2 号 335–352

特集「生存時間解析と信頼性解析」
[総合報告]

生存解析における樹木法と最近の発展

杉本 知之[1]・丸尾 和司[2]・下川 敏雄[3]

(受付 2024 年 12 月 2 日;改訂 2025 年 5 月 6 日;採択 5 月 14 日)

要　　旨

決定木やランダムフォレストなどの樹木法は,データサイエンス分野をはじめとして,他の実質科学領域やビジネスの実務で利用されている統計・機械学習ツールである.過去 10 年において,樹木法やフォレストからの変数重要度指標の推測,フォレストの一致性や漸近正規性に基づいた因果ツリーや因果フォレストなどの因果推論ツールとしての方法論や理論の発展も目覚ましい.本稿では,このような樹木法における方法や理論の最近の発展を,生存時間解析での利用の観点で各手法とその特徴を調査し総合的に報告する.

キーワード:CART,Cox 回帰,生存木,ランダムフォレスト,変数重要度,因果フォレスト.

1. はじめに

関心のあるイベントに対するリスク要因分析,将来起こるイベントからのリスク予測のため,生存解析に基づく様々な回帰モデリング手法が適用できる.統計的性質,解釈性,計算可能性といった多くの利便性の観点から,対数ハザードの観点で線形な Cox 回帰モデリングをデフォルトの選択肢として使用することも多い.樹木法は,より柔軟な仮定のもとで,高次の交互作用や変数の非線形要因に自動的に対応する分析手法として人気が高く,とくに,解釈性の高い決定木(CART)(Breiman et al., 1984; 下川 他, 2013),予測性能の高いランダムフォレスト(Breiman, 2001)は,科学研究での利用だけでなく,ビジネスにおける実務でも広く応用されるようになっている.そして,樹木法の構成方法の対案(例えば,Loh, 2002; Hothorn et al., 2006b),経時データや生存データへの方法の拡張(例えば,Segal, 1992; LeBlanc and Crowley, 1992; Arano et al., 2010; 江村, 2023)も数多く提案されてきた.生存解析におけるランダムフォレスト法(Ishwaran et al., 2008)は,高性能なリスク予測モデルを構築する Cox 回帰の代替手法を提供する(例えば,Morvan et al., 2020; Liao et al., 2024 など).また,このような生存フォレスト法に基づく変数重要度は,予後因子の同定のため,Cox 回帰のような全所的線形モデリングの代替もしくは補完として有用であることも多く報告されるようになっている(例えば,Hsich et al., 2011; He et al., 2020; Liu et al., 2021 など).

樹木法,ランダムフォレスト法は広く利用されているにもかかわらず,複数の要素を巧妙に組み合わせたブラックボックスの性質をもっているため,数理統計的な性質の解明の難度は高い.ランダムフォレストの一致性の早期の研究では,Breiman (2004)はランダムフォレストの

[1] 大阪大学 基礎工学研究科:〒 560–8531 大阪府豊中市待兼山町 1–3; sugimoto.tomoyuki.es@osaka-u.ac.jp
[2] 筑波大学 医学医療系:〒 305–8575 茨城県つくば市天王台 1–1–1; kazushi.maruo@gmail.com
[3] 和歌山県立医科大学 医学部:〒 641–8509 和歌山市紀三井寺 811–1; toshibow2000@gmail.com

簡略化バージョンに対して技術ノートを与えている．その後，Scornet et al. (2015)は，ブートストラップが非復元抽出サブサンプリングにおき換えられている点を除き，元のランダムフォレストの現実的な設定のもとで，共変量次元を固定し，2 乗期待値の意味での一致性を与え，さらに，Wager (2016) [Theorem 4.3] は，共変量次元を固定しない高次元の状況において，ランダムフォレストの各点一致性を示している．Wager and Athey (2018)は，正直性(Honesty)などの概念とダブルサンプリング手法を導入して，ランダムフォレストによる応答関数の各点推定において漸近正規性を確立することに成功し，フォレスト法に基づく条件付きの処理効果の(因果)推定法の因果フォレストを提案している (中村, 2020)．応答関数の各点推定値に関する分散の一致推定値を，ランダムフォレストなどのバギング学習器に対する分散推定の無限小ジャックナイフ法 (Wager et al., 2014; Efron, 2014)を適合させて計算可能としている点も実用的なツールとして著しく有用である．因果フォレスト法 (Wager and Athey, 2018)の着想は，ランダムフォレスト法の推定方程式による定式化として一般化され (Athey et al., 2019)，その生存解析への応用として Cui et al. (2023)は右側打ち切りデータに利用可能な理論とツールへと発展させている．また，Athey and Wager (2019)では，クラスター化された連続観測値に対して，因果フォレストを用いてノンパラメトリックなランダム効果あてはめの変化形を与えている．本論文では，このような樹木法の発展を体系的に概説する．

2. 生存樹木法とその例示

2.1 生存データ

生存解析における基本的データ集合の設定を与える．個人 i の真のイベント時間 T_i^*, 真の右側中途打ち切り時間 C_i に対して, 観測イベント時間は $T_i = \min(T_i^*, C_i)$ である．$\Delta_i = \mathbb{1}(T_i^* \leq C_i)$ (1：完全データ，0：打ち切りデータ)は右側中途打ち切り指標であり，個人 i の共変量は p 個の共変量のベクトル $\boldsymbol{Z}_i = (Z_{i1}, \ldots, Z_{ip})$ で構成される．また，観測機構は独立打ち切りの仮定を満たすとする，すなわち，T_i^* と C_i は，共変量 $\boldsymbol{Z}_i$ を所与として条件付きに独立であると仮定する．データは独立した N 人の被験者 $\mathcal{L}_N = \{(T_i, \Delta_i, \boldsymbol{Z}_i), i = 1, \ldots, N\}$ に対して利用可能である．ここでは，一貫して，共変量 $\boldsymbol{Z}_i$ は，時点 0 で測定され，時間変化しないとする．

例示として，R パッケージ `randomForestSRC` 内のデータ `peakVO2` を用いる．これは心肺運動負荷試験を受けた 2231 名の収縮期心不全患者の予後におけるリスク因子を同定するための研究 (Hsich et al., 2011)からのデータ集合であり，我々の例示では，全データのうち男性 $N = 1629$ 名を用いる．このとき，死亡までの観測時間(T)と打ち切り指標(Δ)に関係する共変量($\boldsymbol{Z}$)として，糖尿病や冠動脈疾患などの 8 項目の過去の病歴・治療歴，β 遮断薬やカルシウム拮抗薬など 15 項目の薬剤使用状況，トレッドミル運動時間や最高酸素摂取量などの 6 つの心機能指数，BUN などの 5 つの血液検査値，および年齢，人種，BMI，喫煙からなる $p = 38$ 個の変数を含む．図 1 は，Cox 回帰を適用し，情報量規準 AIC によって変数選択されたハザード比に関する結果であり，変数の順番はハザード比に関する統計的有意性の順に並び替えている．Cox 回帰では，予後因子の同定とその大きさの評価をハザード比の観点で解釈できる点で優れている．また Cox 回帰のあてはめから，有意義な予後予測モデルの作成を行うこともできる．この予測性能については，2.3 節においてランダムフォレストなどの樹木法と比較する．

2.2 生存木

生存木の開発は，CART 法 (Breiman et al., 1984)や自動交互作用検出(Auto Interaction Detection; AID)法 (Morgan and Sonquist, 1963)などの既存の樹木構造接近法を生存データに拡張することを目的として展開された (例えば，Gordon and Olshen, 1985; Ciampi et al., 1986;

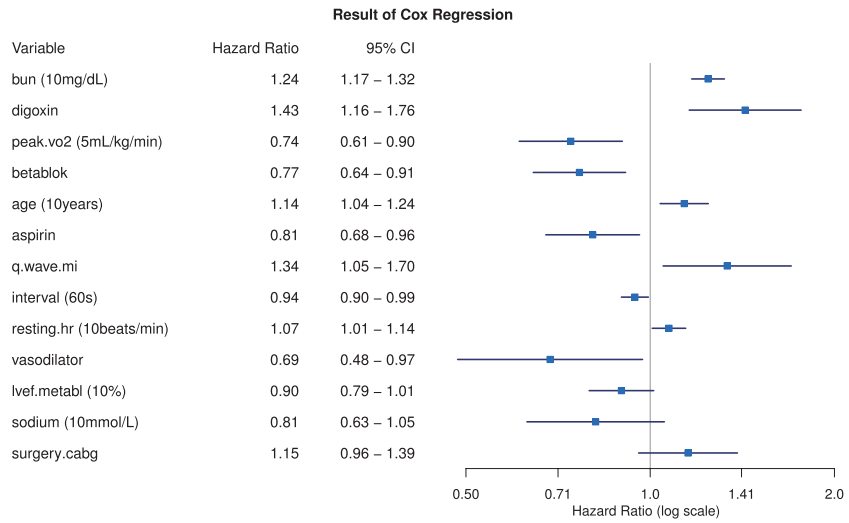

図 1. データ peakVO2(男性)に対する AIC に基づいて選択された 13 個の共変量の Cox 回帰の結果.

```
tree<-rpart(Surv(ttodead, died)~., data=peakVO2m, cp=0)
tree0<- prune(tree,cp=tree$cptable[which.min(tree$cptable[,"xerror"]),"CP"])
plot(as.party(tree0))
```

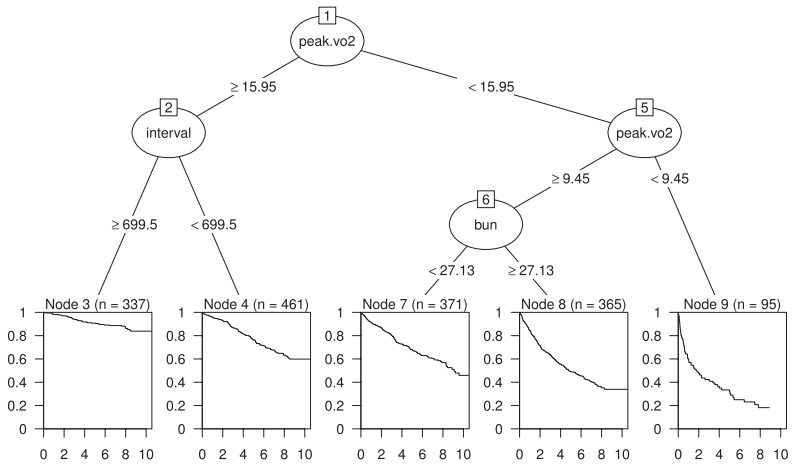

図 2. データ peakVO2(男性)に対する最小交差検証生存木.

