PharmacoEconomics
https://doi.org/10.1007/s40273-025-01562-4

ORIGINAL RESEARCH ARTICLE

Individualized Treatment Rules Based on Cost-Effectiveness Criteria in Microsimulations

Niklaus Meier[1,2] · **Ana Cecilia Quiroga Gutierrez**[2] · **Mark Pletscher**[3] · **Matthias Schwenkglenks**[1,4]

Received: 15 March 2025 / Accepted: 19 October 2025

Abstract

Background and Objective In cost-effectiveness analysis, treatment decisions are analysed at the population level. Combinations of treatment strategies that account for the heterogeneity of costs and effects across patients can be more cost-effective than a "one size fits all" approach. Individualized treatment rules (ITRs) assign a specific treatment to every patient based on their relevant characteristics, such that overall cost-effectiveness is optimized, but do not include feasibility or ethical considerations. We propose an approach for the design of ITRs based on simulated patient data from microsimulation models using statistical learning techniques.

Methods We mathematically define the optimal ITR and how to measure the value of an ITR in a cost-effectiveness context. We explore least absolute shrinkage and selection operator (LASSO) regression, classification trees, and policy trees to illustrate how standard statistical learning techniques can be used to derive ITRs. We compare the strengths and limitations of these three approaches in terms of three criteria: the incremental value of the ITRs compared to optimal treatment assignment in terms of net monetary benefit (NMB), computational speed, and the interpretability of the ITRs. We propose methods to describe the impact of parameter uncertainty on the ITRs. We also explore how stochastic uncertainty can impact the ITR incremental value. We illustrate the methods by applying them to a microsimulation model for haemophilia B comparing four treatment strategies as a case study. The relevant patient characteristics in this model are the annualized bleeding rate, age, and sex.

Results In our case study, a simple two-layer-deep classification tree is best suited based on the three criteria. This classification tree allocates treatments depending on whether the annualized bleeding rate of a patient is above or below 30 and whether their age is above or below 51. The optimal threshold values are uncertain based on the 95% credible ranges from the probabilistic analysis: 21–46 for annualized bleeding rate and 42–56 for age. Scenarios show that stochastic uncertainty has an impact on the incremental value of the ITR.

Discussion Based on methodological considerations and the empirical findings in our case study, we expect the superiority of classification trees for the derivation of ITRs to be generalizable to other microsimulation models. This finding needs to be confirmed in future applications. Stochastic uncertainty has significant impacts on the ITRs, such that accurate representations of individual patient pathways are particularly crucial when designing ITRs. Future research could explore further empirical models and analytical approaches for ITRs or consider the translation of ITRs into the real-world decision-making context.

Key Points for Decision Makers

Microsimulation cost-effectiveness models can be used to design individualized treatment rules.

Individualized treatment strategies can be more cost-effective than uniform treatment strategies but are also subject to uncertainty.

Individualized treatment rules can be expressed graphically as decision trees or as a list of IF statements.

✉ Niklaus Meier
niklaus.meier@bfh.ch

[1] Institute of Pharmaceutical Medicine (ECPM), University of Basel, Basel, Switzerland

[2] Institute of Health Economics and Health Policy, Bern University of Applied Sciences, Bern, Switzerland

[3] Novartis Pharma AG, Basel, Switzerland

[4] Health Economics Facility, Department of Public Health, University of Basel, Basel, Switzerland

1 Introduction

In cost-effectiveness analysis, treatments are typically evaluated at the cohort level, overlooking heterogeneity in prognostic factors and outcomes. However, the most cost-effective treatment for a population may not be optimal for every individual. Considering variations in treatment costs and effects across patients in the treatment assignment can lead to more cost-effective approaches than uniform treatment strategies [1–4]. Beyond understanding the value of individualized care, it is important to present individual-level cost-effectiveness findings in an easily interpretable manner. Individualized treatment rules (ITRs) can be used to determine how treatments can be assigned on the basis of patient characteristics [5, 6] that influence cost-effectiveness outcomes. They provide useful insight by identifying subgroups or individuals for whom a given treatment is cost-effective, without pre-specification of subgroups by the modellers. ITRs may not always be suitable for real-world implementation due to the costs of implementation [3] or due to ethical concerns, e.g. about discrimination or exacerbation of inequity [7–9]. Although ethical and feasibility considerations, as well as clinical relevance, are crucial for real-world applications, they fall outside the scope of this study, which focuses solely on methods for designing ITRs.

