
Article:
Glynn, David orcid.org/0000-0002-0989-1984, Giardina, John, Hatamyar, Julia Farideh et al. (3 more authors) (2024) Integrating decision modelling and machine learning to inform treatment stratification. Health Economics. ISSN 1057-9230

https://doi.org/10.1002/hec.4834

Received: 6 March 2023 | Revised: 18 March 2024 | Accepted: 29 March 2024

DOI: 10.1002/hec.4834

RESEARCH ARTICLE

Health Economics WILEY

Integrating decision modeling and machine learning to inform treatment stratification

David Glynn[1] | John Giardina[2] | Julia Hatamyar[1] | Ankur Pandya[2] | Marta Soares[1] | Noemi Kreif[1]

[1]Centre for Health Economics, University of York, York, UK

[2]Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Correspondence
David Glynn, Centre for Health Economics, University of York, York, UK.
Email: david.glynn@york.ac.uk

Funding information
Medical Research Council, Grant/Award Number: MR/T04487X/1

Abstract

There is increasing interest in moving away from "one size fits all (OSFA)" approaches toward stratifying treatment decisions. Understanding how expected effectiveness and cost-effectiveness varies with patient covariates is a key aspect of stratified decision making. Recently proposed machine learning (ML) methods can learn heterogeneity in outcomes without pre-specifying subgroups or functional forms, enabling the construction of decision rules ('policies') that map individual covariates into a treatment decision. However, these methods do not yet integrate ML estimates into a decision modeling framework in order to reflect long-term policy-relevant outcomes and synthesize information from multiple sources. In this paper, we propose a method to integrate ML and decision modeling, when individual patient data is available to estimate treatment-specific survival time. We also propose a novel implementation of policy tree algorithms to define subgroups using decision model output. We demonstrate these methods using the SPRINT (Systolic Blood Pressure Intervention Trial), comparing outcomes for "standard" and "intensive" blood pressure targets. We find that including ML into a decision model can impact the estimate of incremental net health benefit (INHB) for OSFA policies. We also find evidence that stratifying treatment using subgroups defined by a tree-based algorithm can increase the estimates of the INHB.

KEYWORDS
causal inference, decision modeling, heterogeneity, machine learning, microsimulation, optimal policy, precision medicine

1 | INTRODUCTION

An overarching goal throughout health care decision making is to give the right treatment to the right individual at the right time. Health decision makers, however, often rely on evidence about whether a policy improves health outcomes *on average* for a given population, leading to "one size fits all" treatment decisions that can mask important heterogeneity in the effectiveness and cost-effectiveness of treatments.

Health decision analysis, including cost-effectiveness analysis, aims to quantitatively compare different treatment options to identify the optimal choice. Subgroup analysis techniques have been used to characterize heterogeneity in decision analytic models so that recommendations can be targeted toward specific groups of patients and more individuals receive their particular optimal treatment or intervention. The approach to subgroup analyses can range from a simple analysis in which a population is split into two groups, to complex methods that aim to stratify decision making by multiple groups based on patient characteristics (Basu & Meltzer, 2007; Coyle et al., 2003; Espinoza et al., 2014).

The key issue when conducting these subgroup analyses is to find ways to use the available data to appropriately parameterize the decision analytic model for each subgroup. There is an increasing focus on parameterizing decision models directly with individual patient data (IPD)—in these cases, the number of subgroups that can be parameterized in a model is limited by the size of the data set and distribution of patient characteristics across the population.

Machine learning (ML) techniques can be used to find the patient characteristics that best predict risk and variation in treatment effects, and have been proposed for use in health decision analysis and economic evaluation (Bonander & Svensson, 2021; Padula et al., 2022; Sadique et al., 2022; Xu et al., 2022). There has been limited discussion, however, of the best approach to integrate ML-derived parameter estimates into decision models. This paper demonstrates how to integrate ML-derived parameter estimates into a decision analytic microsimulation model, focusing on integrating heterogeneous treatment effect (HTE) estimates derived using ML and causal inference methods, and shows how this approach to parameterization can be used to stratify decisions across subgroups. A framework for conducting these analyses is illustrated using a case study on personalizing hypertension treatment decisions.

2 | BACKGROUND

2.1 | Incorporating heterogeneity and machine learning into decision analysis

There has been extensive work discussing the integration of heterogeneity into decision analytic models. The "value of heterogeneity (VoH)" framework has been proposed to describe the marginal payoff from subdividing a population into an increasing number of subgroups while also taking account of how uncertainty changes as subgroups are added (Espinoza et al., 2014; Pataky, Bryan, Sadatsafavi, Peacock, & Regier, 2022). The "expected value of individualized care (EVIC)" has also been defined as the additional value gained from assigning each individual in a population their optimal treatment, based on their observed characteristics, compared to assigning all individuals to the same treatment option (Basu & Meltzer, 2007).

In addition to these general frameworks for incorporating heterogeneity, there have recently been more specific discussions about how to incorporate different types of ML methods into decision-making. These approaches often promise an increased ability to stratify and target decisions based on observable patient characteristics. Padula et al. discuss the use of ML for various areas of health economics and outcomes research, highlighting the difference between using ML for predictive analytics (e.g., constructing prognostic risk scores) and causal inference (Padula et al., 2022).

Predictive ML has an important place in the personalization of health interventions for example, using ML based screening algorithms to identify patients with a high risk of developing a disease (Hill et al., 2020). However, predictive ML cannot on its own derive the causal effect of a health policy, which is often the key parameter when deciding between different treatment options. More recently, a combination of ML and causal inference has been used to estimate HTEs for policy relevant outcomes using IPD. Sadique et al. estimated individual and subgroup treatments effects of a trial conducted among critically ill patients, on 90 days mortality (Sadique et al., 2022). Kreif and colleagues estimated HTEs on infant mortality and maternal health care utilization outcomes in order to evaluate the benefits of extending health insurance in Indonesia (Kreif et al., 2021).

These ML-based methods to estimate HTEs can be contrasted with traditional subgroup analyses. Subgroup analyses are often pre-specified to avoid "cherry picking" statistically significant results, but this approach does not offer flexibility in discovering drivers of heterogeneity that were not posited a priori in the original analysis plan

(Wang et al., 2007). In contrast, recent approaches in ML and causal inference allow for capturing heterogeneity in predicted outcomes by modeling them as a flexible function of patient baseline covariates (X). The appeal of ML approaches is their ability to adaptively capture unknown functional form relationships (e.g., how the covariates and the treatment relate to the outcome), and use techniques to effectively select among a potentially large number of covariates (Hahn et al., 2020).

To translate these estimated HTEs into decisions, the literature on "optimal treatment regimes" formalizes treatment allocation by defining policy assignment rules that map X into a treatment decision, where optimality is understood with respect to a pre-specified objective function of the health decision maker (e.g., minimizing a health measure such as mortality), subject to constraints (e.g., size of the available budget). See Murphy et al. and Manski et al. for the initial theory, and Luedtke and van der Laan and Athey and Wager for recent proposals (Athey & Wager, 2021; Luedtke & van der Laan, 2016; Manski, 2004; Murphy, 2003).

One simple policy rule is the so-called "plug-in policy" or "thresholding rule" that assigns treatment to patients according to the sign of the estimated HTE in each patient (e.g., positive in terms of expected survival). However, such plug-in rules might appear to a policymaker as a black box, and calculating them for a new patient may be beyond the expertize of health decision makers. Policy rules that follow a policy tree of fixed depth provide an alternative with straightforward interpretation (Amram et al., 2022; Athey & Wager, 2021; Bertsimas et al., 2019).

In the context of cost-effectiveness analysis, methods have been developed to estimate heterogeneity in more comprehensive outcomes such as incremental net benefit, a measure which combines health and cost outcomes into a single outcome. Bonander and Svensson apply causal forests to IPD to flexibly model cost, quality adjusted life years (QALYs) and net benefit outcomes (Bonander & Svensson, 2021). This approach only considered the information contained in the IPD and so we refer to it as a "pure IPD" approach. This is equivalent to a "within trial economic evaluation", which as described by Sculpher et al. has a number of well recognized limitations for medical decision making: the dataset must include all relevant interventions; trial follow up must be sufficient to capture all relevant differences between treatments; and the trial must include all relevant patient groups (Sculpher et al., 2006). These IPD requirements may be so high as to be impractical and rarely applicable to policy decisions. In a clinical trial case study Xu et al. address the trial follow up limitation of IPD by using a decision model to predict long term outcomes for the IPD (Xu et al., 2022). We classify this as an "augmented IPD" method which, in contrast to the "pure IPD" approach, uses disease simulation modeling methods to augment IPD.

The decision modeling and evidence synthesis literature provides a suite of methods to address the limitations of analysis based on a single IPD (Sculpher et al., 2006). In this paper we demonstrate a novel "full decision modeling approach" which addresses the limitations of IPD based approaches by combining the strengths of ML and decision modeling. We use ML to model heterogeneity in decision model parameters in addition to quantifying uncertainty. We then show how to compare the payoffs with different policy rules.

3 | METHODS

3.1 | Analytic overview

We propose a three-step workflow to integrate ML and decision modeling and use this approach to stratify decision making. First, we use ML to model heterogeneity in key decision model inputs (e.g., risk of primary outcome). Second, we use these heterogeneous inputs in a decision model to generate individual specific long-term payoffs for example, incremental net health benefit (INHB). Third, using these individual specific payoffs, we evaluate the expected benefit of various policy allocation rules in a VoH framework. The three policy rules are (1) one size fits all policies where everyone in the population gets the same treatment decision; (2) a plug-in policy in which treatment choice is based on a complex function of all patient characteristics and; (3) subgroup based decision making in which a finite number of subgroups is learned using policy tree algorithms. Below we first discuss some general points about integrating decision modeling and ML, we then describe in detail the three-step workflow summarized in Figure 1.