Segal, 1988; Davis and Anderson, 1989; LeBlanc and Crowley, 1992)．基本的な生存木の方法論が確立された後は，生存木のアンサンブル拡張を含め，多変量生存データ，相関のある生存データ，時間依存共変量，離散尺度イベント時間変数などを扱う問題などの研究も進められた．生存木に関するレビューは LeBlanc and Crowley (1995)，Bou-Hamad et al. (2011)，Zhou and McArdle (2015)などが詳しく，Zhou and McArdle (2015)の図 2 は，公表された生存木の分岐ルール，刈り込み，実装の分類を明確に与えている．

生存木を含む樹木構造接近法では，トレーニング・データとして，全データ集合 $\mathcal{L}_N$ のすべ

て，もしくは $\mathcal{L}_N$ から復元抽出もしくは非復元抽出した部分集合を用いて木を構築する．図 2 は，データ peakVO2(男性)を用いて，R の rpart 関数により計算された生存木を partykit を通して描画したものである．$\mathcal{T}$ を構築された生存木とする．図 2 の生存木 $\mathcal{T}$ は 9 個のふし(node)$\hbar_\ell$, $\ell = 1, \ldots, 9$ により構成されている．各ふしは対応するルールに該当する個人ラベル i の集合を表すものである．とくに，最も上位にあるふし $\hbar_1$ はルート(root)と呼ばれる．終端ふし(terminal node)の集まり $\widetilde{\mathcal{T}} = \{\hbar_3, \hbar_4, \hbar_6, \hbar_8, \hbar_9\}$ は，葉(leaf)とも呼ばれ，通常，終端ふし内のデータを用いて(生存)予測子を構築する．図 2 の終端ふしの要約には，Kaplan-Meier 推定値が描かれており，終端ふし 3 から 9 にかけて予後が悪くなる様子が見て取れる．最高酸素摂取量(peak.vo2)は Cox 回帰の結果と同様に予後に重要な因子となったが，トレッドミル運動時間(interval)や尿素窒素値(bun)との組み合わせ(交互作用)に基づいたより解釈しやすい予後予測が得られている．

R の rpart 関数で作成される現在の生存木は，CART 法で用いる残差を生存時間データに適応させるためデビアンス残差を利用した方法（LeBlanc and Crowley, 1992)の指数分布版が採用されている．$\hbar_L$ と $\hbar_R$ を，ふし $\hbar$ から生成される娘ふしとする．説明変数 Z_{ij} が数値型であれば，分岐点 c_j を用いて，$\{i \in \hbar_L : Z_{ij} < c_j\}$ と $\{i \in \hbar_R : Z_{ij} \geq c_j\}$ の 2 分割を行う．カテゴリ型であれば，組み合わせに基づく 2 分割となる．$R(\hbar)$ をふし内残差とする．親ふし $\hbar$ から娘ふし $\hbar_L, \hbar_R$ への 2 分岐ルールの選択は，ふし $\hbar$ 内の不純度指標の $R(\hbar)$ を最も減少させる $\hbar_L, \hbar_R$ のペア，すなわち，減少量 $\Delta R_\hbar(s_\hbar) = R(\hbar) - \{R(\hbar_L) + R(\hbar_R)\}$ を最も大きくする 2 分岐ルール $s_\hbar \in \mathcal{S}_\hbar$ が選択される．ここに，$\mathcal{S}_\hbar$ は，各ふし $\hbar$ ですべての説明変数の対応するすべての 2 分岐ルールの候補集合である．適切な樹木の大きさを決定するために，CART 法(Breiman et al., 1984)では，複雑度コスト(cost-complexity)刈り込み法と交差検証法を提案している．一方，生存木では，ログランク検定統計量に基づく分岐ルール（Segal, 1988)のニーズも高い．LeBlanc and Crowley（1993)は，2 分岐ルールに検定統計量の最大化を採用する場合に，複雑度分岐(split-complexity)法を提案している．複雑度パラメータの統計的解釈は，前者の場合は予測指標の罰則項，後者の場合は有意性指標の自由度調整項に該当する．

決定木はこのような形で作成されるが，データの見方についての一つの提案であって，実地では，生存木を含め，決定木の公式的な作り方に忠実に拘る必要はない．例えば，分岐候補値は，解釈のしやすいカテゴリ分けにしてもよいし，分岐基準もデータの特性から関心のある統計的指標に変更してよい．また，決定木は，例えば，観測データの 90% で学習させた場合に得られる木構造が大きく異なるなど，データのばらつきにより異なる木が生成される．安定した生存木を利用したい場合には，例えば 90% のデータでランダムサンプリングを行い推定された木構造の安定性を確かめることが推奨される．決定木は人が解釈のしやすい思考構造を提供するため，例えば，ブートストラップ標本を用いて類似する性能をもつ生存木を複数作り出して，分析者の仮説に合致する木を探し出すことも考えられる．

2.3 生存フォレスト

ランダムフォレスト(RF)（Breiman, 2001)は，大きく成長させて過剰あてはめを引き起こした樹木予測子を平均化することで過学習を回避するという他の機械学習法(多くは縮小により過学習を回避する)の中では見られない独自性をもち，高い予測性能をもつノンパラメトリックな機械学習手法を提供するにも関わらず，並列計算も可能という点が顕著な特徴である．そのような RF を生存データに適応させる研究が Breiman（2003)，Hothorn et al.（2006a)，Ishwaran et al.（2008）などにより進められた．生存フォレストの構成方法は標準的な RF と同様であるが，生存解析に特有の分岐方法や予測指標が必要になる．生存フォレストの一般的な計算方法は以下の通りである：

ステップ 1．B 個のブートストラップ標本 $\mathcal{L}_N^{\mathrm{IB}(b)}$ $b = 1, \ldots, B$ を抽出する．

ステップ 2．各ブートストラップ標本 $\mathcal{L}_N^{\mathrm{IB}(b)}$ に基づいて生存木を成長させる（刈り込み無し）：

(i) 各生存木のふしにおいて，ある説明変数（予測変数，特徴量）の集まり（単一，線形結合）を選択する．

(ii) (i) で選択した説明変数によって分割されるすべての 2 群の中から，右側打ち切りデータに適した分割基準（例えば，ログランク検定）に基づいて，最良の二分割（娘ふしへの分割）を見つける．

(iii) 各娘ノードに対して，(i) - (iii) の手順を停止基準に達するまで再帰的に繰り返す．

ステップ 3．B 個の生存木の終端ふしからの生存情報を集約し，生存予測モデルのアンサンブルを得る．

生存フォレストによるリスク予測方法を記述するため，$\mathcal{T}^{(b)}$ を b 番目のブートストラップ標本に対して生成された生存木とし，共変量 $\boldsymbol{z}$ をもつ個体が位置する終端ふしを $\mathcal{T}^{(b)}(\boldsymbol{z})$ とする．ブートストラップ標本 $\mathcal{L}_N^{\mathrm{IB}(b)}$ 内には，元のデータ集合 $\mathcal{L}_N$ の個体が複数含まれ得る．個体 i が b 番目のブートストラップ標本 $\mathcal{L}_N^{\mathrm{IB}(b)}$ に出現する回数を $\aleph_i^{(b)}$ とする．$\mathcal{L}_N^{\mathrm{IB}(b)}$ に個体 i が含まれていない場合，$\aleph_i^{(b)} = 0$ である．共変量 $\boldsymbol{z}$ に対応する生存木の終端ふし $\mathcal{T}^{(b)}(\boldsymbol{z})$ において，$\bar{\mathcal{N}}^{(b)}(t|\boldsymbol{z})$ を時刻 t までの打ち切りされていないイベント数，$\bar{\mathcal{Y}}^{(b)}(t|\boldsymbol{z})$ を時刻 t におけるアトリスク数とすると，これらは，計数過程の記法（Fleming and Harrington, 1991）を用いて

$$\bar{\mathcal{N}}^{(b)}(t|\boldsymbol{z}) = \sum_{i=1}^{N} \aleph_i^{(b)} \mathbb{1}(\boldsymbol{Z}_i \in \mathcal{T}^{(b)}(\boldsymbol{z})) \mathcal{N}_i(t), \quad \bar{\mathcal{Y}}^{(b)}(t|\boldsymbol{z}) = \sum_{i=1}^{N} \aleph_i^{(b)} \mathbb{1}(\boldsymbol{Z}_i \in \mathcal{T}^{(b)}(\boldsymbol{z})) \mathcal{Y}_i(t).$$

と書ける．ここに，$\mathcal{N}_i(t) = \mathbb{1}(T_i \leq t, \Delta_i = 1)$, $\mathcal{Y}_i(t) = \mathbb{1}(T_i > t)$ である．Ishwaran et al.（2008）の生存フォレストでは，各生存木に基づく Nelson-Aalen 推定量 $\hat{H}^{(b)}(t|\boldsymbol{z}) = \int_0^t d\bar{\mathcal{N}}^{(b)}(u|\boldsymbol{z})/\bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})$ を集約することで予測子を構築する．すなわち，ブートストラップ標本のインバック（in-bag）データを用いて，各木の終端ふしに対する条件付き累積ハザード関数を $\hat{H}^{(b)}(t|\boldsymbol{z})$ により推定することで，生存フォレストの予測生存関数を

$$\hat{S}^{\mathrm{RSF}}(t|\boldsymbol{z}) = \exp\left(-\frac{1}{B}\sum_{b=1}^{B} \hat{H}^{(b)}(t|\boldsymbol{z})\right) \tag{2.1}$$