Recent studies have presented novel approaches for designing ITRs on the basis of individual patient data for cost-effectiveness analysis via machine learning [10–12]. Machine learning methods can help to individualize care in different ways [13, 14]. First, they can be used to identify heterogeneity in treatment effects [15], and thereby make counterfactual predictions of the cost-effectiveness of different treatments for individual patients. Second, machine learning can be used to identify optimal ITRs that maximize mean cost-effectiveness outcomes, based on the observed individual patient data as well as the predicted counterfactuals [16].

However, cost-effectiveness analysis based on individual patient data is not appropriate for decision-making if the observation period is too short to capture all relevant consequences of treatment decisions over the patients' lifetime or if the data do not include the entire patient population affected by a decision. In such cases, decision-analytic models become necessary [17]. Microsimulations are patient-level decision-analytic models in which the heterogeneity of patient characteristics, health events, and treatment sequences can be emulated. Microsimulations, like other modelling approaches, can be used to extrapolate costs and effects beyond the observation period of clinical data [18–20].

We propose a conceptual framework for ITRs based on cost-effectiveness microsimulations. We compare multiple approaches for designing ITRs in microsimulation models via machine learning. These leverage the availability of counterfactual outcomes under alternative treatments for every simulated patient. The availability of these simulated counterfactuals removes the need to predict counterfactuals via statistical machine learning techniques and instead allows for the direct identification of ITRs in the simulated patient data. This approach can be pursued with any cost-effectiveness microsimulation model and does not require any modifications to the underlying model.

As with other cost-effectiveness analyses, uncertainty should be considered when determining an optimal ITR. Uncertainty can be categorized as parameter uncertainty, stochastic uncertainty, or structural uncertainty [21]. Parameter uncertainty refers to the statistical uncertainty of the microsimulation model parameter values. We use probabilistic analysis to assess the impact of parameter uncertainty on the uncertainty of the ITR parameters and present it as credible intervals and graphically. We study how stochastic uncertainty can influence the ability of ITRs to optimally individualize care by varying the degree of stochastic uncertainty in the model. Structural uncertainty about the underlying cost-effectiveness model can affect ITRs but is not the focus of this article. Patient heterogeneity is distinct from uncertainty but modelled using uncertain model parameters. Hence, parameter uncertainty also creates uncertainty about heterogeneity, which we aimed to capture in the probabilistic analysis.

The remainder of the article is structured as follows. First, we provide a mathematical definition of the optimal ITR and outline how to measure the value of an ITR in a cost-effectiveness context. Second, we explore regressions with the least absolute shrinkage and selection operator (LASSO), classification trees, and policy trees to demonstrate how standard statistical learning techniques can be used to derive ITRs based on microsimulation models. Third, we compare the strengths and limitations of these three approaches in terms of three criteria: the incremental value of ITRs, computational speed, and the ability to express ITRs in an easily interpretable manner. Fourth, we propose methods to describe the impact of parameter uncertainty on ITRs. Fifth, we explore how random events, which, due to stochastic uncertainty, cannot be predicted with certainty based on patient characteristics and have heterogeneous impacts on patient outcomes, can impact the incremental value of an ITR. Sixth, we introduce an extension of a previously published microsimulation for haemophilia B [22], in which the gene therapy etranacogene dezaparvovec (ED) is compared with prophylactic and on-demand administration of extended half-life factor IX (EHL-FIX), as a case study. The relevant and observable characteristics in the model are the annualized bleeding rate (ABR), baseline age, and sex of patients.