3.2 | Integrating machine learning and decision modeling

Decision models simulate outcomes for individuals or cohorts of patients, taking account of the differences in outcomes with different baseline characteristics for example, age, sex, disease history etc. The set of characteristics associated with

1. **Model heterogeneity in a time-to-event endpoint using Bayesian modelling**
   a. For each trial arm, choose AFT BART parameters (ζ, ξ and ϕ) via k-fold cross-validation (where k = 10):
   b. Predict counterfactual survival curves for all individuals using cross fitting (Figure 2).
2. **Incorporate counterfactual time-to-event estimates into a decision model**
   a. Convert the counterfactual survival curve predictions to transition probabilities and insert the transition probabilities into the decision model.
   b. Generate individual specific predictions of long-term payoffs ($INHB_{xi}$).
3. **Define treatment allocation rules and estimate payoffs**
   a. 'One size fits all' policy (π^{OSFA}): all individuals get the same treatment decision:
      i. Policy assignment: allocate all patients to the treatment with the highest expected $INHB_{xi}$ over the population.
      ii. Policy payoff: calculate out of fold payoff for tree with depth d, using k-fold cross validation (where k = 5).
   b. 'Subgroup' policy ($\pi^{T}(x_i)$): treatment choice is based on subgroups that are learned using optimal policy algorithms.
      i. Policy assignment: using the policy tree algorithm, find the tree (with depth d) which maximises the average of the $\hat{\pi}^{Td}(xi).INHB_{xi}$ of those individuals that receive treatment according to the policy assignment.
      ii. Policy payoff: calculate out of fold payoff for tree with depth d, using k-fold cross validation (where k = 5).
   c. 'Plug in' policy ($\pi^{PI}(x_i)$): treatment choice is based directly on decision model prediction.
      i. Policy assignment: allocate patients to the new treatment if they have a positive $INHB_{xi}$.
      ii. Policy payoff: average $INHB_{xi}$ according to the policy assignment.

FIGURE 1 Algorithm summarizing the three steps taken in the analysis. AFT BART, accelerated failure time Bayesian additive regression tree; $INHB_{x_i}$, incremental net health benefit for patient x_i; OSFA, one size fits all.

patient i is referred to as their profile ($x_i \in X$). Predictions are based on a set of input parameters θ, these parameters are associated with parametric uncertainty, the joint distribution of uncertainty in all model inputs is represented by $p(\theta) = \Theta$ (Baio, 2013; Bojke et al., 2022; Briggs et al., 2006; O'Hagan et al., 2006). ML can be used to flexibly model heterogeneity in these input parameters. Allowing input parameters to flexibly depend on x_i will naturally impact the heterogeneity in outcome predictions across patient profiles.

We identified two key issues to consider when integrating ML-based estimates into a decision model. First, decision models usually include a focus on uncertainty quantification (e.g., through probabilistic sensitivity analyses), so it is often necessary to know the uncertainty around each parameter in the model and draw from an underlying distribution. Many ML methods, however, are developed to optimize prediction success and are less focused on uncertainty quantification (Athey & Imbens, 2019). To accommodate this in our case study, we used the Bayesian additive regression tree (BART) method to estimate the key HTE parameters in the model. BART provides posterior distributions of the causal effect of the intervention for each individual in the patient population, based on their observed covariates (J. Hill, Linero, & Murray, 2020).

The second issue we encountered was that it was unclear how to incorporate concerns about overfitting, which is a major focus of the implementation of ML-based methods, into the development of a decision analytic model. To solve this issue in our case study, we made sure that, for each individual in the model, the parameters used to predict survival for that individual came from models that were trained on data that did not include that particular individual. The logic we use is in line with state-of-the-art causal ML methods that aim to estimate conditional average treatment effects (CATEs). For example, the "honest" causal forest approach by Athey et al. produces so-called "out-of-bag" estimates of the CATE (Athey et al., 2019). A similar sample splitting approach is taken in the double-robust learner proposed by Kennedy (2020). We propose methods to address these two issues. Note that the methods described are specific to our case study, but the principles underlying the methods are generally applicable.

3.3 | Step 1: Model heterogeneity in a time-to-event endpoint using Bayesian modeling

Step 1 in our workflow is to model treatment specific heterogeneity in a time-to-event (or survival) outcome using ML. We chose a time-to-event outcome as these are commonly estimated in decision modeling. The method chosen is an accelerated failure time (AFT) Bayesian additive regression tree (BART). In this section we describe the AFT-BART in general, in the following section we describe how to use it to model treatment effects. In principle the methods described here could be generalized to binary and continuous outcomes.

While there has been a recent explosion of ML methods that estimate heterogeneous parameters for survival data (see Hu et al., Curth et al. and Xu et al., for reviews and comparative simulation studies), the estimation of uncertainty

derived from most of these methods is not well understood (Curth et al., 2021; Hu et al., 2021; Xu et al., 2022). This is a general phenomenon, as many ML methods are developed to optimize prediction success and are less focused on uncertainty quantification (Athey & Imbens, 2019). Below we describe methods to address this.

Consider an AFT approach which models log survival time ($\log T$) as an additive function of two components (1) a flexible function $f(x)$, based on patient characteristics and (2) a residual component W.

$$\log T = f(x) + \mathrm{W}.$$

In order to quantify the input uncertainty needed to integrate ML parameter estimates into decision modeling we chose to utilize a flexible Bayesian ML implementation of survival modeling, the BART method (Henderson et al., 2020; J. Hill, Linero, & Murray, 2020; Hu et al., 2021). This approach nests a random forest within a Bayesian estimation procedure, hence naturally providing uncertainty quantification required to specify Θ (Baio, 2013; Briggs et al., 2006; J. Hill, Linero, & Murray, 2020).

The BART is a tree-based method in which decision trees are used to characterize $f(x)$. Decision trees assign individuals to different predictions by splitting them into subcategories based on cut-off values of their characteristics. Tree depth controls the number of splits per tree. The final subcategories are associated with predictions and these are called the terminal nodes. The BART is an ensemble method in which the prediction for each individual results from a combination of many different decision trees. The BART adaptation to survival analysis described by Henderson imposes statistical priors on the trees and terminal nodes to constrain tree depth and give a high prior probability to the range of survival time seen in the data.

If W is assumed Gaussian, then the BART AFT model will be semi parametric. Following Henderson et al. we instead model W using a non-parametric Dirichlet process mixture model constrained to center on zero (Henderson et al., 2020). Here W is characterized by cluster ($l = 1, ..., L$) specific mixture proportions γ_l, location parameters τ_l and a common scaling parameter σ.

The complexity of the model can be controlled by a set of so-called tuning parameters. Henderson identifies the three most important as: ζ, which controls the prior variance of the residual component; ξ, which determines prior variance of the node values; and φ, for the number of trees. Different combinations imply different degrees and forms of regularization and so can be used to control over fitting. Henderson et al. suggest a k-fold cross-validation grid search approach. K-fold cross validation involves (1) splitting the data into k subsamples (i.e., folds), (2) training the BART with a given set of hyperparameters, (3) calculating absolute log prediction error on the out-of-sample kth fold, (4) repeating this k times and calculate the average error in the out-of-sample folds. A grid-search means we calculate the out of sample error for each set of hyperparameters and choose the set that minimizes the error.

In this model censoring is addressed by imputing unobserved survival times during posterior sampling. Censoring is assumed to be noninformative that is, censoring time and survival time are independent given treatment assignment and patient baseline characteristics (Henderson et al., 2020; Hu et al., 2021).

This non-parametric BART AFT model has been demonstrated to perform well in time-to-event simulation studies (Hu et al., 2021). For technical details see Henderson et al. (2020). The result of fitting the BART model is a set of posteriors for the model parameters. These include estimates of predicted log survival $\widehat{f}(x_i)$, for each patient profile, in addition to the components of the non-parametric residual model: $\widehat{\gamma}_l$, $\widehat{\tau}_l$ and $\widehat{\sigma}$.

The BART AFT can be used for a variety of purposes in decision modeling. For example, we can predict heterogenous survival times for just one treatment group (to characterize heterogeneity in the baseline survival function) and then apply a literature-derived treatment effect parameter to shift these individual-level survival parameters. However, in this paper our main focus is estimating causal heterogenous treatment effect parameters, so we modified this general approach as described below.

3.3.1 | Estimate heterogenous treatment effects as a contrast between predicted survival times

To allow patient characteristics to interact with the treatment in the most flexible way, we follow Hu et al. who propose modeling log T separately for the treatment and control group (similar in spirit to a so-called two-model learner, or 't-learner') (2021). From these separate models, various aspects of the survival distribution can be predicted for each individual, under the treatment and control states (Henderson et al., 2020; Hu et al., 2021). We generate so-called

counterfactual survival probabilities—for each individual and each time period, we predict the survival probability that would be observed if the individual received either the standard care ($a = 0$) or the new treatment ($a = 1$). Contrasting the two counterfactual survival probabilities provides information of the heterogenous treatment effect corresponding to a given patient profile.

In addition to choosing BART parameters using cross-validation, to further protect against overfitting, we use "cross-fitting" to ensure that each survival prediction is out of sample (Chernozhukov, Chetverikov, et al., 2018). This means that we predict the survival of a given individual using models that were fitted without that individual. To do this, we randomly split the sample into two subsamples ($S1$ and $S2$) and stratify by treatment ($a1$) and control ($a0$), for example, $S1_{a1}$ denotes the treated group within subsample $S1$. This results in four strata ($S1_{a0}, S1_{a1}, S2_{a0}, S2_{a1}$) with a separate BART model fit to each. Figure 2 illustrates how cross-fitted counterfactual predictions are produced for all individuals in the sample. As an example, to predict survival with treatment for those individuals in $S1$ who in fact received the control ($S1_{a0}$), we use the model that has been fitted on the individuals in S1 who received treatment ($S1_{a1}$). The end result is a set of posteriors which describe tailored survival probabilities under each treatment condition: $\widehat{f}(a, x_i)$, $\widehat{\gamma_l^a}$, $\widehat{\tau_l^a}$ and $\widehat{\sigma_a}$.