として構成する．一方，Hothorn et al.（2004）では，予測生存関数の構成に

$$\hat{S}^{\mathrm{CSF}}(t|\boldsymbol{z}) = \prod_{u \leq t}\left(1 - \frac{\sum_{b=1}^{B} d\bar{\mathcal{N}}^{(b)}(u|\boldsymbol{z})}{\sum_{b=1}^{B} \bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})}\right) \approx \exp\left(-\int_0^t \frac{\sum_{b=1}^{B} d\bar{\mathcal{N}}^{(b)}(u|\boldsymbol{z})}{\sum_{b=1}^{B} \bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})}\right) \tag{2.2}$$

を用いることを提案している．予測生存関数の (2.1) と (2.2) の違いに関して，Mogensen et al.（2012）は関係式

$$\frac{\sum_{b=1}^{B} d\bar{\mathcal{N}}^{(b)}(u|\boldsymbol{z})}{\sum_{b=1}^{B} \bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})} = \frac{1}{B}\sum_{b=1}^{B}\left[\frac{\bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})}{\frac{1}{B}\sum_{b=1}^{B} \bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})}\right] \frac{d\bar{\mathcal{N}}^{(b)}(u|\boldsymbol{z})}{\bar{\mathcal{Y}}^{(b)}(u|\boldsymbol{z})}$$

から，累積ハザード関数の集約において，(2.2) ではリスク数の多い終端ふしほどより多くの重みを割り当てる一方，$\hat{S}^{\mathrm{RSF}}(t|\boldsymbol{z})$ はすべての終端ふしに等しい重みを割り付ける違いがあることを指摘している．

良い性能をもつ RF を構築するための本質は，樹木あてはめの残差の相関が樹木間で小さくなるようにランダム量を上手く注入することである．そのような調整（turning）パラメータとして代表的なものは，各分岐でランダムに選択される候補変数の数（`mtry`）と最小ふしサイズ（`nodesize`）である（例えば，Segal, 2004; Probst et al., 2019）．変数の線形結合を分岐に用いる

RF（Forest-RC）（Breiman, 2001）は計算上のコストから R，Python では実装されていないが，調整パラメータの選択肢が増え，予測性能のさらなる向上が見込める（杉本 他, 2005）．各調整パラメータの目安となる（デフォルト）値はあるが，実際には，データの特徴によって最適な調整パラメータは異なるため，計算コストに配慮しつつ，調整パラメータのいくつかの組み合わせを事前に試行した上で，良い予測性能を与えるものを選択することが望ましい．このことは 3.2 節以降で紹介する因果フォレストにおいても同様である．

予測性能 あてはめたモデルの予測性能の評価では，生存木やフォレストに尤度基準の導入が難しいことから，Harrell et al.（1982）の C-指標（Heagerty and Zheng, 2005）や Brier スコアが用いられる．また，生存予測では C-指標が 2 分類予測の ROC 曲線 AUC ほど理解しやすい指標ではなく，期待 Brier スコア $\mathrm{BS}^*(t;\hat{S}) = \mathrm{E}\big[\{\mathbb{1}(T_i^* > t) - \hat{S}_i(t)\}^2\big]$ には，個人生存関数の推定値が真の生存関数と等しい場合に最小になるという絶対的基準があるため，生存時間データの予測性能評価では，Brier スコアが好まれることも多い．ここに，$\hat{S}_i(\cdot)$ は，個人 i の生存関数の推定値である．右側中途打ち切りの生存データにおける Brier スコアの推定値として

$$\widehat{\mathrm{BS}}(t) = \frac{1}{N}\sum_{i=1}^{N}\big[\{1-\hat{S}_i(t)\}^2\mathbb{1}(T_i > t)\hat{S}_i^C(t)^{-1} + \hat{S}_i(t)^2\mathbb{1}(T_i \le t, \Delta_i = 1)\hat{S}_i^C(T_i)^{-1}\big] \tag{2.3}$$

が提案されている（Graf et al., 1999）．ここに，$\hat{S}_i^C(t)$ は個人 i の右側中途打切り時間の生存関数の推定値であり，Kaplan-Meier 推定法を用いることが基本的であるが，Cox 回帰など共変量を用いたモデル推定も有用で，性能を向上させることが報告されている．期待 Brier スコアは小さいほど良いという性質をもつが，実際には (2.3) のように推定する必要ある．Brier スコアによる予測性能の評価は，交差検証法などを用いてモデル構築と期待値に関する推定のために起こるバイアスを防ぐことが望ましい．全体的な時間の総合指標として，$\widehat{\mathrm{BS}}(t)$ を対象となる時間の上（ここでは $(0, T_{\max}]$ とする）で合計した累積 Brier スコア

$$\widehat{\mathrm{IBS}} = \frac{1}{T_{\max}}\int_0^{T_{\max}} \widehat{\mathrm{BS}}(t)dt \approx \frac{1}{T_{\max}}\sum_m \widehat{\mathrm{BS}}(t_m)(t_m - t_{m-1})$$

が利用できる．図 3 は，二つの Cox 回帰（`Cox.full`, `Cox.AIC`），生存木（`SCART`），Ishwaran et al.

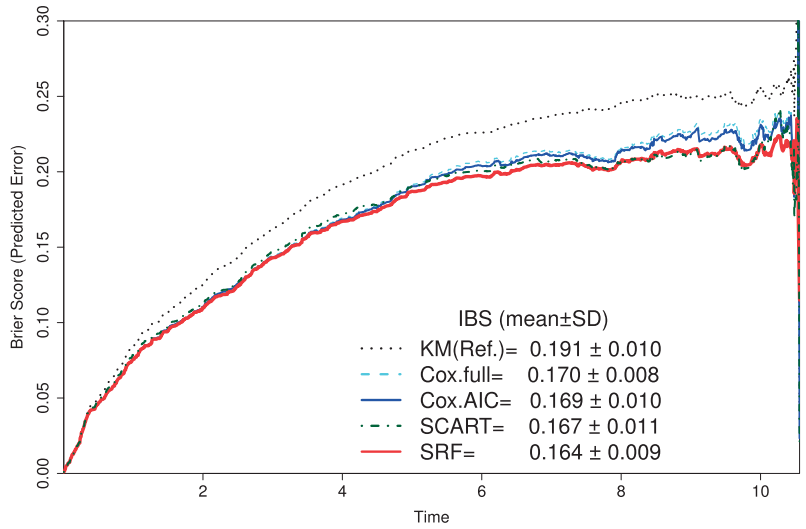

図 3. データ `peakVO2`（男性）における Cox 回帰，生存木（SCART），生存フォレスト（RSF）の Brier スコアの比較．

(2008)の生存フォレスト(RSF)(ライブラリ randomForestSRC の rfsrc 関数)をデータ peakVO2(男性)に適用した結果を示す．予測性能は，20 回の内的妥当性検証(ホールドアウト法：訓練 80%，検証 20%)を用いた Brier スコアの平均で要約している．このデータでは，AIC 選択の Cox 回帰(Cox.AIC)よりも生存フォレストの方が僅かではあるが累積 Brier スコアの性能はよく，とくに RSF と SCART は時間的に後半の予測でより優位である．生存木は，一般に，AIC 選択 Cox 回帰よりも累積 Brier スコアの性能が落ちることが多い(今回のデータの傾向とは異なる)．なお，Brier スコア算出のライブラリ pec は現時点で rpart 関数からの生存木の結果に対応していない．実装において，生存木の予測生存関数に，終端ふしの部分集団の Kaplan-Meier 推定値を用いると標本サイズが小さくなり Brier スコアが不安定となり得るため，図 3 では，終端ふしをダミー変数とする Cox 回帰の Breslow 推定値に基づく推定生存関数を用いて予測 Brier スコアを安定化させている．なお，生存フォレストの調整パラメータは，ここでは，予測性能と計算時間の観点で，nodesize は 15(rfsrc 関数のデフォルト値)，mtry は 5 (デフォルト値は $\lceil\sqrt{p}\rceil = 7$)に設定した．調整パラメータの設定，N が小さいときの挙動に関するさらなる議論は，Web Supplementary Materials A に与える．

3. 変数効果の推測

本節では，生存フォレストによる変数重要度指標，因果生存フォレストによる変数効果の統計的推測の最近の発展を通して，フォレストからの変数効果の推測方法を概説する．

3.1 変数重要度指標

樹木法では，各手法に応じて変数重要度(VIMP)の推定ツールが提案されている (杉本 他, 2005)．手法間で多少の違いはあるものの，概念的には，変数の並び替えやマスキングを活用することにあり，それらの操作のもとで，各変数に起因する予測誤差を評価し，各変数の重要度を評価する．フォレスト法では，重要度の評価において計算コストが高い交差検証法を避けることができる (Breiman, 1996, 2001)．すなわち，オリジナルな CART 法とは異なり，ブートストラップ標本 $\mathcal{L}_N^{\mathrm{IB}(b)} := \mathcal{L}_N^{\mathrm{IB}}(\Theta^{(b)})$ を訓練データに用いて各樹木 $T^{(b)}$ を構成するため，訓練データであるブートストラップ標本で用いられなかった，いわゆるアウトオブバック(OOB)データ $\mathcal{L}_N^{\mathrm{OOB}}(\Theta^{(b)}) := \mathcal{L}_N \setminus \mathcal{L}_N^{\mathrm{IB}}(\Theta^{(b)})$ を評価データに用いて予測誤差を推定することができる．ここに，$\Theta^{(b)}$ は木を成長させることと OOB データの並び替えに関するランダム性を含む要因とする．並び替え(Breiman-Cutler)重要度の計算では，OOB データの j 番目の共変量のすべての値 $Z_{ij}, i \in \mathcal{L}_N^{\mathrm{OOB}}(\Theta^{(b)})$ をランダムに並び替える．このとき，元の Z_{ij} から並び替えられた共変量値を $\tilde{Z}_{ij}$ と書くと，これらの新たなデータの共変量

$$\tilde{\boldsymbol{Z}}_i^{(j)} = (Z_{i1}, \ldots, Z_{ij-1}, \tilde{Z}_{ij}, Z_{ij+1}, \ldots, Z_{ip}), i \in \mathcal{L}_N^{\mathrm{OOB}}(\Theta^{(b)})$$