Seventh, we present the described methods and their results by applying them to the haemophilia B model.

2 Methods

2.1 Decision Rules

A decision rule, denoted as $\pi(X)$, is a function of a matrix of patient characteristics X that returns a vector of treatment assignments for all patients. $\pi(X)$ determines which patient $i \in \{1,2,\ldots,N\}$ should receive which treatment $j \in \{1,2,\ldots,M\}$. The expression $\pi(X) = j$ describes a situation in which the decision rule assigns a specific but unspecified treatment scheme j. In cost-effectiveness analysis, the value of treatment j for a specific patient i can be defined as the patient's net monetary benefit, NMB_{ij}, which is calculated by multiplying the effectiveness E_{ij} with the willingness to pay, WTP, and subtracting the costs, C_{ij} [23, 24].

$$\mathrm{NMB}_{ij} = E_{ij} * \mathrm{WTP} - C_{ij}$$

Traditional cost-effectiveness analysis applies a uniform treatment rule $\bar{j}$ to all patients [1, 2]. $\pi = \bar{j}$ indicates that each patient is assigned the same treatment, independently of their individual characteristics. The value $V(\bar{j})$ of a uniform treatment rule $\pi = \bar{j}$ is the expected value of individual net monetary benefit, $\mathrm{NMB}_{i\bar{j}}$, under that treatment, given the matrix of patient characteristics X in the population. The expected value of $\mathrm{NMB}_{i\bar{j}}$ is determined by the conditional expectation $E\left[\mathrm{NMB}|X = x, \bar{j}\right]$ and the joint density of observable characteristics $p(x)$ in the population. In a simulation setting, the expected value of a uniform treatment $V(\bar{j})$ can be calculated as the mean $\mathrm{NMB}_{i\bar{j}}$ across all patients.[1]

$$\begin{aligned} V(\bar{j}) &= E\left[\mathrm{NMB}_{i\bar{j}}\right] \\ &= \int E\left[\mathrm{NMB}|X = x, \pi = \bar{j}\right] p(x)\mathrm{d}x \\ &\simeq \frac{1}{N}\sum_{i=1}^{N} E\left[\mathrm{NMB}|X = x_i, \pi = \bar{j}\right] \end{aligned}$$

The optimal treatment in a uniform treatment strategy $\bar{j}^*$ (* denotes an optimal treatment strategy) is chosen to maximize the expected NMB when all patients receive the same treatment. In regular (non-individualized) cost-effectiveness analysis, $\bar{j}^*$ is simply the cost-effective treatment strategy.

$$\bar{j}^* = \underset{\bar{j}}{\operatorname{argmax}} V(\bar{j})$$

An ITR $\pi(X)$ that allows different treatments to be assigned to individual patients can be described by a set of IF statements based on threshold values ρ, which depend on patient characteristics X. We denote specific patient characteristics $k \in \{1,2,\ldots,K\}$ as $x^{(k)}$. In some but not all instances, $\pi(X)$ can be described by a combination of multiple IF statements based on multiple characteristics with different threshold values ρ for patient characteristics X. For example, an ITR $\pi(X)$ for two relevant patient characteristics $x^{(1)}, x^{(2)}$ that sort patients into treatments $j \in \{1, 2, 3\}$ could follow the structure below:

1. If $x_i^{(1)} \geq \rho^{(1)}$ and $x_i^{(2)} < \rho^{(2)}$, patient i is treated with $j = 1$
2. If $x_i^{(1)} \geq \rho^{(1)}$ and $x_i^{(2)} \geq \rho^{(2)}$, patient i is treated with $j = 2$
3. If $x_i^{(1)} < \rho^{(1)}$, patient i is treated with $j = 3$

The value of an ITR $V(\pi(X))$ is the expected NMB for the individual treatments $\pi(X) = j$ assigned to each patient based on all patients' observable characteristics X.

$$\begin{aligned} V(\pi(X)) &= \int E\left[\mathrm{NMB}|X = x, \pi(X) = j\right] p(x)\mathrm{d}x \\ &\simeq \frac{1}{N}\sum_{i=1}^{N} E\left[\mathrm{NMB}|X = x_i, \pi(X) = j\right] \end{aligned}$$