Using these estimates, we then derive individualized survival curves that are specific to patient covariate profile, time and treatment status (Hu et al., 2021):

$$S(t, a, x_i) = 1 - \sum_l \widehat{\gamma_l^a} \Phi\left(\frac{\log t - \widehat{f}(a, x_i) - \widehat{\tau_l^a}}{\widehat{\sigma^a}}\right)$$

3.4 | Step 2: Incorporate counterfactual time-to-event estimates into a decision model

To accommodate the state-based discrete time decision model used in our case study, we discretize the output of a survival analysis to calculate transition probabilities. To implement this we first derive the cumulative hazard function (Briggs et al., 2006) (for simplicity we drop the arguments a and x_i).

$$H(t) = -\ln(S(t))$$

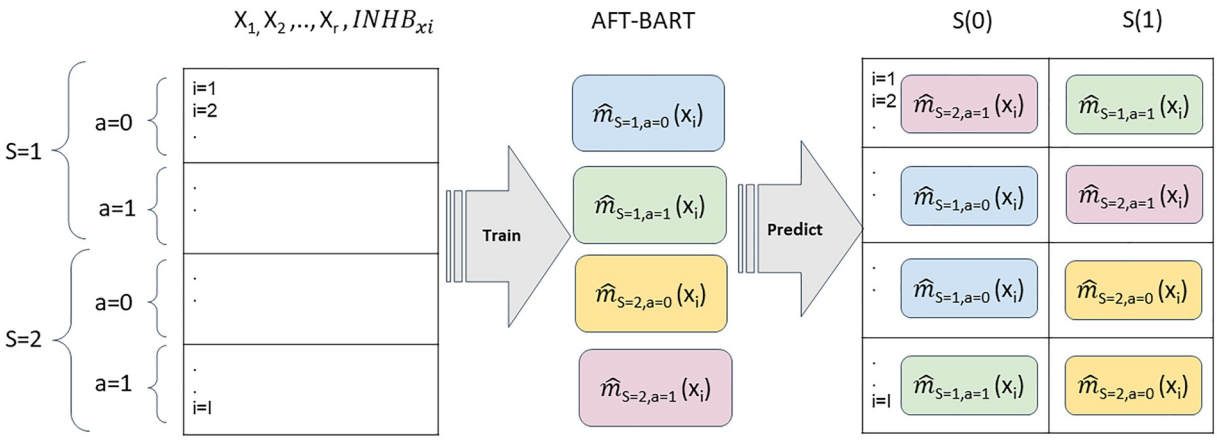

FIGURE 2 Workflow for the cross-fitting approach to predict counterfactual survival curves. The diagram illustrates how AFT-BART models are trained, and how the trained models are used to predict counterfactual survival curves. The dataset is randomly split into two subsamples (S1 and S2), and these subsamples are further stratified by treatment group. For each of the four strata, a separate AFT-BART model is fitted. For example, the model $m_{s=1,a=0}$ has been trained on the control observations within subsample S1. This trained model is then used to generate survival predictions under the control (S(0)) state, for those individuals that have not been used in the training of this model. These are the treated observations in subsample S1, and the control observations in subsample S2. This same logic is followed for the other three trained models, and as a result, two out-of-sample predictions are generated for each individual (i) in the dataset. AFT BART, accelerated failure time Bayesian additive regression tree; $INHB_{x_i}$, incremental net health benefit for patient x_i.

The transition probability for cycle length u is given by (Diaby et al., 2013):

$$tp(t_u) = 1 - \exp(H(t-u) - H(t))$$

Note that because the joint posteriors for $\widehat{f(a,x_i)}$, $\widehat{\gamma_l^a}$, $\widehat{\tau_l^a}$ and $\widehat{\sigma_a}$ embody the sampling uncertainty in these parameters, the survival probabilities and transition probabilities derived from them will embody this joint uncertainty. Incorporating heterogeneous transition probabilities into a decision model will result in heterogeneous estimates of lifetime costs and health outcomes (typically measured in QALYs Briggs et al., 2006).

3.4.1 | Decision making metrics

For decision making within a cost effectiveness analysis framework, the benefits of different treatment options are weighed against the costs they impose (Briggs et al., 2006; Ramponi et al., 2021). The main result of a cost-effectiveness analysis is the incremental cost effectiveness ratio (ICER), which is the ratio of incremental costs (ΔCosts) over incremental QALYs (ΔQALYs). The ICER is compared to a threshold value, v. The threshold value is used to represent the health displaced (generated) by the increased (decreased) intervention costs. If the ICER < v then this indicates that the new treatment will create more health than it displaces and so may be considered value for money. An equivalent decision-making approach is to use INHB. This captures the expected net health produced by adopting a new intervention over current practice. The formula is given by:

$$\text{INHB} = \Delta\text{QALYs} - \frac{\Delta\text{Costs}}{v}$$

A positive INHB indicates that the new intervention is expected to generate more health than it displaces. As mentioned above, incorporating heterogeneous transition probabilities will result in heterogenous estimates of costs (ΔCosts_{x_i}) and health outcomes (ΔQALYs_{x_i}) for each treatment. In this paper we refer to these as an "individual level" estimates and use the subscript x_i.[1] This allows us to calculate INHB as a function of observed patient characteristics:

$$\text{INHB}_{x_i} = \text{INHB}|X = x_i$$

Algorithms to calculate INHB_{x_i} with cohort and microsimulation models (while reflecting posterior uncertainty) are described formally in appendix A2. Briefly, a probabilistic sensitivity analysis must be carried out for each unique patient profile x_i. This results in a distribution of INHB_{x_i} predictions for each individual. Depending on the algorithm, this distribution reflects parametric uncertainty in the model (i.e., a posterior distribution for INHB_{x_i}) or a combination of parametric uncertainty plus stochastic uncertainty.

Note that heterogeneity is a common feature of decision models and can arise due to one or a combination of model inputs. A common example of an input which drives heterogeneity in decision models is the difference in all-cause mortality between men and women (Briggs et al., 2006). As the case study will show, the dependence of INHB on x_i does not necessarily require ML and so arises in both standard decision models and decision models incorporating ML, albeit to different degrees.

3.5 | Step 3: Define treatment allocation rules and estimate payoffs

Below we describe how three different treatment allocation "policies" can be implemented and evaluated using the a given set of INHB_{x_i} estimates from the distribution calculated above. Each approach is described using a policy rule π which assigns individuals to a binary treatment decision in which 1 means treat with the new treatment and 0 means keep with current practice $\pi : x \rightarrow \{0, 1\}$. Depending on the rule, π may depend on patient characteristics.

3.5.1 | One size fits all policy

A "one size fits all" (OSFA) approach entails making a treatment decision for the full population based on its average payoff. This is the standard (non-stratified) approach to decision making. The policy rule is given by:

$$\pi^{\text{OSFA}} = \begin{cases} 1, & \frac{1}{I}\sum_{i=1}^{I} \text{INHB}_{x_i} \geq 0 \\ 0, & \frac{1}{I}\sum_{i=1}^{I} \text{INHB}_{x_i} < 0 \end{cases}$$

If the average payoff in the population is positive then all are treated with the new intervention, otherwise all get current practice. Note that once the OSFA policy has been learned for a given population, the rule does not depend on specific patient characterizes.

3.5.2 | Plug-in policy

This approach makes tailored decisions for each individual using a function of their characteristics. A prediction for INHB_{x_i} is made for each individual, δ_{x_i}. Those with $\delta_{x_i} \geq 0$, are allocated to the new intervention, otherwise they are allocated to current practice:

$$\pi^{PI}(x_i) = \begin{cases} 1 & \delta_{x_i} \geq 0 \\ 0 & \delta_{x_i} < 0 \end{cases}$$

This is known as the plug-in policy ($\pi^{PI}(x_i)$) in the ML literature (Athey & Wager, 2021). The difference between the expected payoff with $\pi^{PI}(x_i)$ and π^{OSFA} is the EVIC, that is, the expected gain from individualizing care as defined by Basu and Meltzer (2007). A direct approach to specifying δ_{x_i} is to use the decision model prediction, $\delta_{x_i} = \text{INHB}_{x_i}$, this is the approach taken in this article.[2]

Plug-in policies based directly on decision model predictions have some specific disadvantages. First is the reporting of results. Predictions will often be a complex function of patient characteristics, therefore it may be hard to summarize the results of an analysis. For use in decision making, it would be necessary to create an online tool which would take in x_i and return a yes/no treatment decision. Second is acceptability. The policy takes x_i as an input and outputs a treatment decision. Decision makers may be uncomfortable not knowing the specific grounds for a decision. Variable importance analysis such as the best linear projection approach can be used to provide some "explainability" to decision makers (Bonander & Svensson, 2021; Semenova & Chernozhukov, 2020).

3.5.3 | Subgroup policies

Decision making based on subgroups can address the reporting and acceptability limitations of $\pi^{PI}(x_i)$. Rather than a complex function to make decisions, subgroup-based approaches use a subset of explicit patient characteristics to define subgroups. These subgroups are then linked to treatment decisions. This makes the link between individual characteristics and decision making more transparent and easier to summarize and report.

In this paper we utilize an approach to subgroup definition based on optimal policy tree learning algorithms (Athey & Wager, 2021). In this approach, an algorithm is used to map a patient's covariate profile and INHB_{x_i} to an individualized treatment decision (Bonander & Svensson, 2021; Hirano & Porter, 2009; Kitagawa & Tetenov, 2021; Manski, 2004). Specifically, we adapt the tree-based approach proposed by Athey and Wager (2021).[3]

For tree-based methods the task is to find the optimal policy tree $\pi^{T}(x_i)$ from the set of trees of a given depth Π. In our application the optimal policy tree is the one which maximizes the average of INHB_{x_i} over the population of I individuals. Formally this can be written as:

$$\widehat{\pi}^{T}(x_i) = \text{argmax}\left(\frac{1}{I}\sum_{i=1}^{I}\left(2\pi^{T}(x_i) - 1\right)\text{INHB}_{x_i} : \pi^{T} \in \Pi\right)$$

Expanding Π to allow trees of greater depth results in allocation rules which utilize a greater number of subgroups by incorporating more elements of X and creating a greater number of splits. Increasing tree depth may better approximate treatment assignment from $\pi^{PI}(x_i)$. However, increasing depth has two downsides. First, it will increase the complexity of the treatment algorithm and so undermine its usefulness for clinical decision making. Second, an inappropriately deep tree may overfit the data resulting in poor out of sample performance.

3.5.4 | Estimating payoffs

Given a sample of I patient profiles from the population of interest $\tilde{x} = x_1, x_2, ...x_i...x_I$, the expected payoff with a given policy in this sample can be calculated as:

$$\frac{1}{I}\sum_{i=1}^{I}\pi(x_i)\text{INHB}_{x_i}$$

When estimating the payoff with a given policy we want to avoid using the same estimates of INHB_{x_i} to both define the treatment allocation rules and estimate the payoffs. This is because variance in estimation of decision model inputs (such as the AFT-BART) can lead to spurious variation in INHB_{x_i} estimates. Using the same INHB_{x_i} estimates to define the treatment allocation rules and to estimate the payoffs from these rules will result in treating this spurious variation as real heterogeneity. This may lead to an optimistic bias in estimating the benefits of treatment stratification.

Cross validation can be used to ameliorate this bias for π^{OSFA} and $\pi^{T}(x_i)$. For each policy option π (i.e., OSFA, and each tree depth considered): (1) for data fold 1: k − 1, find the policy π that maximizes average of $\pi(x_i).\text{INHB}_{x_i}$ in the population, (2) for the selected policy, calculate the average $\pi(x_i).\text{INHB}_{x_i}$ for individuals in the kth fold, (3) repeat steps 1-2 for all k folds, (4) The payoff for each policy option is the average out of sample expected $\pi(x_i).\text{INHB}_{x_i}$ over the k folds, (5) repeat steps 1–4 for all policy options. The decision model based plug-in policy is defined and evaluated using INHB_{x_i}, so it is not possible to carry out the cross-validation step. Therefore the benefits of $\pi^{PI}(x_i)$ should be understood as an upper bound for the benefits of stratification and interpreted with caution.