を木にあてはめて誤差率を計算する．この新しい誤差率が元の OOB 誤差率を上回る量が，木における j 番目の変数の重要度

$$Imp(Z^{(j)}, \Theta^{(b)}, \mathcal{L}_N) = \frac{\sum_{i \in \mathcal{L}_N^{\mathrm{IB}}(\Theta^{(b)})} \{\ell(S_i, \hat{S}_i(\tilde{\boldsymbol{Z}}_i^{(j)}, \Theta^{(b)}, \mathcal{L}_N)) - \ell(S_i, \hat{S}_i(\boldsymbol{Z}_i, \Theta^{(b)}, \mathcal{L}_N))\}}{|\mathcal{L}_N^{\mathrm{IB}}(\Theta^{(b)})|}$$

として定義して，フォレスト全体の木に対する平均

$$Imp(Z_j, \Theta^{(1)}, \ldots, \Theta^{(B)}, \mathcal{L}_N) = \frac{1}{B} \sum_{b=1}^{B} Imp(Z^{(j)}, \Theta^{(b)}, \mathcal{L}_N)$$

を取ることで VIMP を算出する．ここに，S_i と $\hat{S}_i$ は，それぞれ，個人 i の生存関数の真値と

推定値，$\ell(S_i, \hat{S}_i(\boldsymbol{Z}_i, \Theta^{(b)}, \mathcal{L}_N))$ は生存木に関する損失関数であり，典型的なものとして Brier スコアもしくは C-指標を想定する．無限に生成した木に対する VIMP 推定量は，$B \to \infty$ の極限を取ることで次のように定義される：

$$Imp(Z_j, \mathcal{L}_N) = \mathbb{E}_\Theta[Imp(Z^{(j)}, \Theta, \mathcal{L}_N)].$$

上記の定式化では，いずれもフォレスト予測子を明示的に使用せず，各々が単一の木の予測子のみで書かれることに注意したい．

フォレストからの VIMP の分散推定における深刻な問題として，ブートストラップの直接的な適用が機能しないことが挙げられる．フォレストの構築において，既にブートストラップ標本を使用しており，分散推定のために改めてブートストラップを適用すれば，二重ブートストラップ標本を生成する必要があり，計算上の問題に加えて，通常の OOB 利用における妥当性に問題が生じる．Ishwaran and Lu (2018)は，二重ブートストラップ標本において，妥当な OOB の生成確率は 0.164 であることを導き，VIMP の分散推定のための.164 ブートストラップ法，およびこの方法から派生して，サブサンプリング法と d 例削除ジャックナイフ法を提案している．Williamson et al. (2021)は生存フォレストへの直接的な適用ではないものの，変数のマスキングに基づく VIMP 指標に対して，ノンパラメトリックな影響関数を導出し，VIMP 指標の補正推定値を提案し，提案の推定量の漸近ガウス性と漸近分散の一致推定量を与えている．ただし，VIMP 指標がゼロである帰無仮説では，パラメータ空間の境界上にある設定のため，帰無仮説のもとでは不向きな補正方法であり，さらなる工夫を要する課題としている．

3.2 因果フォレスト

VIMP 指標の性質の理論的研究が徐々に進行する中で，Wager and Athey (2018)は，樹木法に基づく条件付きの処理効果の(因果)推定に関する画期的な方法として因果フォレスト法を提案した．因果木や因果フォレストに対する主要な着想は，利用可能なデータ集合 $\mathcal{L}_N$ を，樹木を構成するデータ集合 ($\mathcal{L}_\bullet^{\mathcal{I}}$) と，処理効果を推定するためのデータ集合 ($\mathcal{L}_\bullet^{\mathcal{J}}$) に分割して各プロセスを実施することで彼らの用語「正直さ(honesty)」を実現することである．つまり，この正直さの着想によって，樹木の分岐の決定(予測モデルの構造推定)と処理効果の推定において，同じデータを用いないことで各推定の独立性を与え，因果推定の偏りを減らすことができる．このとき，樹木構成でトレーニングデータの半分を捨てるため，関数推定に関する情報損失が生じるが，$\mathcal{L}_\bullet^{\mathcal{I}}$ と $\mathcal{L}_\bullet^{\mathcal{J}}$ のデータ分割を再度ランダムサンプリングし直すため，この意味においてデータを無駄にすることがないと主張している．このような樹木の生成方式は，ダブルサンプル木(double-sample tree)と呼ばれ，以下にそのアルゴリズムに纏めたものを与える：

ステップ 0．データ集合 $\mathcal{L}_N$ の個人 i は応答観測値 $Y_i(= y(T_i, \Delta_i))$ と特徴量 $\boldsymbol{Z}_i$ をもつ．因果木では，特徴量から関心のある処理変数を分離し，その処理変数を W_i と書く．葉の最小サイズを k とする．

ステップ 1．個体ラベル $\{1, \ldots, N\}$ から M_N 個の要素をランダムに非復元抽出して得たデータ集合を $\mathcal{L}_{M_N}^{\mathrm{sub}}$ とする．$\mathcal{L}_{M_N}^{\mathrm{sub}}$ を二つの互いに素なデータ集合 $\mathcal{L}_{\lfloor M_N/2 \rfloor}^{\mathcal{I}(b)}$, $\mathcal{L}_{\lceil M_N/2 \rceil}^{\mathcal{J}(b)}$ にランダム分割する．

ステップ 2．再帰分割により木を成長させる．樹木成長では，$\mathcal{L}_{\lceil M_N/2 \rceil}^{\mathcal{J}(b)}$ からの任意のデータと，$\mathcal{L}_{\lfloor M_N/2 \rfloor}^{\mathcal{I}(b)}$ からの特徴量のみを用いて分岐を選択する(因果木の場合では処理変数 W は葉の例数管理のため用いる)．ただし，データ集合 $\mathcal{L}_{\lfloor M_N/2 \rfloor}^{\mathcal{I}(b)}$ の応答観測値は用いない．また各分岐の構成において，回帰木では，$\mathcal{L}_{\lfloor M_N/2 \rfloor}^{\mathcal{I}(b)}$ の観測値を k 個以上含むように，因果木では，$\mathcal{L}_{\lfloor M_N/2 \rfloor}^{\mathcal{I}(b)}$ の観測値を各処理群ごとに k 以上含むように管理する必要がある．

ステップ 3．ステップ 2 で得た木 $\mathcal{T}^{(b)}$ に，データ集合 $\mathcal{L}^{\mathcal{I}(b)}_{\lfloor M_N/2 \rfloor}$ をあてはめて，葉ごとの応答観測値の要約を求める．

ダブルサンプル木は，因果推論のためだけでなく，通常の樹木法にも適用可能であり，関数近似の区間推定等に利用可能である．因果ツリーの文脈では，因果推論における SUTVA，無交絡性（Unconfoundedness），重複あり（overlap）の基本条件は成り立つとする．十分に小さい葉内においては，個人 $i \in \mathcal{T}^{(b)}(\boldsymbol{z})$ に対応する応答と処理変数のペアがランダム化実験で生じたように見えるため潜在的な交絡調整効果が期待できる．このとき，条件付き処理効果（CATE）$\tau(\boldsymbol{z}) = \mathrm{E}[Y_i^{(1)} - Y_i^{(0)} | \boldsymbol{Z}_i = \boldsymbol{z}]$ を，終端ふし $\mathcal{T}^{(b)}(\boldsymbol{z})$ を用いて

$$\hat{\tau}^{(b)}_{\mathrm{tree}}(\boldsymbol{z}) = \frac{\sum_{\{i \in \mathcal{T}^{(b)}(\boldsymbol{z}) : W_i = 1\}} Y_i}{|\{i \in \mathcal{T}^{(b)}(\boldsymbol{z}) : W_i = 1\}|} - \frac{\sum_{\{i \in \mathcal{T}^{(b)}(\boldsymbol{z}) : W_i = 0\}} Y_i}{|\{i \in \mathcal{T}^{(b)}(\boldsymbol{z}) : W_i = 0\}|} \tag{3.1}$$

により推定する．ここに，$Y_i^{(1)}$ と $Y_i^{(0)}$ はそれぞれ処理 $W_i = 1, 0$ を与えたときの Y_i の潜在的な結果である．因果フォレストは，B 個の単一のダブルサンプル木からの結果 $\hat{\tau}^{(b)}_{\mathrm{tree}}(\boldsymbol{z})$ を集約して処理効果のより強固な推定値 $\hat{\tau}(\boldsymbol{z}) = B^{-1} \sum_{b=1}^{B} \hat{\tau}^{(b)}_{\mathrm{tree}}(\boldsymbol{z})$ をもたらす．Wager and Athey（2018）はこの推定値 $\hat{\tau}(\boldsymbol{z})$ が，いくつかの正則条件（サブ標本サイズ M_N の条件，正直性，ランダム分岐，α-正則性，対称性，Lipschitz 連続）のもとで，$N \to \infty$ として

$$(\hat{\tau}(\boldsymbol{z}) - \tau(\boldsymbol{z})) / \sqrt{\mathrm{V}[\hat{\tau}(\boldsymbol{z})]} \xrightarrow{D} N(0, 1) \tag{3.2}$$

となる漸近正規性を確立し，分散 $\mathrm{V}[\hat{\tau}(\boldsymbol{z})]$ の一致推定値

$$\hat{V}_{IJ}(z) = \frac{N-1}{N} \left(\frac{N}{N - M_N} \right)^2 \sum_{i=1}^{N} \widehat{\mathrm{Cov}}[\hat{\tau}_{\mathrm{tree}}(\boldsymbol{z}), \aleph_i]^2 \tag{3.3}$$