The incremental value $\Delta(\pi(X))$ of an ITR compared with that of the most cost-effective uniform treatment $\bar{j}^*$ is the difference between the values of the two treatment strategies.

$$\Delta(\pi(X)) = V(\pi(X)) - V(\bar{j}^*)$$

The optimal ITR $\pi^*(X)$ maximizes the average NMB across the treated population by assigning each patient the most cost-effective treatment.

$$\pi^*(X) = \underset{\pi(X)}{\operatorname{argmax}} V(\pi(X))$$

The Expected Value of Individualized Care (EVIC) has been defined in the literature as the incremental value of the optimal ITR $\pi^*(X)$ compared with that of the optimal uniform treatment strategy $\bar{j}^*$ [1, 2].

$$\mathrm{EVIC}(\pi^*(X)) = \Delta(\pi^*(X)) = V(\pi^*(X)) - V(\bar{j}^*)$$

The optimal ITR $\pi^*(X)$ can also serve as a benchmark for ITRs $\pi(X)$. The regret $R(\pi(X))$ of an ITR $\pi(X)$ measures how much the ITR $\pi(X)$ deviates from the optimal $\pi^*(X)$ in terms of NMB and is calculated as the difference between

---

[1] The integral represents the population expectation with respect to the joint distribution of characteristics. The Monte Carlo sample average is a consistent estimator of this expectation and asymptotically converges to it as the simulation sample size increases under standard assumptions. In the simulation context, the sample average is exact because we compute the expected NMB over a finite set of simulated patients.

the value of the optimal ITR $V(\pi^*(X))$ and the value of the assessed ITR $V(\pi(X))$. As $V(\pi^*(X)) \geq V(\pi(X))$, regret $R(\pi(X))$ is never negative. While regret is not relevant for decision-making, it is an important technical metric for understanding how close to a theoretical optimum a given ITR is, and is therefore often used to tune machine learning models [16, 25].

$$R(\pi(X)) = V(\pi^*(X)) - V(\pi(X)) \geq 0$$

2.2 Designing Data-Driven Individualized Treatment Rules in Microsimulation Data

In this study, we use machine learning techniques to design ITRs by re-analysing simulated patient-level data from a microsimulation model of haemophilia B. This analysis is based on efficacy parameters derived from a published randomized clinical trial and secondary sources.

The main strength of microsimulations for designing ITRs is that they compute counterfactual outcomes under alternative treatments in the same population and thus provide conditional average treatment effects without the need to employ causal machine learning techniques as in observational data settings. Thanks to this property of microsimulation models, standard predictive statistical learning methods can be used to identify predictors of NMB under alternative treatments and to derive ITRs based on these predictors.

Let $f(X,j)$ be a statistical model's conditional expectation function of the NMB. For regression models, this conditional expectation function is a parametric function of patient characteristics and treatment. For tree-based models, the conditional expectation function is a set of average NMBs under alternative treatments for subgroups with different combinations of patient characteristics X. The conditional expectation function can be used to calculate the fitted $\widehat{\text{NMB}}_{ij}$ for each simulated patient i under all alternative treatments j.

$$\widehat{\text{NMB}}_{ij} = f(x_i, j)$$

The ITR $\pi(x_i)$ assigns to each patient the treatment $\widehat{\pi}(x_i)$ with the highest expected individual net monetary benefit $\widehat{\text{NMB}}_{ij}$ based on the conditional expectation function $f(x_i, j)$.

$$\widehat{\pi}(x_i) = \underset{j}{\operatorname{argmax}} \widehat{\text{NMB}}_{ij}$$

In microsimulation data, the modelled "true" NMB_{ij} under all counterfactual therapies j are recorded, and the "truly" optimal individual treatment $\widehat{\pi}^*(x_i)$ can be determined. This benchmark does not necessarily represent the perfectly individualized treatment for observed heterogeneity in real patients but only the perfectly individualized treatment for the modelled heterogeneity.