Depending on the algorithm used to calculate INHB_{x_i} (see appendix A2), it is possible to have a distribution of INHB_{x_i} which describes their parametric uncertainty. In this case, the payoffs with each policy option, π, with current information should be calculated by: (1) taking a large number of samples from the distribution of INHB_{x_i}, (2) calculating the payoff from π in each sample (including the cross-validation step if possible), (3) taking the average over this distribution of payoffs.

3.5.5 | Value of heterogeneity

Because the policy options considered here take advantage of heterogeneity, this evaluation fits into the VoH framework, which considers the marginal gains from stratifying decision making, or "static" VoH[4] (Espinoza et al., 2014; Pataky et al., 2022). This framework was originally defined conditional on the number of subgroups for stratification. To allow comparison of all policy options, we condition on the number of variables or stratification factors.

Figure 3 shows different policy options for a hypothetical decision problem, with a given evidence base and set of treatment options. There is a potential increase in population net health benefit achieved by moving from a OSFA policy (point A) to policies which stratify the population based on two variables (points B or C) and six variables (points D and E). The two points B and C represent two different ways in which treatment can be allocated, based on two variables, using the subgroup policy (point B) or plug-in policy (point C).

For a given number of variables there is an optimal configuration which gives the highest population net health benefit, this is represented by the points A, C and D. The set of these optimal points is called the efficiency frontier. When no variables are conditioned on the OSFA policy will be on the efficiency frontier (point A). At all other points, the plug-in policy will be on the frontier because it represents the upper bound from the benefits of stratification.

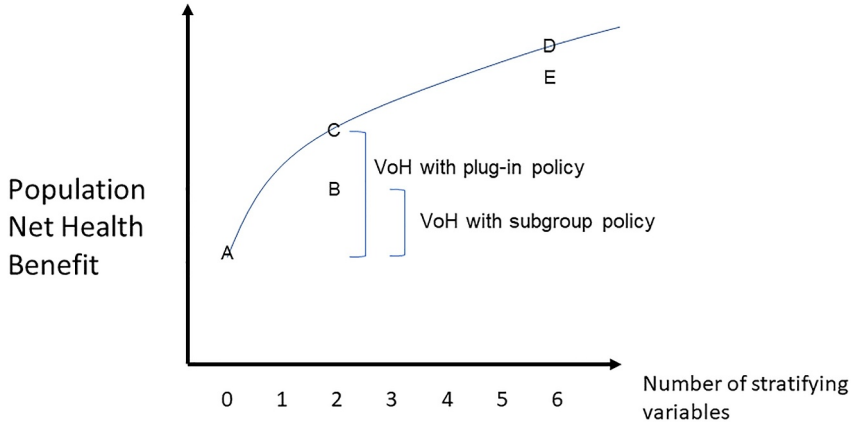

FIGURE 3 The static value of heterogeneity (VoH) and the policy efficiency frontier. Moving from left to right is associated with stratifying on a greater number of variables. Point A gives the payoff with no stratification that is, the one size fits all policy. Point B shows the payoff when stratifying by 2 variables using a subgroup policy, this payoff is greater than A. The static VoH is the vertical distance between these points. Point C is the payoff with the plug-in policy and is the upper bound for VoH when stratifying on 2 variables. Points D and E show payoffs with 6 variables, D is from a plug-policy and E is a subgroup policy.

Because the subgroup policy approximates the plug-in policy, the plug-in policy will always be as close as or closer to the efficiency frontier.

Theoretically, if we observed an infinite number of patients with all possible combinations of patient characteristics we could identify the true functional form for the relationship between X and population net health benefit. We could then identify the true optimal frontier and make perfectly individualized decisions with the plug-in policy. In practice this ideal can only be approximated given inherently finite datasets.

4 | CASE STUDY

Here we demonstrate how to integrate ML and decision modeling using a cardiovascular case study. It is expected that reducing blood pressure for patients at risk of cardiovascular disease will improve both health outcomes and reduce costs. However, there is uncertainty about the blood pressure targets that should be adopted. More intensive targets are expected to reduce the risk of cardiovascular events but may be associated with a higher risk of side effects and increased treatment costs.

One key US trial which aimed to answer this question was the Systolic Blood Pressure Intervention Trial (SPRINT) ($n = 9361$) which compared outcomes between "standard" (<140 mm Hg) versus "intensive" (<120 mm Hg) systolic blood pressure control targets. The trial showed reduced rates of death and cardiovascular events with the intensive target (SPRINT research group, 2015).

Because the SPRINT IPD did not contain all the information required for decision making, Bress et al. built a microsimulation model to combine the results of SPRINT with data from the wider literature to estimate the cost effectiveness of intensive treatment (Bress et al., 2017). In the model, patients began in a "well" state and were simulated over their full lifetimes. In every time period patients were at risk of: a composite cardiovascular disease (CVD) event, serious adverse event and non-CVD death. Time to first composite CVD event was the primary outcome in SPRINT and included myocardial infarction, stroke, non-MI acute coronary syndrome, chronic heart failure and CVD death.

The Bress model is split into two sections: the first 5 years (representing the trial follow up period) and beyond 5 years (the extrapolation period). For the first 5 years, Bress et al. used SPRINT data to model the rate of experiencing the first composite CVD event using a constant hazard for "intensive" (1.65 per 100 person-year) and "standard" (2.19 per 100 person-year). There was no adjustment for patient characteristics making this equivalent to an exponential survival model with no covariates. After experiencing the composite CVD event once, risk of repeated events came from the pooled cohort equations which accounted for a range of patient characteristics, see appendix A1 for further details (Goff et al., 2014). Constant hazard survival models estimated from SPRINT were also used to model serious adverse events and non-CVD deaths.

For the extrapolation period, the authors modeled various scenarios with different assumptions. To bracket issues around extrapolation and ML, we replicate the worst-case scenario modeled by Bress et al.[5] In this scenario, after 5 years all patients were assumed to be non-adherent, meaning all patients experienced the composite CVD event at rates derived from the pooled cohort equations. Serious adverse events were as observed in the standard treatment arm in SPRINT and non-CVD death came from life tables. For this scenario Bress found that, for a one size fits all policy, "intensive" was more costly and resulted in more QALYs. It was found to be cost effective at a threshold (v) of $100,000/QALY.

For our case study application we keep all Bress parameters as in the original model, but use ML to model the primary outcome (time to first composite CVD event). Note that our case study is based on a reproduction of the model developed by Bress. This reproduction was not validated to the standards required for policy making, so all case study results are meant to illustrate the methods under consideration in this paper, as opposed to directly informing blood pressure clinical guidelines in the US.

4.1 | Step 1: Model heterogeneity in a time-to-event endpoint using Bayesian modeling

We used ML to estimate the risk of the primary outcome in SPRINT and incorporate this into the Bress model. 8746 patients provided complete data for the primary outcome (93.4% of SPRINT patients). The baseline characteristics used to model outcomes included demographics, risk factors/biomarkers, Framingham 10 years risk score, patient history and medication use. Summary statistics (means and standard deviations) matched those reported in Bress et al. See the appendix for full details and for comparison of the variables used in the BART with variables used in the pooled cohort equations (A1).

The AFT BART models were fitted using the AFTrees package in R (Henderson, 2019). We did not add an estimated propensity score covariate (as suggested by Hahn et al.), due to excellent covariate balance resulting from the randomized setting (Hahn et al., 2020).

To choose the BART hyperparameters we used a 10-fold cross-validation grid search using the hyperparameter candidate values suggested by Henderson et al. (2020). This was carried out for the treatment and control model separately. The best performing values were the same for both treatment and control: $\zeta = 0.9$, $\xi = 3$, $\varphi = 50$. To estimate the time-to-event probabilities, as described above we fit four BART models, one for each sub-sample of the treatment and control arms.

Posterior estimation was carried out using a Metropolis-within-Gibbs sampler (Henderson et al., 2020). 10,000 posterior samples were taken, 2000 were burnt to allow the model to converge. Due to the large number of objects used to characterize the BART, this results in large memory demands, therefore, we thinned the 8000 posterior samples, by keeping approximately every 8th draw to give 1000 posterior draws for use in the probabilistic sensitivity analysis.

4.2 | Step 2: Incorporate counterfactual time-to-event estimates into a decision model

To compute estimates of INHB_{x_i} we jointly propagated both uncertainty in Θ and within patient variability evaluating the model 10,000 times for each of the ($I = 8746$ individual profiles (87.46 million simulations)). Note that because of computational cost, the algorithm chosen did not allow for characterization of parametric uncertainty for the estimates of INHB_{x_i} (see appendix A2 for further details), therefore we took the average over the 10,000 simulations for each patient profile.

4.3 | Step 3: Define treatment allocation rules and estimate payoffs

For each policy, the treatment allocation rules and their payoffs were all derived from the above estimates of INHB_{x_i} as described in the methods section. For the subgroup approach we learned trees of depth 2 and 3 using exhaustive tree search in the policytree package in R (Sverdrup et al., 2022). Deeper trees were not estimated due to computational expense. We converted continuous characteristics into ordered integers to obtain clinically interpretable tree outputs, and to improve run-time of the algorithm (see appendix A1 for further details). For the subgroup and OSFA policy, we used 5-fold cross validation to estimate the out-of-sample payoffs.

4.4 | Decision model without ML

To demonstrate the impact of including ML derived inputs into decision models we compare results with: (1) all model inputs kept at their original values reported in Bress but replacing the model for the primary outcome with a BART; (2) all model inputs kept at their original values as reported in Bress. The former we refer to as a "decision model incorporating ML", the latter as a "standard decision model".

The only difference between these two scenarios is the method used to model the time to composite CVD event (the primary outcome). In the "standard decision model" the survival model for the primary outcome is time invariant and does not contain any heterogeneity. However, there will be heterogeneity in the overall model. This is because the pooled cohort equations are conditioned on a set of patient characteristics. This means that $INHB_{x_i}$ will differ across individuals in both the decision model incorporating ML and the standard decision model. The difference between the models is the extent of heterogeneity modeled and the number of model parameters. Table A1.1 compares the patient characteristics used to model overall heterogeneity in each decision model. Because $INHB_{x_i}$ can be calculated for the standard decision model, the payoff from all of the policy options (one size fits all, plug-in policy and subgroups) can be calculated using the methods previously outlined.

It should also be noted that the differences in results between "standard" and "intensive" treatment arms are not only due to differences in the primary outcome. As described above, in the first 5 years the Bress decision model also takes into account differences between treatments in serious adverse events and non-CVD deaths.