を与えている．ここに，$\widehat{\mathrm{Cov}}[\hat{\tau}_{\mathrm{tree}}(\boldsymbol{z}), \aleph_i] = \sum_b (\hat{\tau}^{(b)}_{\mathrm{tree}}(\boldsymbol{z}) - \hat{\tau}(\boldsymbol{z}))(\aleph_i^{(b)} - \bar{\aleph}_i)/B$ であり，$\bar{\aleph}_i = B^{-1} \sum_b \aleph_i^{(b)}$ であり，$\aleph_i^{(b)}$ は第 i サンプルが b 番目の木 $\mathcal{T}^{(b)}$ に使用されれば 1，そうでなければ 0 を示す（ダブルサンプル木では，第 i サンプルが，データ集合 $\mathcal{L}^{\mathcal{I}(b)}_{\lfloor M_N/2 \rfloor}$ または $\mathcal{L}^{\mathcal{J}(b)}_{\lceil M_N/2 \rceil}$ のいずれかに含まれることを意味する）．正則条件についての詳細，定理の証明は Wager（2016），Wager and Athey（2018）を参照されたい．

3.3 因果生存フォレスト

Wager and Athey（2018）の因果フォレストは，連続な応答観測値 Y_i に対する方法であるため，生存時間データに適用するためには，更なる工夫が必要である．因果フォレストそのものを生存解析用に開発し直すか，生存時間情報 (T_i, Δ_i) のある変換 y を与えて，$Y_i = y(T_i, \Delta_i)$ とおくことが考えられる．例えば，変換 y として，通常の CART 樹木を生存データに適用するために帰無モデルのマルチンゲール残差を利用する方法（Therneau et al., 1990）は直裁的である（ただしこの場合の結果の解釈は難しい）．また，中途打ち切りがなければ，ある定数 $x > 0$（horizon）に対して，RMST（制限付き平均生存時間）の推測では，$y(T_i, 1) = T_i \wedge x$, 生存率の推測では，$y(T_i, 1) = \mathbb{1}(T_i > x)$ と定めることが可能である．これに先立って，Athey et al.（2019）は，因果フォレストに基づく処理効果の推測において，ダブルサンプルフォレストと傾向フォレストを統合する動機から，観測値を直接的に用いる形式から推定方程式

$$\sum_{i=1}^{N} \alpha_i(\boldsymbol{z}) \psi_{\hat{\tau}(\boldsymbol{z}), \hat{\eta}(\boldsymbol{z})}(Y_i) = 0 \tag{3.4}$$

に基づく方法へと拡張し，因果フォレストを局所推定の枠組みで一般化した．ここに，

$\psi_{\hat{\tau}(\boldsymbol{z}),\hat{\eta}(\boldsymbol{z})}(\cdot)$ は関心のある推定量 $\hat{\tau}(\boldsymbol{z})$ と局外推定量 $\hat{\eta}(\boldsymbol{z})$ を含む式であり，重み $\alpha_i(\boldsymbol{z})$ は $B^{-1}\sum_{b=1}^{B}\alpha_i^{(b)}(\boldsymbol{z})$ と構成され，それぞれの $\alpha_i^{(b)}(\boldsymbol{z})=\mathbb{1}(i\in\mathcal{T}^{(b)}(\boldsymbol{z}))$ は，生成した b 番目の単一の木 $\mathcal{T}^{(b)}(\boldsymbol{z})$ から算出される．このような Athey et al. (2019) の一般化 RF の枠組みでの因果木の生成でも，正直性を保証するため，互いに素な二つのデータ集合に分け，一方のデータ集合を樹木構成に用い，もう一方のデータ集合を $\alpha_i^{(b)}(\boldsymbol{z})$ の計算に用いる．ただし，樹木構成における分割基準は，通常の CART 法の分割基準とは異なり，親ふし $\hbar$ において

$$\widetilde{\Delta R}(\hbar_L,\hbar_R)=\frac{1}{|\{i\in\hbar_L\}|}\left(\sum_{i\in\hbar_L}\rho_i\right)^2+\frac{1}{|\{i\in\hbar_R\}|}\left(\sum_{i\in\hbar_R}\rho_i\right)^2$$

を最大にする娘ふし $\hbar_L,\hbar_R$ を選択する．ここに，ρ_i は疑似アウトカムと呼ばれ，親ふし $\hbar$ における $\tau(\boldsymbol{z})$ の推定値の計算に対する第 i 観測値の影響関数に負符号を付けたものに対応する．疑似アウトカム ρ_i は，最小二乗回帰のときには回帰残差であり，因果推論の文脈では処理効果の個人推定値に対応するため，疑似アウトカムを目的変数にすることで標準的な機械学習手法をそのまま適用できる（さらなる詳細は Athey et al., 2019 を参照）．一般化 RF の因果木生成のアルゴリズムを以下に纏める：

ステップ 0-1．因果フォレストのステップ 0-1 と同様に行う（全データから部分標本を非復元抽出し，互いに素なデータ集合 $\mathcal{L}^{\mathcal{J}(b)}_{\lceil M_N/2\rceil}$，$\mathcal{L}^{\mathcal{I}(b)}_{\lfloor M_N/2\rfloor}$ にランダム分割する）．

ステップ 2．$\mathcal{L}^{\mathcal{J}(b)}_{\lceil M_N/2\rceil}$ のデータ集合を用い，再帰的分割により樹木を成長させる．なお，樹木成長では，各ふし $\hbar$ において $\widetilde{\Delta R}(\hbar_L,\hbar_R)$ を最大にする娘ふし $\hbar_L,\hbar_R$ を選択する分割基準を用いる．このような方式で十分成長させた樹木を $\mathcal{T}^{\mathcal{J}(b)}$ とする．

ステップ 3．ステップ 2 で作成した樹木 $\mathcal{T}^{\mathcal{J}(b)}$ に，$\mathcal{L}^{\mathcal{I}(b)}_{\lfloor M_N/2\rfloor}$ のデータ集合をあてはめて，$\alpha_i^{\mathcal{J}(b)}(\boldsymbol{z}),\ i\in\mathcal{L}^{\mathcal{I}(b)}_{\lfloor M_N/2\rfloor}$ を計算する．

ステップ 4．ステップ 1-3 を B 回繰り返して，集計された $\{\alpha_i(\boldsymbol{z})\}$ を用いて，(3.4) の解を求める．

Cui et al. (2023) の因果生存フォレストの方法は，この一般化 RF の局所推定の方法に沿って，Robins et al. (1994) の推定方程式を組み合わせて生存解析に応用している．Robinson (1988) の推定量を構成する式を

$$\psi_\tau(T_i^*,\boldsymbol{Z}_i,W_i;\hat{e},\hat{m})=\{W_i-\hat{e}(\boldsymbol{Z}_i)\}\{y(T_i^*)-\hat{m}(\boldsymbol{Z}_i)-\tau(W_i-\hat{e}(\boldsymbol{Z}_i))\}$$

とする．ここに，真のイベント時間 T_i^* は必ずしも観測可能ではなく，$\hat{e}(\boldsymbol{Z}_i)$ と $\hat{m}(\boldsymbol{Z}_i)$ は $e(\boldsymbol{z})=P(W_i=1\mid\boldsymbol{Z}_i=\boldsymbol{z})$, $m(\boldsymbol{z})=\mathrm{E}[y(T_i^*)\mid\boldsymbol{Z}_i=\boldsymbol{z}]$ の交差あてはめ（cross-fitting, 例えば，Schick, 1986; Chernozhukov et al., 2018）によって得られた推定値とし，$y(T)$ は $y(T)=T\wedge x$ または $y(T)=\mathbb{1}(T>x)$ のいずれかである．

Cui et al. (2023) は，中途打切りデータに対して，処理効果を推定するためのある推定方程式 $\mathrm{E}[\tilde{\psi}_\tau(T_i,\Delta_i,\boldsymbol{Z}_i,W_i,x)]=0$ を得るために，Robins et al. (1994) を一般化して，次のスコア関数

$$\begin{aligned}(3.5)\qquad \widetilde{\psi}_\tau(T_i,\Delta_i,\boldsymbol{Z}_i,W_i,x)&=\Delta_i^x\psi_\tau(T_i,\boldsymbol{Z}_i,W_i)/S^C_{W_i}(T_i\wedge x|\boldsymbol{Z}_i)\\&\quad+(1-\Delta_i^x)\mathrm{E}[\psi_\tau(T_i^*,\boldsymbol{Z}_i,W_i)|T_i^*\wedge x>T_i\wedge x,\boldsymbol{Z}_i,W_i]/S^C_{W_i}(T_i\wedge x|\boldsymbol{Z}_i)\\&\quad-\int_0^{T_i\wedge x}\frac{\lambda^C_{W_i}(t\mid\boldsymbol{Z}_i)}{S^C_{W_i}(t\mid\boldsymbol{Z}_i)}\mathrm{E}[\psi_\tau(y(T_i^*),\boldsymbol{Z}_i,W_i)|T_i^*\wedge x>t,\boldsymbol{Z}_i,W_i]dt\end{aligned}$$

を用いることを提案している．ここに，$\Delta_i^x=\Delta_i\mathbb{1}(T_i\geq x)$, $S^C_w(t|\boldsymbol{z})=P(C_i\geq t\mid W_i=w,\boldsymbol{Z}_i=\boldsymbol{z})$ は中途打切り時間の条件付き生存関数，$\lambda^C_w(t|x)=-\frac{d}{dt}\log S^C_w(t|x)$ は対応するハザード関数

である．(3.5)に，完全データの場合のスコア関数 $\psi_\tau(T_i^*, \boldsymbol{Z}_i, W_i; \hat{e}, \hat{m})$ をあてはめて，実際の推定計算で用いる 2 重ロバスト・スコア関数

$$
\begin{aligned}
(3.6)\quad & \widetilde{\psi}_\tau(T_i, \Delta_i, \boldsymbol{Z}_i, W_i, x; \hat{e}, \hat{m}, \hat{\lambda}_w^C, \hat{S}_w^C, \hat{Q}_w) \\
&= \Bigg(\frac{\hat{Q}_{W_i}(T_i \wedge x \mid \boldsymbol{Z}_i) + \Delta_i^x \big[y(T_i) - \hat{Q}_{W_i}(T_i \wedge x | \boldsymbol{Z}_i) \big] - \hat{m}(\boldsymbol{Z}_i) - \tau(W_i - \hat{e}(\boldsymbol{Z}_i))}{\hat{S}_{W_i}^C(T_i \wedge x | \boldsymbol{Z}_i)} \\
&\qquad - \int_0^{T_i \wedge x} \frac{\hat{\lambda}_{W_i}^C(t \mid \boldsymbol{Z}_i)}{\hat{S}_{W_i}^C(t \mid \boldsymbol{Z}_i)} \big[\hat{Q}_{W_i}(t \mid \boldsymbol{Z}_i) - \hat{m}(\boldsymbol{Z}_i) - \tau(W_i - \hat{e}(\boldsymbol{Z}_i)) \big] dt \Bigg) (W_i - \hat{e}(\boldsymbol{Z}_i))
\end{aligned}
$$