$$\widehat{\pi}^*(x_i) = \underset{j}{\operatorname{argmax}} \text{NMB}_{ij}$$

This information enables the estimation of the value of the empirical ITR $V\left(\widehat{\pi}(X)\right)$ and the value of the "truly" optimal ITR $V(\pi^*(X))$ in the microsimulation data. Note that the value of the ITR is the expected value of the "true" NMB under the treatments that were favoured by the conditional expectation function, not the expected value of the conditional expectation function.

$$V(\hat{\pi}(X)) = \frac{1}{N}\sum_{i=1}^{N} E[\text{NMB} \mid X = x_i,\ \pi(X) = \hat{\pi}(x_i)]$$

$$V(\pi^*(X)) = \frac{1}{N}\sum_{i=1}^{N} E[\text{NMB} \mid X = x_i,\ \pi(X) = \pi^*(x_i)]$$

The incremental value of an ITR, $\Delta\left(\widehat{\pi}(X)\right)$, can be measured by the distance of the value of the ITR, $V\left(\widehat{\pi}(X)\right)$, from the value of the optimal uniform treatment strategy, $V\left(\bar{j}^*\right)$.

$$\begin{aligned}\Delta(\hat{\pi}(X)) &= V(\hat{\pi}(X)) - V\left(\bar{j}^*\right)\\ &= \frac{1}{N}\sum_{i=1}^{N} E\left[\text{NMB}|X = x_i, \pi(X) = \hat{\pi}\left(x_i\right)\right] - E\left[\text{NMB}|X = x_i, \pi = \bar{j}^*\right]\end{aligned}$$

2.3 Empirical Models

There are many potentially viable approaches to derive ITRs via outcome modelling in microsimulation patient data. In this study, we explore LASSO regression, classification trees, and policy trees to demonstrate how standard statistical learning techniques can be used to derive ITRs from microsimulation patient data and examine key considerations for selecting an appropriate approach. We chose these three empirical models, all capable of assigning treatments, because they represent widely used off-the-shelf machine learning approaches, yet vary sufficiently in their properties, offering differing advantages and disadvantages.

LASSO regression: Regression models are flexible tools for analysing patient characteristics and predicting the treatment with the highest NMB for individual patients. LASSO regression [26] is a machine learning regularization technique that reduces the coefficients of covariates with low predictive power, and allows the selection of a subset of covariates with non-zero coefficients, simplifying the decision rule.

We first select covariates with non-zero coefficients via LASSO and then use these covariates in a nonregularized regression model. In the candidate set of covariates, we include treatment dummies $\boldsymbol{T}$, binary patient characteristics $\boldsymbol{X}_b$, continuous patient characteristics $\boldsymbol{X}_c$, polynomials of continuous characteristics up to the 5th degree $\boldsymbol{Z}^d = poly(\boldsymbol{X}_c, d)$, interactions between the treatment dummies and binary patient characteristics $\boldsymbol{T} \cdot \boldsymbol{X}_b$, interactions between the treatment and continuous patient characteristics $\boldsymbol{T} \cdot \boldsymbol{Z}_c^d$, and interactions between all patient characteristics $\boldsymbol{X} \cdot \boldsymbol{X}$, excluding polynomial terms. This regression model is described by the following equation:

$$\mathrm{NMB} = \beta_0 + 2T\boldsymbol{\beta} + \boldsymbol{X}_b\boldsymbol{\gamma}_b + \sum_{d=1}^{5} \boldsymbol{Z}_c^d\boldsymbol{\gamma}_c^d + \left(\boldsymbol{T} \cdot \boldsymbol{X}_b\right)\boldsymbol{\delta}_b + \sum_{d=1}^{5} \left(\boldsymbol{T} \cdot \boldsymbol{Z}_c^d\right)\boldsymbol{\delta}_c^d + (\boldsymbol{X} \cdot \boldsymbol{X})\boldsymbol{\theta} + \varepsilon$$