All analysis was carried out using R with the Viking Cluster, a high performance computer facility provided by the University of York.

5 | CASE STUDY RESULTS

5.1 | BART results

Figure 4 shows the counterfactual survival curves predicted by the BART model for a single patient. This plot includes both the sampling uncertainty in estimation and the expected survival curve.

Figure 4 shows the survival for one individual with parametric uncertainty, while Figure 5 summarizes the expected survival for all 8746 of the individuals in SPRINT that is, the between patient variability in survival. The left panel of Figure 5 shows the distribution of individual level survival predictions for "standard" and "intensive" at year five. The right panel of Figure 5 shows the difference in individual level survival at year five. The mean of the distribution is above zero indicating that survival (aggregated across individuals) is higher with "intensive" treatment. However, there is significant between-patient variability with 14% of patients predicted to have worse survival with "intensive" compared to "standard" treatment. The estimate used in Bress (and the standard decision model) is also shown in the right side panel of Figure 5. As expected this closely models the average patient.

Figure 5 shows the magnitude of differences in predicted individual survival from the BART model. It can also be useful to understand what is driving these differences. A best linear projection analysis fits an ordinary least squares

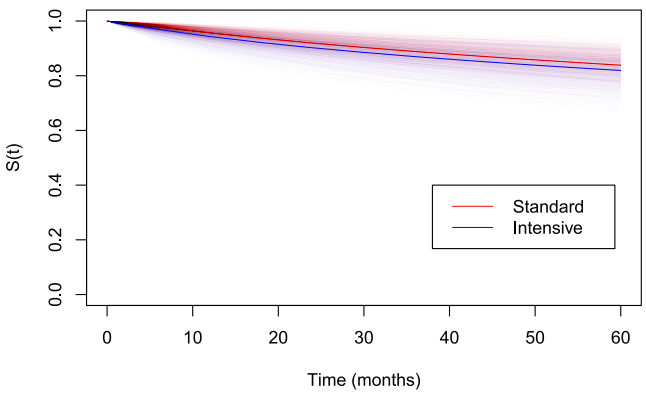

FIGURE 4 Individual counterfactual survival curve for an example patient with standard and intensive treatment. Darker lines indicate expected survival, faint lines indicate parametric uncertainty in estimates.

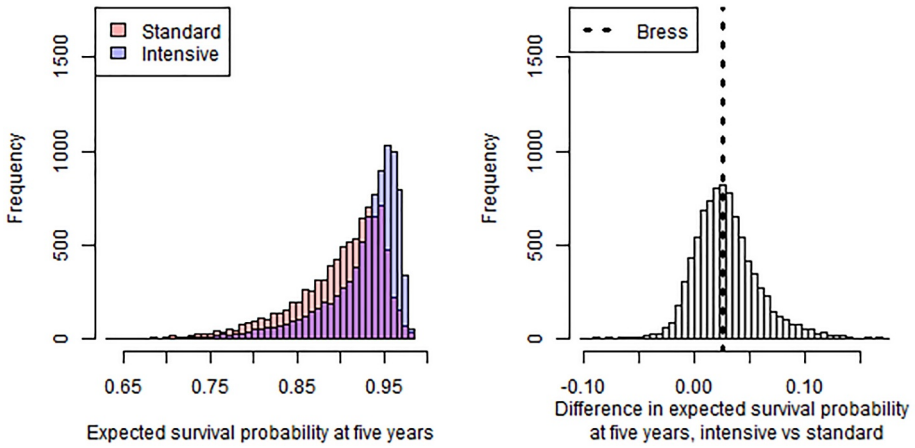

FIGURE 5 Between individual variation in expected survival at 5 years estimated by the Bayesian Additive Regression Trees (BART) survival model. All results show predictions conditional on patient characteristics with parametric uncertainty integrated out. The left panel summarizes the between patient distribution of 5 year survival predictions for both intensive and standard. The right panel shows the distribution of individual differences in expected survival at 5 years. As expected this distribution is centered on the estimate from Bress et al. which assumed the same survival for all individuals in the trial.

regression with the difference in expected survival as the independent variables and patient characteristics as dependent variables (Chernozhukov, Demirer, Duflo, & Fernandez-Val, 2018; Semenova & Chernozhukov, 2020). This allows us to explore the relative importance of different variables in predicting the survival outcomes. Figure A3.1 in the appendix (A3) plots the results of this analysis. Characteristics which are associated with statistically significant reductions in the effect of "intensive" treatment are: female sex, Hispanic ethnicity, black ethnicity, history of clinical CVD, white ethnicity, aspirin use, statin use. Characteristics which are associated with statistically significant increases in the effect of "intensive" treatment are: higher BMI, higher triglycerides, worse diabetic status, higher cholesterol, use of hypertensives, kidney disease, smoking status and age.

The Bress model was built on a 6-month cycle length and we derived individualized transition probabilities (*tp*) which also varied by treatment and over time. The underlying hazards were assumed constant within each 6-month cycle. As transition probabilities are derived from survival probabilities this shows a very similar pattern of results.

5.2 | Decision making results

5.2.1 | One size fits all

Table 1 presents expected absolute and incremental costs and QALYs with the one size fits all policy. For the standard decision model, intensive treatment was expected to cost $25,496 more on average and to also provide 0.344 additional QALYs per person. This implied an ICER of $74,073/QALY. For the decision model incorporating ML, the additional costs of intensive treatment were higher at $27,383 and additional QALYs were also higher at 0.424 QALYs. This translates to an ICER of $64,556/QALY, a reduction of 13%. Note that this policy had the highest expected costs and QALYs of all the policy options.

Table 1 reports that the incremental gain in INHB with the standard decision model was 0.089 QALYs and was 0.150 QALYs with the decision model incorporating ML, implying an increase of 69% in INHB with ML. Columns 1 and 2 in Table 2 report INHB at the population level. This was calculated by multiplying by the eligible SPRINT population in the US of 16.8 million (Bress et al., 2016).

5.2.2 | Plug-in policy

Table 1 reports the expected values for absolute and incremental costs and QALYs with the plug-in policy. Figure 6 plots the incremental costs and QALYs for each individual in SPRINT on a cost effectiveness plane. This shows a broader

TABLE 1 Absolute and incremental per person costs and QALYs with each policy option and model type.

	Treated with intensive	Baseline mean costs	Mean costs with policy	Mean incremental costs with policy	Baseline mean QALYs[a]	Mean QALYs with policy	Mean incremental QALYs with policy	Mean INHB[b]
Decision model with machine learning								
One size fits all	100%	$245,700	$273,083	$27,383	11.498	11.922	0.424	0.150
Plug-in	97.2%	$245,700	$268,514	$22,814	11.498	11.887	0.389	0.161
Tree depth 2	96.6%	$245,700	$271,901	$26,201	11.498	11.91	0.412	0.150
Tree depth 3	87.2%	$245,700	$271,622	$25,922	11.498	11.908	0.41	0.151
Standard decision model								
One size fits all	100%	$247,515	$273,011	$25,496	11.515	11.859	0.344	0.089
Plug-in	91.9%	$247,515	$267,507	$19,992	11.515	11.814	0.299	0.099
Tree depth 2	87.2%	$247,515	$269,747	$22,232	11.515	11.83	0.315	0.093
Tree depth 3	84.3%	$247,515	$268,233	$20,718	11.515	11.817	0.302	0.095

Note: Here baseline is assumed to be treatment according to "standard" blood pressure targets.

[a]QALYs, quality adjusted life years.

[b]INHB, incremental net health benefit.

TABLE 2 Static value of heterogeneity for the decision model with and without the machine learning survival model.

	Treat all with standard	Treat all with intensive	Depth 2 tree	Depth 3 tree	Treat according to plug-in
Decision model incorporating machine learning					
Number of stratifying variables	0	0	2	5	26
Treated with intensive	0%	100%	97.2%	96.6%	87.2%
Population expected incremental net health benefit (QALYs[a])	0	2,525,700	2,528,300	2,543,800	2,706,000
Static value of heterogeneity (QALYs)	-	0	2600	15,500	162,200
Standard decision model					
Number of stratifying variables	0	0	2	3	11
Treated with intensive	0%	100%	91.9%	87.2%	84.3%
Population expected incremental net health benefit (QALYs)	0	1,499,300	1,564,500	1,589,200	1,662,100
Static value of heterogeneity (QALYs)	0	0	65,200	24,700	73,000

Note: It is assumed that decision making is based on incremental net health benefit (INHB) which considers both health outcomes and opportunity costs. The static value of heterogeneity is calculated as the marginal gain from increasing the stratifying variables from left to right on the table. Note that for the subgroup policy the number of stratifying variables depends on the tree specific tree, for the plug-in policy it depends on the total number of patient characteristics used to model heterogeneity.

[a]QALYs, quality adjusted life years.

spread of predictions with the decision model incorporating ML. For both models the average ICER across individuals was below the $100,000/QALY threshold. The dashed line indicates the cost effectiveness threshold, "intensive" treatment is considered cost effective for all those falling below this line. With the standard decision model intensive treatment is expected to be cost effective for 84% of patients. With the decision model incorporating ML this is expected to increase to 87%.

The per person distribution of $INHB_{x_i}$ is shown in Figure 7. For the standard decision model $INHB_{x_i}$ had a mean of 0.089 QALYs, the lowest prediction for any individual was −0.462 and the highest was 0.450 QALYs. For the decision model incorporating ML the mean was 0.150 QALYs and the range across individuals was larger, ranging from −0.517 to 0.964 QALYs. As reported in Table 1, treating only those with $INHB_{x_i} > 0$ is expected to result in a per-person gain of 0.099 QALYs with the standard decision model and 0.161 with the decision model incorporating ML, implying an increase of 63% in INHB relative to the standard model. The final columns in Table 2 reports INHB at the population level with the plug-in policy and will be discussed further under the "VoH" section.

Figures A3.2 and A3.3 in the appendix (A3) show the results of a best linear projection analysis applied to explain heterogeneity in $INHB_{x_i}$. This allows us to understand the importance of different variables in predicting heterogeneity in individual INHB for the models with and without ML. For the decision model with ML, the predictors of heterogeneity in $INHB_{x_i}$ closely track those which predict heterogeneity in the survival outcomes. As expected, for the standard decision model the drivers of heterogeneity in $INHB_{x_i}$ are the inputs to the pooled cohort equations.