を得る．ここに，$\hat{Q}_w(t|x) = \mathrm{E}[y(T_i^*) \mid \boldsymbol{Z}_i = \boldsymbol{z}, W_i = w, T_i \wedge x > t]$ は変換された生存情報 $y(T_i^*)$ の条件付き期待値であり，$\hat{Q}_w(t|x)$，$\hat{S}_w^C(t|x)$，$\hat{\lambda}_w^C(t|x)$ は交差あてはめされた局外パラメータの推定値である．$\hat{Q}_w(t|x)$ は推定生存関数の数値積分によって，$\hat{\lambda}_w^C(t|x)$ は $-\log(\hat{S}_w^C(t|x))$ の前向き差分によって推定される．生存木成長の分割ルール「$\widetilde{\Delta R}(\hbar_L, \hbar_R)$ の最大化」で用いる疑似アウトカム ρ_i $(i \in \hbar)$ は，属するふし $\hbar$ において，$\rho_i = \widetilde{\psi}_{i,\hat{\tau}_\hbar} / (|\hbar|^{-1} \sum_{j \in \hbar} U_j)$ を用いる（Cui et al., 2023）．ここに，$\widetilde{\psi}_{i,\tau}$ は $\widetilde{\psi}_\tau(T_i, \Delta_i, \boldsymbol{Z}_i, W_i, x; \hat{e}, \hat{m}, \hat{\lambda}_w^C, \hat{S}_w^C, \hat{Q}_w)$ の簡略表記であり，U_j は

$$
U_j = U(T_j, \boldsymbol{Z}_j, W_j, x; \hat{e}, \hat{\lambda}_w^C, \hat{S}_w^C) = (W_j - \hat{e}(\boldsymbol{Z}_j))^2 \left(\frac{1}{\hat{S}_{W_j}^C(T_j \wedge x | \boldsymbol{Z}_j)} - \int_0^{T_j \wedge x} \frac{\hat{\lambda}_{W_j}^C(t|\boldsymbol{Z}_j)}{\hat{S}_{W_j}^C(t|\boldsymbol{Z}_j)} dt \right)
$$

であり，$\hat{\tau}_\hbar$ は $\sum_{i \in \hbar} \widetilde{\psi}_{i,\tau} = 0$ を解くことで得られるふし $\hbar$ 内での τ の推定値である．

一般化 RF のアルゴリズムのステップ 0-4 に従って，因果木を生成し，各重みを評価することで，フォレスト全体としての重み $\{\alpha_i(\boldsymbol{z})\}$ が算出される．このステップ 4 における処理効果のフォレスト型の推定を行う際には，スコア関数(3.6)の中で必要とされる局外成分 $\hat{e}$, $\hat{m}$, $\hat{\lambda}_w^C / \hat{S}_w^C$, $\hat{S}_w^C$, $\hat{Q}_w$ を推定し，それらをフォレストの局所推定方式(3.4)に組み合わることで得られる推定方程式

$$
(3.7)\qquad \sum_{i=1}^N \alpha_i(\boldsymbol{z}) \widetilde{\psi}_{\hat{\tau}(\boldsymbol{z})}(T_i, \Delta_i, \boldsymbol{Z}_i, W_i, x; \hat{e}, \hat{m}, \hat{\lambda}_w^C, \hat{S}_w^C, \hat{Q}_w) = 0
$$

を用いる．処理効果のフォレスト推定値 $\hat{\tau}(\boldsymbol{z})$ は(3.7)を満たす解として得られる．

Cui et al. (2023)は Wager and Athey (2018)の因果フォレストに類似する正則条件と局外成分の一致性のもとで推定量 $\hat{\tau}(\boldsymbol{z})$ が真値 $\tau(\boldsymbol{z})$ の周りで漸近的正規であることを与えている．漸近分散 σ_N^2 は，関数的デルタ法により，$\sigma_N^2(\boldsymbol{z}) = \mathrm{Var}(\sum_i \alpha_i(\boldsymbol{z}) \widetilde{\psi}_{i,\tau(\boldsymbol{z})}) / \mathrm{E}[U(T, \boldsymbol{Z}, W, x; e, \lambda_w^C, S_w^C | \boldsymbol{Z} = \boldsymbol{z})]^2$ の形式で与えられる．$\mathrm{E}[U(T, \boldsymbol{Z}, W, x; e, \lambda_w^C, S_w^C) | \boldsymbol{Z} = \boldsymbol{z}]$ はフォレストに基づく経験推定値で推定でき，$\mathrm{Var}(\sum_i \alpha_i(\boldsymbol{z}) \widetilde{\psi}_{i,\tau(\boldsymbol{z})})$ はスモールバックのブートストラップ法（Athey et al., 2019）により推定可能である．なお，スモールバックのブートストラップ法のサンプリング計画や一致性についての詳細は，R パッケージ `grf` の実装コードや Athey et al. (2019)を参照されたい．現時点で因果生存フォレストを応用した研究（例えば，Desai et al. (2024)では駆出率保持心不全に対するスピロノラクトンの有効性評価のために実施された RCT データに因果生存フォレストを適用し，治療効果の異質性に影響を与える要因を検討している）は限られているが，因果生存フォレストは伝統的なロジスティック回帰や Cox 回帰よりも柔軟に集団レベルの曝露効果を個人レベルに分解し，理論的妥当性をもって治療効果の個人不均一性の特定もしくは仮説を生み出すツールとして有望である．

4. データ適用例

ここでは，データ `peakVO2`(男性)を用いて生存フォレストや因果生存フォレストの適用例を

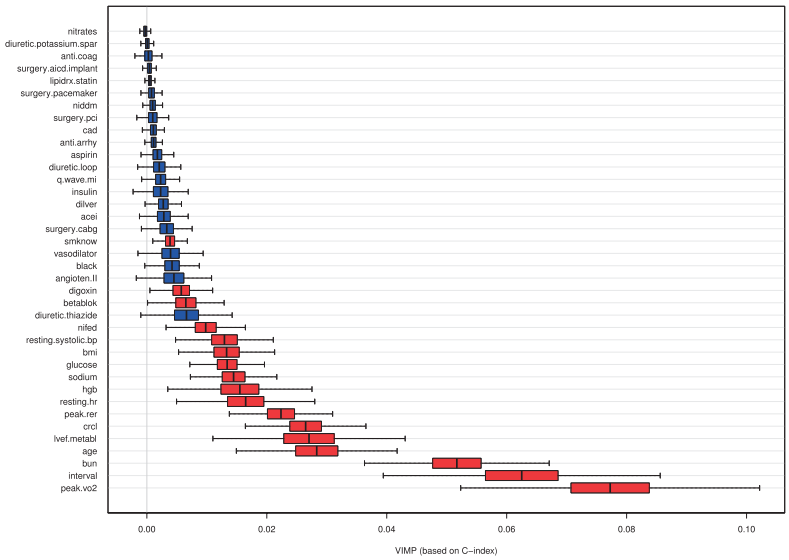

図 4．データ peakVO2(男性)における生存フォレストの変数重要度(VIMP)プロット．

示す．まず，3.1 節で説明した生存フォレストによる VIMP を図 4 に与える．例えば心筋梗塞二次予防において推奨されてきたが近年議論になっている β 遮断薬の投与(betablock)は，Cox 回帰と同様に有意な傾向にある．各 VIMP の近似的 95% 信頼区間は d 例削除ジャックナイフ法（Ishwaran and Lu, 2018)により作成され，この方法の漸近理論の確立は依然として課題であるがシミュレーション研究では良好な性能が示されている．なお，図 4 は R ライブラリの randomForestSRC（Ishwaran et al., 2021)の次のコマンドにより描画した：

```
set.seed(124)
rf <- rfsrc(Surv(ttodead, died)~., data=peakVO2m)
rf.smp.o <- subsample(rf, B=25, subratio=0.5)
plot.subsample(rf.smp.o, xlab="VIMP",sort=T)
```

Hsich et al.（2011)で指摘されたように，選ばれる重要な予後因子は，Cox 回帰と生存フォレストで類似するが，重要度の順位にはいくらかの違いが生じる．そのため，VIMP は，実際の研究では，変数選択の目的や重要な予後因子の同定のために Cox 回帰の代替もしくは補完として有用である（例えば，He et al., 2020; Liu et al., 2021 など)一方，ハザード比のような直接的な解釈ができないため，集団疫学的な視点からは活用しにくい側面があることには留意したい．

次に，処理変数ではないが例示のため，最高酸素摂取量(peak.vo2)を興味ある変数と仮定して生存フォレストと因果生存フォレストの結果を比較する．図 5 の左図は生存フォレストから予測された生存関数に対して，最高酸素摂取量を 12 から 30 まで変化させた部分従属プロットである．この部分従属プロットは(2.1)の推定生存関数ではなく，個々の生存木 $\mathcal{T}^{(b)}$ から推定された生存関数 $\hat{S}^{(b)}(t|z)$ に，特定の変数 Z_j の値(ここでは peak.vo2)を c におき換えた $\mathcal{T}^{(b)}$ の OOB データをあてはめて，フォレストの上で平均化した

$$\hat{S}_j^{\mathrm{OB}}(t|c) = \frac{1}{N}\sum_{i\in\mathcal{L}_N}\frac{1}{|O_i|}\sum_{b\in O_i}\hat{S}^{(b)}(t|\boldsymbol{Z}_i, Z_j = c)$$