We tuned the LASSO model via tenfold cross-validation[2] with mean squared error as the loss function and selected the covariates with non-zero coefficients at the penalty parameter λ that minimizes the cross-validation error.[3] Once the regression coefficients β_0, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}_b$, $\boldsymbol{\gamma}_c^d$, $\boldsymbol{\delta}_b$, $\boldsymbol{\delta}_c^d$, and $\boldsymbol{\theta}$ are estimated, $\widehat{\mathrm{NMB}}_{ij}$ can be predicted for all patients in the simulated data. The treatment with the highest $\widehat{\mathrm{NMB}}_{ij}$ for patient i is the treatment assigned by the ITR based on the linear regression function. This regression function is an ITR $\pi(X)$ in the sense that for any given combination of patient characteristics, a clear prediction of the optimal treatment is made.

Classification tree: Recursive partitioning can be used to create classification trees that attempt to correctly classify observations by repeatedly splitting the patient population into a treatment at patient-characteristic threshold values, creating subgroups in the process [29]. Different combinations of patient characteristics and threshold values ρ are algorithmically tested to construct a tree that minimizes the classification error for each patient in the dataset. In this context, a class refers to the optimal treatment for any given patient, meaning that if a patient receives optimal treatment j, this is also their class. The threshold values for splitting are determined by the Gini impurity G in the data for combinations of patient characteristics and threshold values [30]. In this application, the Gini impurity coefficient G is calculated as 1 minus the sum of the squared proportions of each class $P[j^*]^2$ in the subgroup.

$$G = 1 - \sum_{j}^{M} P[j^*]^2$$

If every patient in a subgroup belongs to the same class, the Gini impurity reaches a minimum of 0. As each split creates two subgroups, the Gini impurity coefficient is calculated for each of the two subgroups following the split, G_1 and G_2. Then, a weighted Gini impurity G_w with weights given by sample sizes in the two subgroups, N_1 and N_2, is calculated.

$$G_\mathrm{w} = \frac{N_1}{N}G_1 + \frac{N_2}{N}G_2$$

The split threshold value ρ that minimizes the weighted Gini impurity G_w is chosen for each split. Within each node, the optimal treatment for most patients (the majority class) is assigned to all patients in that node. By repeating this process multiple times, a tree of nodes, each corresponding to a treatment assignment, is created. This classification tree is the graphical representation of an ITR $\pi(X)$. For this application, we generated a classification tree with a depth of two layers to keep the complexity of the corresponding ITR low.[4] This choice was not based on the empirical properties of the model but rather serves as an example of a case study with the goal of creating a decision rule with limited user-defined complexity.

Policy tree: Bonander and Svensson [10], Glynn et al. [12], and Hattab et al. [11] have previously applied the policy tree algorithm [16] to develop ITRs with the aim of assigning each patient the treatment with the highest value for money for this individual. As with classification trees, policy trees are generated via recursive partitioning, but the algorithm generates nodes and sets split thresholds ρ that minimize regret and thereby maximize the value of a given policy tree, rather than minimizing the Gini impurity as the classification tree algorithm does. Minimizing regret is equivalent to maximizing the incremental value of the ITR [25]. For the present application, we generated a policy tree with a depth of two layers, in order to have a similar level of complexity when comparing it to the classification tree.[5]

2.4 Assessing the Impact of Parameter Uncertainty on Individualized Treatment Rules

The influence of parameter uncertainty on cost-effectiveness outcomes also affects ITRs. Therefore, variations in the optimal ITRs due to parameter uncertainty should be examined via probabilistic analysis [21]. To ensure that the assessments of the parameter uncertainty of cost-effectiveness

---

[2] K-fold cross-validation refers to the splitting of the data into K subsamples, training the model on K − 1 subsamples, and testing the performance of the model on the final subsample. This process is performed iteratively, so that each of the K subsamples is used $K - 1$ times as a training subsample, and once as a testing subsample. These iterations are aggregated to determine the best parameter values as a performance metric [27]. K-fold cross-validation helps train machine learning models on new, unseen data, and thereby reduces overfitting while still making full use of the data that is currently available.