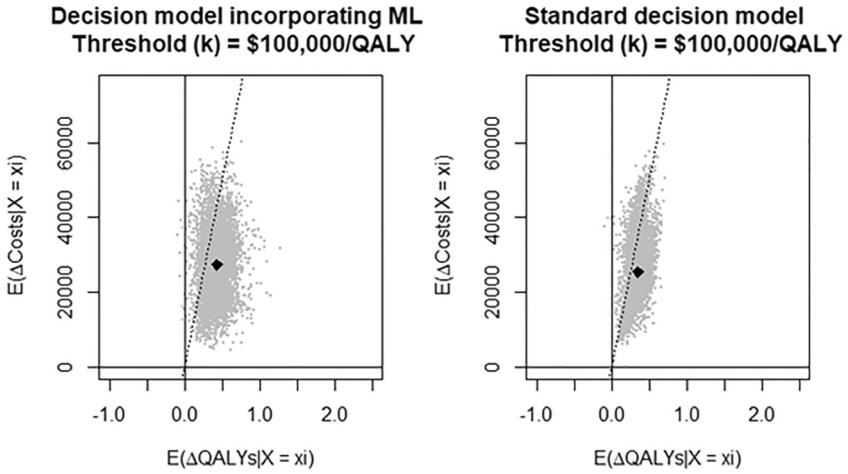

FIGURE 6 Individual level cost effectiveness plane plotting the joint distribution of incremental costs and incremental quality adjusted life years (QALYs) for each of the individuals in SPRINT. Each point represents a prediction for an individual patient profile (xi) in SPRINT. Results are presented for the decision model with machine learning (ML) (left panel) and the standard decision model (right panel). The dashed line indicates the cost effectiveness threshold, "intensive" treatment is considered cost effective for all those falling below this line, 87% for the decision model with machine learning (ML), 84% for the standard decision model. The diamond shows the location of the average incremental cost effectiveness ratio (ICER). ML, machine learning; QALYs, quality adjusted life years.

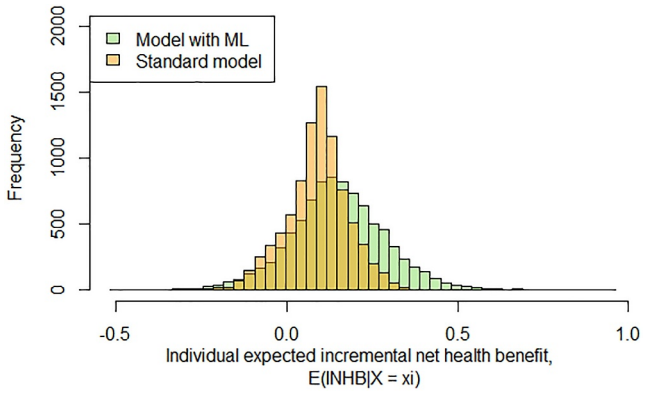

FIGURE 7 Distribution of individual level estimates of incremental net health benefit ($INHB_{x_i}$) with the decision model incorporating ML (green) and the standard decision model (orange). This shows predictions for INHB conditional on patient characteristics (x_i) with parametric uncertainty integrated out. ML, machine learning.

5.2.3 | Subgroups

Figure 8 shows the policy trees of depth 2 and depth 3 for the standard decision model and the decision model incorporating ML. For the standard decision model, the depth 2 tree stratifies on two variables, kidney function and age. For the depth 3 tree SBP is added to the list of splitting nodes (stratifying on a total of 3 variables). For the decision model incorporating ML the depth 2 tree stratifies on two variables, age and sex. To this the depth 3 tree adds systolic blood pressure, sub clinical cardiovascular disease and kidney disease (stratifying on a total of 5 variables). See Table 1 for expected absolute and incremental costs and QALYs for each tree depth.

Table 1 reports the expected values for absolute and incremental costs and QALYs with the depth 2 and depth 3 policy trees.

5.2.4 | Value of heterogeneity

Table 2 below aggregates and compares all the policy options within the VoH framework. Results are reported at the population level assuming decision making based on INHB with a threshold of $100,000/QALY. As expected, population expected INHB increases with increasing stratification (Espinoza et al., 2014). The plug-in policy stratifies on the total number of patient characteristics used to model heterogeneity. For the standard decision model this is the 11 variables in the pooled cohort equations, the model with ML includes these in addition to the variables used in the BART model to give a total of 26 (see Table A1.1 in the appendix).

By definition, the policy with the highest population level INHB was the plug-in policy in which treatment choice was defined and evaluated using $INHB_{x_i}$ (as discussed this should be interpreted as an upper bound). With a threshold of $100,000/QALY and scaled to the population level the estimated gains from the plug-in policy compared to one size fits all were (1,662,100–1,499,300 =) 162,800 QALYs with the standard decision model and (2,706,000–2,525,700 =) 180,300 QALYs for the decision model with ML. The second-best option was with subgroups defined using a depth 3 tree. The lowest population INHB was with the "OSFA policy". We found that the marginal gains from stratification (i.e., the static VoH) differed between the decision model with ML and the standard decision model. Compared to the standard decision model, the decision model with ML gains relatively more from the plug-in policy and relatively less from subgroup stratification.

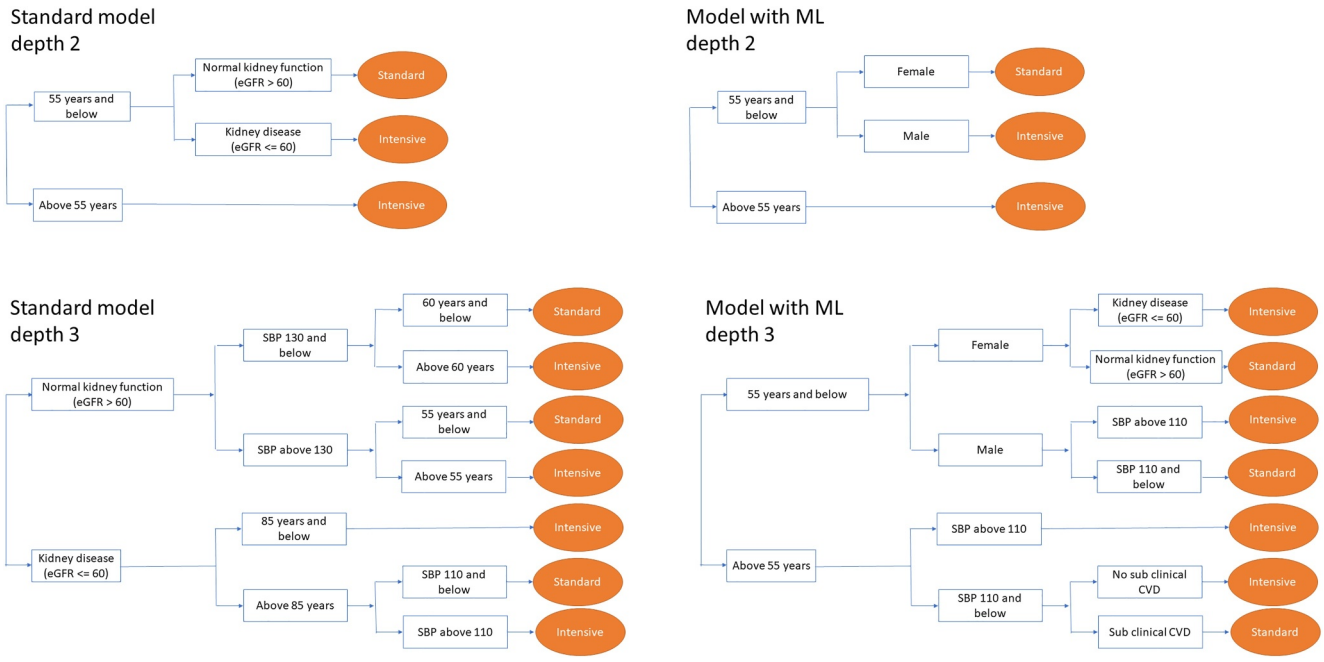

FIGURE 8 Optimal policy trees for the standard decision model and the decision model incorporating machine learning (ML) for depth 2 and depth 3 policy trees. Respectively, the left and right panels show the optimal policy trees learned from the decision model which does not include the machine learning estimate and the decision model which does include the machine learning estimate. Respectively, the upper and lower panels show the optimal policy trees learned with depth 2 and depth 3 trees. ML, machine learning.

6 | DISCUSSION

In this paper we outlined a novel approach to stratified decision making, integrating ML and decision modeling. We demonstrated how a BART ML algorithm could be used to inform a key input into a decision model, and how this decision model can be used to generate individual specific estimates of INHB ($INHB_{x_i}$). We then showed how these individual specific payoffs can be used to calculate the expected benefit of various policy allocation rules in a VoH framework. In defining treatment subgroups we proposed an innovative application of policy tree algorithms from the ML literature.

In our blood pressure control case study we compared payoffs between a standard decision model and the decision model incorporating ML in addition to investigating the gains from increasing stratification. We found that the expected payoff from a standard "one size fits all" cost effectiveness analysis was influenced substantially by incorporating ML with decision model inputs. With a threshold of $100,000/QALY the INHB increased by 69%. This change in the estimated incremental net health benefits of the "one size fits all" policy after including individual heterogeneity in the primary outcome is expected with complex model structures such as microsimulations. This is because here, the expected costs and QALYs for each individual are a nonlinear function of the primary outcome. Hence, even though the average of the heterogenous estimates of the primary outcome corresponds to the estimate used in the model with no heterogeneity, introducing heterogeneity in the primary outcome between individuals will change the expected costs and QALYs (Wilson, 2021).

Note that implementing ML directly models heterogeneity in the primary outcome with both intensive and standard care. This will impact baseline costs and QALYs in addition to incremental costs and QALYs. We note that an alternative implementation could have been to use a homogeneous risk estimate for the standard arm and add HTEs on top of this to model the counterfactual risk under the treatment. This would have resulted in the same baseline cost and QALY estimate in the models with and without ML.

Incorporating heterogeneity had two effects in the model. The first was a non-linear response to reductions in the risk of the primary outcome with the "intensive" target. Though the exponential model and the BART predict the same average risk reduction, the BART predicted large differences in risk across individuals (see Figure 5). This translated to an increase in QALYs because the gain from a reduction in risk was larger than the loss from an increase in risk of the same magnitude. The second effect was due to patient selection. The exponential survival model assumed that all individuals faced the same risk of the primary outcome. However, the BART predicted higher risk for some individuals rather than others. If healthier patients experience greater reductions in the risk of the primary outcome with the intensive target, then they will be selected for in the model and the model population will be at a lower risk of mortality and future cardiovascular events.