により与えられる（Ishwaran and Lu, 2019）．ここに，O_i は個人 i が OOB であることを記録した生存木の番号である．図 5 の中央は生存関数の部分従属プロットを 5 年生存率で表現したもので peak.vo2 が 20 までは緩やかに増加し，20 以降ではあまり変化がないことがわかる．図 5 の右図は，処理変数に $W_i^{(c)} = \mathbb{1}(\texttt{peak.vo2}_i < c)$ を設定して，因果生存フォレストを適用して得た結果を，c を変化させて描いたものである．丸点は，各 c（x 軸）の値において，条件付き処理効果（CATE）$\tau_c(\boldsymbol{z}_{md}) = \mathrm{E}[\mathbb{1}(T_i^{*(1)} > 5) - \mathbb{1}(T_i^{*(0)} > 5) | \boldsymbol{Z}_i = \boldsymbol{z}_{md}]$ に対する 3.3 節で与えた推定値 $\hat{\tau}(\boldsymbol{z}_{md})$ であり，$\boldsymbol{Z}_i = \boldsymbol{z}_{md}$ は（W_i 以外の）すべての共変量を中央値 $\boldsymbol{z}_{md}$ に設定することを意味する（実線の曲線はこれらの点の Lowess あてはめ，破線の Lowess 曲線は対応する 95% 各点信頼区間を示す）．図 5 の右図の三角点は平均処理効果（ATE）$\tau = \mathrm{E}[\mathbb{1}(T_i^{*(1)} > 5) - \mathbb{1}(T_i^{*(0)} > 5)]$ の推定値を表す．ATE の推定量として，Cui et al.（2023）では，Robins et al.（1994）による拡張逆確率重み付け推定量（augmented inverse propensity）を若干変形した 2 重ロバスト・スコア

$$\hat{\Gamma}_i = \hat{\tau}(\boldsymbol{Z}_i) + \frac{\tilde{\psi}_{\hat{\tau}(\boldsymbol{Z}_i)}(T_i, \Delta_i, \boldsymbol{Z}_i, W_i, x; \hat{e}, \hat{m}, \hat{\lambda}_w^C, \hat{S}_w^C, \hat{Q}_w)}{\hat{e}(\boldsymbol{Z}_i)(1 - \hat{e}(\boldsymbol{Z}_i))} \tag{4.1}$$

に基づく平均の使用を示唆している（Cui et al.（2023）では CATE の議論に集中し，ATE に対する言及はないが，Athey and Wager（2019）では（4.1）に類似する形式を用いて ATE 推定量を提案している）．ここで，$\tilde{\psi}_\tau(\cdot)$ は（3.6）で定義されたものであり，$\hat{\tau}(\boldsymbol{Z}_i)$ は因果生存フォレストのあてはめから得られる個人 i の共変量 $\boldsymbol{Z}_i$ に対する CATE 推定値である．Cui et al.（2023）では，CATE の異質性の解釈をなすために，関心のある特徴量から $\{\hat{\Gamma}_i\}_{i=1}^N$ への回帰を行うことも与えている．なお，図 5 の右図のための計算は，R ライブラリの grf（Tibshirani et al., 2024）の causal_survival_forest 関数を用いた以下のコマンドで求めることができる（この例は処理変数 $W_i = \mathbb{1}(\texttt{peak.vo2}< c)$ の設定が $c = 20$ の場合．5 年生存率の推測のため，horizon=5, target="survival.probability"に設定している）：

```
> Y <- peakVO2m$ttodead; D <- peakVO2m$died; Z <- peakVO2m[,-c(30,32,33)]
> W <- ifelse(peakVO2m$peak.vo2 < 20,0,1)
> csf <- causal_survival_forest(Z,Y,W,D, horizon=5,target="survival.probability")
> Zmed<- matrix(apply(Z,2,median),nrow=1)
> predict(csf, newdata=Zmed, estimate.variance=TRUE) # Zmed に対する CATE
  predictions variance.estimates
  0.1839267   0.00386204
> average_treatment_effect(csf) # ATE
  estimate     std.err
  0.07390137   0.01650680
```

図 5 より，5 年生存率の CATE や ATE の推定値（右図）は，最高酸素摂取量（peak.vo2）の変化で平均的に 10% 程度変化すること（中央）を良好に捉えていることが確認できる．最高酸素摂取量（peak.vo2）は，Cox 回帰や生存フォレストにおいて予後への強い統計的有意性を示したが，生存率を目的とした因果フォレストの 95% 各点信頼区間の下限がゼロ付近に留まっていることから，因果フォレストの統計的検出力は低下していることが視察できる．単純なシミュレーション実験から，生存率を目的とした因果フォレストの検出力は，x（horizon）の設定値により大きく変化することに注意したい（Web Supplementary Materials B を参照）．生存率を目的とする因果生存フォレストは，生存期間全体を用いる生存フォレストや Cox 回帰と比べ，5 年生存率という 1 点だけの情報に縮約される点も検出力で不利に働く．なお，Athey and Wager（2019）では，検討する変数が多い場合，最初の因果フォレストで推定された変数重要度の低い

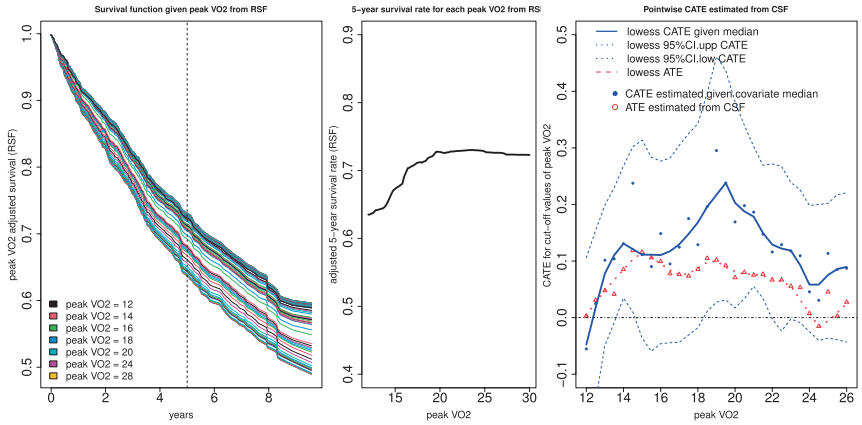

図 5. データ `peakVO2`(男性)における生存フォレスト(RSF)からの `peak.vo2` に対する部分従属プロット(左)，部分従属プロットからの `peak.vo2` による 5 年生存率の推移(中央)，`peak.vo2` のカットオフ値に対する因果生存フォレスト(CSF)からの CATE と ATE の推定結果(左)．

変数を削除してから，改めて因果フォレストをあてはめ直して，CATE や ATE を推定することを提示している(本節ではこのステップは省略した)．正直さ(honesty)の実現のために行うサンプル分割は検出力を下げるかもしれない．因果生存フォレストに有利な状況であれば，ダブルサンプルに基づく検出力低下は，実用的観点ではそれほど問題にならない可能性もある．いずれにしても予測指標や変数重要度指標の表示だけでなく，因果フォレストによって実データから複雑な効果の異質性について有意義な結果を導くための統計的推測の実施のための基本が与えられた意義は極めて大きい．Web Supplementary Materials B では，因果生存フォレストの統計的性能のさらなる数値的検討のため，シミュレーション実験の結果を与えている．

5. 結びに代えて

本論文では，機械学習・統計解析ツールとして，広く応用されるようになっている決定木やフォレスト法などの樹木法における生存データへの適用と最近の因果フォレストへの展開を報告した．近年，様々な機械学習法を因果推論と融合させる方法論の研究が活性化している (Xu et al., 2023; Wan et al., 2024)．因果フォレスト法 (Wager and Athey, 2018)の顕著な点は，難解であったランダムフォレスト法の統計的漸近理論の構築を，正直さ(honesty)などの着想を用いて与えたことである．今後にかけて，関係する理論のさらなる深耕，方法の発展が期待される．因果フォレスト法の適用可能なデータの範囲は徐々に拡大し，Athey and Wager (2019) ではクラスターデータに対する因果フォレストを議論し，Athey et al. (2019)では局所的推定方程式のアプローチによって因果フォレストの適用範囲を広げる一般化ランダムフォレストを提案し，因果生存フォレストの展開が生み出されている．現状の因果生存フォレストでは，CATE，ATE の推測において，RMST と各点生存率の利用に留まっているが，ハザード比や全体的な生存関数への適用可能にすることも興味のある展開の一つである．ダブルサンプルに基づく検出力の低下は予想外に小さい可能性があるがセミパラメトリック有効性にどの程度近づけられるかの検討は興味ある課題である．因果生存フォレスト法の公式的利用は，現時点では，2 値処理変数の枠組みに限定されているが，通常の因果フォレストの場合には連続処理変数への拡張も `grf` に既に実装されている．このような展開から，依然として十分でないフォレ

スト法からの変数効果を調べるための有力なアプローチとしてさらなる研究の進展が期待される．樹木構造アプローチやランダムフォレストはこれまでも統計利用やデータ科学において有用なツールを与えてきたが，今後にかけても，有益な理論と方法論の研究が進展し，有意義な応用が展開されることが期待される．

参 考 文 献

Arano, I., Sugimoto, T., Hamasaki, T. and Ohno, Y. (2010). Practical application of cure mixture model to long-term censored survivor data from a withdrawal clinical trial of patients with major depressive disorder, *BMC Medical Research Methodology*, **10**, https://doi.org/10.1186/1471-2288-10-33.

Athey S. and Wager S. (2019). Estimating treatment effects with causal forests: An application, *Observational Studies*, **5**, 36–51.

Athey, S., Tibshirani, J. and Wager, S. (2019). Generalized random forests, *The Annals of Statistics*, **47**, 1148–1178.

Bou-Hamad, I., Larocque, D. and Ben-Ameur, H. (2011). A review of survival trees, *Statistics Surveys*, **5**, 44–71.

Breiman, L. (1996). Bagging predictors, *Machine Learning*, **24**, 123–140.

Breiman, L. (2001). Random forests, *Machine Learning*, **45**, 5–32.

Breiman, L. (2003). How to use survival forests, Department of Statistics, University of California, Berkeley, California, http://www.stat.berkeley.edu/~breiman/SF_Manual.pdf（最終アクセス日 2025 年 3 月 18 日）.