With ML methods there is always the possibility of "overfitting" to the observed data, where the predicted outcomes are precisely tailored to fit the training data but are not accurate once new data is introduced. In the integration of ML and decision modeling there are two main sources of potential overfitting. The first is overfitting the ML model used to derive the decision model input parameters. We controlled this with techniques to tune hyperparameters to limit model complexity (e.g., minimizing cross validated error) and cross-fitting methods which make out-of-sample predictions for each individual. The second source of overfitting can happen when making stratified treatment decisions. We provided a cross-validation step to avoid using the same $INHB_{x_i}$ estimates to define the treatment allocation rules and to estimate the payoffs from these rules. However, this could not be applied to the plug-in policy from the decision model, which can only be interpreted as an upper bound. Further research is required to formalize this bias and to develop methods to control for it when estimating the value of the plug-in policy.

Our base case analysis assumed that decision making was based on INHB which accounts for both health benefits to the patients treated and the health losses borne by other patients in the health care system which result from increased costs. Table 1 shows that gains in population INHB come from reducing costs rather than increasing QALYs. The same methods can be used to stratify decision making considering only health gains (as captured by QALYs). This results in different subgroups and different individualization of care in the plug-in policy (see appendix A4). The gains from stratification were lower in this case. The plug-in policy is expected to offer modest gains over OSFA and the optimal tree depth was zero that is, OSFA. This is because it was rare that intensive treatment was predicted to reduce QALYs with either the standard decision model or the decision model incorporating ML. Given the potential optimistic bias in the plug-in policy this indicates that there may not be gains from stratification when considering health gains alone.

In the introduction, we classified the methodology of Bonander et al. and Xu et al. as "pure IPD" and "augmented IPD" respectively (Bonander & Svensson, 2021; Xu et al., 2022). One novel contribution in this paper is to extend the application of ML in cost effectiveness analysis to a "full decision modeling" approach. By doing so we combine the strengths of ML and decision modeling. Because Xu et al. applied their method to an augmented dataset generated from SPRINT we can, to a degree, compare our results. Xu et al. found that the intensive treatment was cost effective for 69% of the individuals in SPRINT (with a threshold of $100,000/QALY). With the depth 3 tree we found that it was cost effective for 87.2% of the population. It should be noted that, among other differences, our method used a lifetime follow up, Xu used 15 years and our method incorporated length of life and quality of life, Xu considered only length of life. Further research is required to compare the relative strengths and weaknesses of "pure IPD", "augmented IPD" and "full decision modeling".

Embedding ML survival inputs into decision modeling will allow future applications to draw on the methodology, principles and guidelines developed in decision modeling. Future research should consider: methods to take account of missing data (Faria et al., 2014), issues with the extrapolation of survival curves beyond trial follow up (Latimer, 2011), network analysis of multiple treatment options (Dias et al., 2018); alternative decision model structures (Karnon et al., 2012); model calibration (Jalal et al., 2021) and competing risk models (Putter et al., 2007).

An additional issue is the structural uncertainty implied in choosing a procedure to model heterogeneity (Jackson et al., 2011). In this paper we chose a non-parametric BART to flexibly model individual log survival time, $f(X)$. Other non-parametric approaches such as splines could potentially be used here (Strong et al., 2014). However, these methods are limited by restrictions on the number of interacting covariates. Future research could consider the use of Gaussian process and/or Bayesian neural networks which, like a BART, allow for a high degree of interactions in addition to Bayesian uncertainty quantification (Feng & Zhao, 2021; Fernández et al., 2016). As discussed above, ML methods nested within a Bayesian framework are well suited to integration into decision modeling due to their natural uncertainty quantification. However, it is common in cost effectiveness modeling to utilize frequentist estimation approaches and interpret their point estimates and standard errors as Bayesian posteriors (Briggs et al., 2006). Simulation studies are required to understand how well highly flexible methods (both Bayesian and frequentist) characterize uncertainty.

Though we propagated parameter uncertainty in calculating outcomes, a key limitation of our case study application is the omission of a full parametric uncertainty analysis. The appendix (A2) provides a methodology for characterizing the relationship between payoffs and uncertainty quantified by Θ, however due to the computational cost associated with microsimulation models this was not feasible in this paper. Future work should investigate the use of simpler cohort models or meta-models to make full uncertainty analysis computationally feasible (Degeling et al., 2019). This would allow for an extension of the methods to compute the "dynamic VoH" by using value of information methods to calculate the change in the value of further research as the number of subgroups increases (Espinoza et al., 2014). Other frameworks for decision making under uncertainty exist such as those based on minimizing maximum regret (Manski, 2018). Future research could explore the implications of decision models incorporating ML for this literature.

The policy tree approach to identifying subgroups and the plug-in policy assume that decision makers are unconstrained in the components of X they can stratify on. However, decision makers may face ethical, legal or cost constraints such that they cannot stratify decisions based on some elements of X. For example, Figure 8 shows that, from the space of available trees, the policy tree which stratified treatment by age and sex was "optimal" from the perspective of maximizing population health. This may not be considered legally or ethically acceptable. It is possible to exclude these elements from consideration in the tree search algorithm. The algorithm may then be thought to find subgroups which are on a "constrained efficiency frontier". Further research is required to quantify the trade-offs implied in this approach and to understand its implications for algorithmic fairness (Nabi et al., 2019; Rambachan et al., 2020).

ACKNOWLEDGMENTS

We are very grateful for the insightful comments provided by two peer reviewers. We acknowledge computational support from the University of York High Performance Computing service, Viking and the Research Computing team. This work was prepared using SPRINT Research Materials obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center. The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or SPRINT investigators. We would like to acknowledge all those who provided valuable comments to early iterations of this work: Ben Powell, Richard Grieve, Karl Claxton, Simon Walker and Susan Griffin. This work was funded by the UK Medical Research Council (Grant \#: MR/T04487X/1).

CONFLICT OF INTEREST STATEMENT

David Glynn, Julia Hatamyar, Ankur Pandya, Marta Soares and Noemi Kreif report no relevant conflicts of interest. John Giardina reported receiving financial support from grants from the National Institutes of Health (R01NS104143) and the Harvard University Graduate School of Arts and Sciences (PhD stipend support) during the conduct of this study. Outside of this study, Mr. Giardina reported receiving support from the following entities: Facebook (unrestricted gift to Harvard University), the Agency for Healthcare Research and Quality (PhD student stipend to Harvard University), and the Center for Health Decision Science, Harvard T.H. Chan School of Public Health (student travel fund and Raiffa Award).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

David Glynn https://orcid.org/0000-0002-0989-1984
John Giardina https://orcid.org/0000-0002-6618-1671
Julia Hatamyar https://orcid.org/0000-0003-4145-1265
Ankur Pandya https://orcid.org/0000-0002-6130-2286
Marta Soares https://orcid.org/0000-0003-1579-8513
Noemi Kreif https://orcid.org/0000-0001-9008-5690

ENDNOTES

[1] It should be noted that this is a weaker form of individualization in which unobserved sources of heterogeneity are not accounted for (Basu, 2014).

[2] Note that other approaches to specifying δ_{x_i} are possible, such as those that estimate the relationship between INHB_{x_i} and x_i using linear or non-linear regression models (Jalal et al., 2021; Strong et al., 2014). These may have some advantages over the plug in policy based directly on the decision model prediction but are beyond the scope of this paper.

[3] The Athey and Wager approach originally used ML estimates of 'double-robust scores' to minimize overall regret (Athey & Wager, 2021). Here we use INHB_{x_i} as inputs to the algorithm instead of double-robust scores with the goal for treatment assignment to maximize population-wide net health benefit.

[4] The "dynamic" value of heterogeneity is the change in uncertainty as the number of subgroups increases.

[5] In the extrapolation period of the base case Bress analysis adherence was modeled as a function of the number of anti-hypertensive medications prescribed. After 15 years all patients were assumed non-adherent.

REFERENCES

Amram, M., Dunn, J., & Zhuo, Y. D. (2022). Optimal policy trees. *Machine Learning*, *111*(7), 2741–2768. https://doi.org/10.1007/s10994-022-06128-5

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*(1), 685–725. https://doi.org/10.1146/annurev-economics-080217-053433

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *Annals of Statistics*, *47*(2). https://doi.org/10.1214/18-aos1709

Athey, S., & Wager, S. (2021). Policy learning with observational data. *Econometrica*, *89*(1), 133–161. https://doi.org/10.3982/ecta15732

Baio, G. (2013). *Bayesian methods in health economics*. CRC Press.

Basu, A. (2014). Estimating person-centered treatment (PeT) effects using instrumental variables: An application to evaluating prostate cancer treatments. *Journal of Applied Econometrics*, *29*(4), 671–691. https://doi.org/10.1002/jae.2343

Basu, A., & Meltzer, D. (2007). Value of information on preference heterogeneity and individualized care. *Medical Decision Making*, *27*(2), 112–127. https://doi.org/10.1177/0272989x06297393

Bertsimas, D., Dunn, J., & Mundru, N. (2019). Optimal prescriptive trees. *INFORMS Journal on Optimization*, *1*(2), 164–183. https://doi.org/10.1287/ijoo.2018.0005

Bojke, L., Soares, M. O., Claxton, K., Colson, A., Fox, A., Jackson, C., Jankovic, D., Morton, A., Sharples, L. D., & Taylor, A. (2022). Reference case methods for expert elicitation in health care decision making. *Medical Decision Making*, *42*(2), 182–193. https://doi.org/10.1177/0272989x211028236

Bonander, C., & Svensson, M. (2021). Using causal forests to assess heterogeneity in cost-effectiveness analysis. *Health Economics*, *30*(8), 1818–1832. https://doi.org/10.1002/hec.4263

Bress, B., King, H., Beddhu, Z., Hess, R., Zhang, Z., Berlowitz, D. R., Conroy, M. B., Fine, L., Oparil, S., Morisky, D. E., Kazis, L. E., Ruiz-Negrón, N., Powell, J., Tamariz, L., Whittle, J., Wright, J. T., Supiano, M. A., Cheung, A. K., & Moran, A. E. (2017). Cost-effectiveness of

intensive versus standard blood-pressure control. *New England Journal of Medicine*, *377*(8), 745–755. https://doi.org/10.1056/NEJMsa1616035

Bress, T., Hess, C., Colantonio, L. D., Shimbo, D., & Muntner, P. (2016). Generalizability of SPRINT results to the U.S. Adult population. *Journal of the American College of Cardiology*, *67*(5), 463–472. https://doi.org/10.1016/j.jacc.2015.10.037

Briggs, A., Sculpher, M., & Claxton, K. (2006). *Decision modelling for health economic evaluation*. Oup Oxford.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018a). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1), C1–C68. https://doi.org/10.1111/ectj.12097

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2018b). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. Retrieved from. https://doi.org/10.3386/w24678

Coyle, D., Buxton, M. J., & O'Brien, B. J. (2003). Stratified cost-effectiveness analysis: A framework for establishing efficient limited use criteria. *Health Economics*, *12*(5), 421–427. https://doi.org/10.1002/hec.788

Curth, A., Lee, C., & van der Schaar, M. (2021). SurvITE: Learning heterogeneous treatment effects from time-to-event data. *Advances in Neural Information Processing Systems*, *34*, 26740–26753.

Degeling, K., Koffijberg, H., Franken, M. D., Koopman, M., & IJzerman, M. J. (2019). Comparing strategies for modeling competing risks in discrete-event simulations: A simulation study and illustration in colorectal cancer. *Medical Decision Making*, *39*(1), 57–73. https://doi.org/10.1177/0272989x18814770

Diaby, V., Adunlin, G., & Montero, A. (2013). Survival modeling for the estimation of transition probabilities in model-based economic evaluations in the absence of individual patient data: A tutorial. *PharmacoEconomics*, *32*(2), 101–108. https://doi.org/10.1007/s40273-013-0123-9

Dias, S., Ades, A. E., Welton, N. J., Jansen, J. P., & Sutton, A. J. (2018). *Network meta-analysis for decision-making*. John Wiley & Sons.

Espinoza, M. A., Manca, A., Claxton, K., & Sculpher, M. J. (2014). The value of heterogeneity for cost-effectiveness subgroup analysis: Conceptual framework and application. *Medical Decision Making*, *34*(8), 951–964. https://doi.org/10.1177/0272989x14538705

Faria, R., Gomes, M., Epstein, D., & White, I. R. (2014). A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *PharmacoEconomics*, *32*(12), 1157–1170. https://doi.org/10.1007/s40273-014-0193-3

Feng, D., & Zhao, L. (2021). BDNNSurv: Bayesian deep neural networks for survival analysis using pseudo values. *Journal of Data Science*, 542–554. arXiv preprint arXiv:2101.03170. https://doi.org/10.6339/21-jds1018

Fernández, T., Rivera, N., & Teh, Y. W. (2016). Gaussian processes for survival analysis. *Advances in Neural Information Processing Systems*, *29*.

Goff, D. C., Lloyd-Jones, D. M., Bennett, G., Coady, S., D'agostino, R. B., Gibbons, R., Greenland, P., Lackland, D. T., Levy, D., O'Donnell, C. J., Robinson, J. G., Schwartz, J. S., Shero, S. T., Smith, S. C., Sorlie, P., Stone, N. J., Wilson, P. W. F., & O'donnell, C. J. (2014). 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation*, *129*(25_suppl_2), S49–S73. https://doi.org/10.1161/01.cir.0000437741.48606.98

Hahn, P. R., Murray, J. S., & Carvalho, C. M. (2020). Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis*, *15*(3), 965–1056. https://doi.org/10.1214/19-ba1195

Henderson, N. (2019). AFTrees R package: Accelerated failure trees (Version 0.6.6).

Henderson, N. C., Louis, T. A., Rosner, G. L., & Varadhan, R. (2020). Individualized treatment effects with censored data via fully nonparametric Bayesian accelerated failure time models. *Biostatistics*, *21*(1), 50–68. https://doi.org/10.1093/biostatistics/kxy028

Hill, J., Linero, A., & Murray, J. (2020b). Bayesian additive regression trees: A review and look forward. *Annual Review of Statistics and Its Application*, *7*(1), 251–278. https://doi.org/10.1146/annurev-statistics-031219-041110

Hill, S., Mokgokong, L., Ward, B., Lister, S., Farooqui, U., & Gordon, J. (2020a). Cost-effectiveness of targeted screening for the identification of patients with atrial fibrillation: Evaluation of a machine learning risk prediction algorithm. *Journal of Medical Economics*, *23*(4), 386–393. https://doi.org/10.1080/13696998.2019.1706543

Hirano, K., & Porter, J. R. (2009). Asymptotics for statistical treatment rules. *Econometrica*, *77*(5), 1683–1701.

Hu, L., Ji, J., & Li, F. (2021). Estimating heterogeneous survival treatment effect in observational data using machine learning. *Statistics in Medicine*, *40*(21), 4691–4713. https://doi.org/10.1002/sim.9090

Jackson, C. H., Bojke, L., Thompson, S. G., Claxton, K., & Sharples, L. D. (2011). A framework for addressing structural uncertainty in decision models. *Medical Decision Making*, *31*(4), 662–674. https://doi.org/10.1177/0272989x11406986

Jalal, H., Trikalinos, T. A., & Alarid-Escudero, F. (2021). BayCANN: Streamlining bayesian calibration with artificial neural network metamodeling. *Frontiers in Physiology*, *12*, 662314. https://doi.org/10.3389/fphys.2021.662314

Karnon, J., Stahl, J., Brennan, A., Caro, J. J., Mar, J., & Möller, J. (2012). Modeling using discrete event simulation: A report of the ISPOR-SMDM modeling good research practices task force-4. *Value in Health*, *15*(6), 821–827. https://doi.org/10.1016/j.jval.2012.04.013

Kennedy, E. H. (2020). Towards optimal doubly robust estimation of heterogeneous causal effects. arXiv preprint arXiv:2004.14497.

Kitagawa, T., & Tetenov, A. (2021). Equality-minded treatment choice. *Journal of Business & Economic Statistics*, *39*(2), 561–574. https://doi.org/10.1080/07350015.2019.1688664

Kreif, N., DiazOrdaz, K., Moreno-Serra, R., Mirelman, A., Hidayat, T., & Suhrcke, M. (2021). Estimating heterogeneous policy impacts using causal machine learning: A case study of health insurance reform in Indonesia. *Health Services & Outcomes Research Methodology*, *22*(2), 1–36. https://doi.org/10.1007/s10742-021-00259-3

Latimer, N. (2011). NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. Report by the decision support unit.

Luedtke, A. R., & van der Laan, M. J. (2016). Super-learning of an optimal dynamic treatment rule. *International Journal of Biostatistics*, *12*(1), 305–332. https://doi.org/10.1515/ijb-2015-0052

Manski, C. F. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica*, *72*(4), 1221–1246. https://doi.org/10.1111/j.1468-0262.2004.00530.x

Manski, C. F. (2018). Reasonable patient care under uncertainty. *Health Economics*, *27*(10), 1397–1421. https://doi.org/10.1002/hec.3803

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B*, *65*(2), 331–355. https://doi.org/10.1111/1467-9868.00389

Nabi, R., Malinsky, D., & Shpitser, I. (2019). Learning optimal fair policies. In *Paper presented at the proceedings of the 36th international conference on machine learning, proceedings of machine learning research*. Retrieved from https://proceedings.mlr.press/v97/nabi19a.html

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., Rakow, T. (2006). Uncertain judgements: Eliciting experts' probabilities. https://doi.org/10.1002/0470033312

Padula, W. V., Kreif, N., Vanness, D. J., Adamson, B., Rueda, J. D., Felizzi, F., Jonsson, P., IJzerman, M. J., Butte, A., & Crown, W. (2022). Machine learning methods in health economics and outcomes research-the PALISADE checklist: A good practices report of an ISPOR task force. *Value in Health*, *25*(7), 1063–1080. https://doi.org/10.1016/j.jval.2022.03.022

Pataky, B., Sadatsafavi, P., Sadatsafavi, M., Peacock, S., & Regier, D. A. (2022). Tools for the economic evaluation of precision medicine: A scoping review of frameworks for valuing heterogeneity-informed decisions. *PharmacoEconomics*, *40*(10), 931–941. https://doi.org/10.1007/s40273-022-01176-0

Putter, H., Fiocco, M., & Geskus, R. B. (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine*, *26*(11), 2389–2430. https://doi.org/10.1002/sim.2712

Rambachan, A., Kleinberg, J., Ludwig, J., & Mullainathan, S. (2020). An economic perspective on algorithmic fairness. In *Paper presented at the AEA papers and proceedings*.

Ramponi, F., Walker, S., Griffin, S., Parrott, S., Drummond, C., Deluca, P., Coulton, S., Kanaan, M., & Richardson, G. (2021). Cost-effectiveness analysis of public health interventions with impacts on health and criminal justice: An applied cross-sectoral analysis of an alcohol misuse intervention. *Health Economics*, *30*(5), 972–988. https://doi.org/10.1002/hec.4229

Sadique, Z., Grieve, R., Diaz-Ordaz, K., Mouncey, P., Lamontagne, F., & O'Neill, S. (2022). A machine-learning approach for estimating subgroup- and individual-level treatment effects: An illustration using the 65 trial. *Medical Decision Making*, *42*(7), 923–936. https://doi.org/10.1177/0272989x221100717

Sculpher, M. J., Claxton, K., Drummond, M., & McCabe, C. (2006). Whither trial-based economic evaluation for health care decision making? *Health Economics*, *15*(7), 677–687. https://doi.org/10.1002/hec.1093

Semenova, V., & Chernozhukov, V. (2020). Debiased machine learning of conditional average treatment effects and other causal functions. *The Econometrics Journal*, *24*(2), 264–289. https://doi.org/10.1093/ectj/utaa027

SPRINT research group. (2015). A randomized trial of intensive versus standard blood-pressure control. *New England Journal of Medicine*, *373*(22), 2103–2116. https://doi.org/10.1056/NEJMoa1511939

Strong, M., Oakley, J. E., & Brennan, A. (2014). Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: A nonparametric regression approach. *Medical Decision Making*, *34*(3), 311–326. https://doi.org/10.1177/0272989x13505910

Sverdrup, E., Kanodia, A., Zhou, Z., Athey, S., & Wager, S., (2022). Policytree R package: Policy learning via doubly robust empirical welfare maximization over trees (version 1.2.1).

Wang, R., Lagakos, S. W., Ware, J. H., Hunter, D. J., & Drazen, J. M. (2007). Statistics in medicine—Reporting of subgroup analyses in clinical trials. *New England Journal of Medicine*, *357*(21), 2189–2194. https://doi.org/10.1056/nejmsr077003

Wilson, E. C. (2021). Methodological note: Reporting deterministic versus probabilistic results of Markov, partitioned survival and other non-linear models. *Applied Health Economics and Health Policy*, *19*(6), 789–795. https://doi.org/10.1007/s40258-021-00664-2

Xu, Y., Ignatiadis, N., Sverdrup, E., Fleming, S., Wager, S., & Shah, N. (2022). Treatment heterogeneity with survival outcomes. *arXiv preprint arXiv:2207.07758*. https://arxiv.org/abs/2207.07758

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Glynn, D., Giardina, J., Hatamyar, J., Pandya, A., Soares, M., & Kreif, N. (2024). Integrating decision modeling and machine learning to inform treatment stratification. *Health Economics*, 1–21. https://doi.org/10.1002/hec.4834