Breiman, L. (2004). Consistency for a simple model of random forests, Technical Report, **670**, University of California, Berkeley, California, https://www.stat.berkeley.edu/~breiman/RandomForests/consistencyRFA.pdf（最終アクセス日 2025 年 3 月 18 日）.

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees*, Chapman and Hall, Wadsworth, New York.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters, *The Econometrics Journal*, **21**, 1–68.

Ciampi, A., Thiffault, J., Nakache, J.-P. and Asselain, B. (1986). Stratification by stepwise regression, correspondance analysis and recursive partition: A comparison of three methods of analysis for survival data with covariates, *Computational Statistics & Data Analysis*, **4**, 185–204.

Cui, Y., Kosorok, M. R., Sverdrup, E., Wager, S. and Ruoqing, Z. (2023). Estimating heterogeneous treatment effects with right-censored data via causal survival forests, *Journal of the Royal Statistical Society Series B*, **85**, 179–211.

Davis, R. B. and Anderson, J. R. (1989). Exponential survival trees, *Statistics in Medicine*, **8**, 947–961.

Desai, R. J., Glynn, R. J., Solomon, S. D., Claggett, B., Wang, S. V. and Vaduganathan, M. (2024). Individualized treatment effect prediction with machine learning-salient considerations, *New England Journal of Medicine Evidence*, **3**, https://doi.org/10.1056/EVIDoa2300041.

Efron, B. (2014). Estimation and accuracy after model selection (with discussion), *Journal of the American Statistical Association*, **109**, 991–1007.

江村剛志 (2023). 安定化スコア検定を用いた高次元生存データに基づく決定木の構築法, 日本統計学会誌, **52**, 373–390.

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*, John Wiley & Sons, New York.

Gordon, L. and Olshen, R. A. (1985). Tree-structured survival analysis, *Cancer Treatment Reports*, **69**, 1065–1069.

Graf, E., Schmoor, C., Sauerbrei, W. and Schumacher, M. (1999). Assessment and comparison of prognostic classification schemes for survival data, *Statistics in Medicine*, **18**, 2529–2545.

Harrell, F., Califf, R., Pryor, D., Lee, K. and Rosati, R. (1982). Evaluating the yield of medical tests, *The Journal of the American Medical Association*, **247**, 2543–2546.

He, J., Zhang, J. X., Chen, C. T., Ma, Y., De Guzman, R., Meng, J. and Pu, Y. (2020). The relative importance of clinical and socio-demographic variables in prognostic prediction in non-small cell lung cancer: A variable importance approach, *Medical Care*, **58**, 461–467.

Heagerty, P. J. and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves, *Biometrics*, **61**, 92–105.

Hothorn, T., Lausen, B., Benner, A. and Radespiel-Troger, M. (2004). Bagging survival trees, *Statistics in Medicine*, **23**, 77–91.

Hothorn, T., Buhlmann, P., Dudoit, S., Molinaro, A. and van der Laan, M. J. (2006a). Survival ensembles, *Biostatistics*, **7**, 355–373.

Hothorn, T., Hornik, K. and Zeileis, A. (2006b). Unbiased recursive partitioning: A conditional inference framework, *Journal of Computational and Graphical Statistics*, **15**, 651–674.

Hsich, E., Gorodeski, E. Z., Blackstone, E. H., Ishwaran, H. and Lauer, M. S. (2011). Identifying important risk factors for survival in patient with systolic heart failure using random survival forests, *Circulation*: *Cardiovascular Quality and Outcomes*, **4**, 39–45.

Ishwaran, H. and Lu, M. (2018). Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival, *Statistics in Medicine*, **38**, 558–582.

Ishwaran, H. and Lu, M. (2019). Random survival forests, *Wiley StatsRwf*: *Statistics Reference Online*, 1–13, John Wiley & Sons, https://doi.org/10.1002/9781118445112.stat08188.

Ishwaran, H., Kogalur, U. B., Blackstone, E. H. and Lauer, M. S. (2008). Random survival forests, *Annals of Applied Statistics*, **2**, 841–860.

Ishwaran, H., Lu, M. and Kogalur, U. B. (2021). Variable importance (VIMP) with subsampling inference, http://randomforestsrc.org/articles/vimp.html（最終アクセス日 2025 年 3 月 18 日）.

LeBlanc, M. and Crowley, J. (1992). Relative risk trees for censored survival data, *Biometrics*, **48**, 411–425.

LeBlanc, M. and Crowley, J. (1993). Survival trees by goodness of split, *Journal of the American Statistical Association*, **88**, 457–467.

LeBlanc, M. and Crowley, J. (1995). A review of tree-based prognostic models, *Journal of Cancer Treatment and Research*, **75**, 113–124.

Liao, T., Su, T., Lu, Y., Huang, L., Wei, W. Y. and Feng, L. H. (2024). Random survival forest algorithm for risk stratification and survival prediction in gastric neuroendocrine neoplasms, *Scientific Reports*, **14**, https://doi.org/10.1038/s41598-024-77988-1.

Liu, B., Niu, L., Boscoe, F. and Lee, F. F. (2021). Predictors of survival among male and female patients with malignant pleural mesothelioma: A random survival forest analysis of data from the 2000-2017 surveillance, epidemiology, and end results program, *Journal of Registry Management*, **48**, 118–125.

Loh, W. Y. (2002). Regression trees with unbiased variable selection and interaction detection, *Statistica Sinica*, **12**, 361–386.

Mogensen, U., Ishwaran, H. and Gerds, T. (2012). Evaluating random forests for survival analysis using prediction error curves, *Journal of Statistical Software*, **50**, 1–23.

Morgan, J. N. and Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal, *Journal of the American Statistical Association*, **58**, 415–434.

Morvan, L., Carlier, T., Jamet, B., Bailly, C., Bodet-Milin, C., Moreau, P., Kraeber-Bodéré, F. and Mateus, D. (2020). Leveraging RSF and PET images for prognosis of multiple myeloma at diagnosis, *International Journal of Computer assisted Radiology and Surgery*, **15**, 129–139.

中村知繁 (2020). ランダムフォレストによる因果推論, 慶應義塾大学経済研究所, https://ies.keio.ac.jp/

upload/20201201econo_nakamura_Slide.pdf（最終アクセス日 2025 年 3 月 18 日）.

Probst, P., Boulesteix, A.-L. and Bischl, B. (2019). Tunability: importance of hyperparameters of machine learning algorithms. *Journal of Machine Learning Research*, **20**, 1–32.

Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed, *Journal of the American Statistical Association*, **89**, 846–866.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression, *Econometrica*, **56**, 931–954.

Schick, A. (1986). On asymptotically efficient estimation in semiparametric models, *The Annals of Statistics*, **14**, 1139–1151.

Scornet, E., Biau, G. and Vert, J.-P. (2015). Consistency of random forests, *The Annals of Statistics*, **43**, 1716–1741.

Segal, M. R. (1988). Regression trees for censored data, *Biometrics*, **44**, 35–48.

Segal, M. R. (1992). Tree-structured methods for longitudinal data, *Journal of the American Statistical Association*, **87**, 407–418.

Segal, M. R. (2004). Machine learning benchmarks and random forest regression, Center for Bioinformatics & Molecular Biostatistics, University of California, San Francisco, California, https://escholarship.org/uc/item/35x3v9t4（最終アクセス日 2025 年 3 月 18 日）.

下川敏雄, 杉本知之, 後藤昌司 (2013).『樹木構造接近法』, 共立出版, 東京.

杉本知之, 下川敏雄, 後藤昌司 (2005). 樹木構造接近法と最近の発展, 計算機統計学, **18**, 123–164.

Therneau, T., Grambsch, P. and Fleming, T. (1990). Martingale-based residuals for survival models, *Biometrika*, **77**, 147–160.

Tibshirani, J., Athey, S., Sverdrup, E. and Wager, S. (2024). grf: Generalized random forests, https://github.com/grf-labs/grf（最終アクセス日 2025 年 3 月 18 日）.

Wager, S. (2016). Causal inference with random forests, Ph.D. Thesis, Stanford University, Stanford, California.

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

Wager, S., Hastie, T. and Efron, B. (2014). Confidence intervals for random forests: the jackknife and the infinitesimal jackknife, *The Journal of Machine Learning Research*, **15**, 1625–1651.

Wan, K., Tanioka, K. and Shimokawa, T. (2024). Survival causal rule ensemble method considering the main effect for estimating heterogeneous treatment effects, *Statistics in Medicine*, **43**, 5234–5271.

Williamson, B. D., Gilbert, P. B., Carone, M. and Simon, N. (2021). Nonparametric variable importance assessment using machine learning techniques, *Biometrics*, **77**, 9–22.

Xu, Y., Ignatiadis, N., Sverdrup, E., Fleming, S., Wager, S. and Shah, N. H. (2023). Treatment heterogeneity for survival outcomes, *Handbook of Matching and Weighting Adjustments for Causal Inference* (eds. J. R. Zubizarreta, E. A. Stuart, D. S. Small and P. R. Rosenbaum), 445–482, Chapman and Hall/CRC, New York.

Zhou, Y. and McArdle, J. J. (2015). Rationale and applications of survival tree and survival ensemble methods, *Psychometrika*, **80**, 811–833.

352 Proceedings of the Institute of Statistical Mathematics Vol. 73, No. 2, 335–352 (2025)

Tree-structured Approaches and Recent Advances in Survival Analysis

Tomoyuki Sugimoto[1], Kazushi Maruo[2] and Toshio Shimokawa[3]

[1]Graduate School of Engineering Science, The University of Osaka
[2]Institute of Medicine, Tsukuba University
[3]Graduate School of Medicine, Wakayama Medical University

Tree-based methods such as decision trees and random forests are statistical and machine learning tools widely used in data science and many other fields. Over the past decade, there has been remarkable progress in methodologies and theories related to causal inference tools based on causal trees and causal forests, leveraging variable importance measures, the consistency of forests, and asymptotic normality. This paper investigates the recent developments in methodologies and theories of such tree-based methods, focusing on their application in survival analysis, and comprehensively reports on the characteristics of each method.

Key words: CART, causal forest, Cox regression, random forest, survival tee, variable importance.