[3] The “glmnet” package in R was used for this [28].

[4] The “rpart” package in R was used for this [30].

[5] The “policytree” package in R was used for this [31]. Unlike classification trees generated via the “rpart” package, policy trees are always symmetrical, meaning that with each additional layer of depth, a split is introduced for each node in the prior layer. Even at a similar level of depth, classification trees and policy trees will differ to some degree in terms of structure.

outcomes and ITRs are congruent, they should be performed using the same random parameter draws. The assessment of the parameter uncertainty of ITRs can be implemented as a direct extension of the probabilistic analysis functionality in the microsimulation model.

We assume that P populations of N patients are simulated, with random population-level parameters resampled for each simulated population, so that we draw a unique set of parameter values for each population. The resulting simulated dataset with $P \times N$ simulated patients can be used to design the ITR with the highest expected value under consideration of uncertainty of the model inputs and outputs. We can also design a new ITR in each simulated population P and describe the variation in ITRs to assess the impact of parameter uncertainty on the ITRs themselves.

Variation in ITRs can occur in two forms: structural variation or ITR parameter value variation. It occurs based on the counterfactual patient outcomes that are used as inputs into the empirical models for deriving ITRs. Structural variation refers to changes in the included patient characteristics and, for decision trees, also to changes in the order in which the patient characteristics are considered. ITR parameter value variation refers to differences in estimated coefficients for regression models or differences in estimated threshold values for decision trees. In our usage of the terms, structural variation and ITR parameter value variation both refer specifically to variation in the ITRs themselves, not the underlying microsimulation model.

Allowing for structural variation makes summarizing parameter uncertainty complex and difficult to interpret. We thus focus purely on the impact of parameter uncertainty on ITR parameter value variation. We design the ITR structure using data from all the simulated populations and then estimate the optimal ITR parameter values while keeping the structure fixed. The resulting ITR parameter value variation can be quantified, for example, using a credible interval. If 95% of the ITR parameter values fall within a specific range, this provides a useful measure of uncertainty stemming from parameter variability.

2.5 Criteria for Model Assessment

We formulated three criteria for assessing the three methods for designing ITRs.

Incremental value: The ITR's incremental value $\Delta\left(\widehat{\pi}(X)\right)$ should be as large as possible. When the incremental value of an ITR reaches the possible maximum, it becomes equivalent to the EVIC.

Computational speed: From a practical perspective, tools that require less time to run on widely available hardware grant an advantage when performing cost-effectiveness analyses. We measure the time needed to apply each of the three empirical models on a standard laptop[6] for our case study. The runtime of the empirical models will differ depending on the exact hardware, but the recorded times serve as a reasonable benchmark.

Interpretability: To be useful in practice, an ITR should not only assign treatments to each individual on the basis of individual characteristics. It should also be possible to express it as a set of easily interpretable rules, determining which patients would receive which treatments on the basis of characteristic value thresholds. The corresponding rule thresholds should be determined by the empirical model itself and optimized for any number of characteristics across the simulated patient data. It is a further advantage if the complexity of the rules can be determined by the user who is technically implementing the ITRs.

2.6 Stochastic Uncertainty

In the microsimulation model, due to stochastic uncertainty, certain random events occur only for some patients or at random times based on event probabilities [21]. Due to randomness, even patients with identical characteristics and/or treatments can experience different outcomes. If these different outcomes cannot be predicted with certainty from included patient characteristics, then no ITR is able to assign the optimal treatment to all patients, despite being familiar with all counterfactuals. Random events with impacts on outcomes in the model therefore influence the incremental value and regret of ITRs.

2.7 Case Study: Haemophilia B Microsimulation

We adapt a previously published microsimulation cost-effectiveness model of multiple treatments for haemophilia B in Germany [22] as a case study to illustrate the described methods. In this model, the gene therapy ED is compared with prophylactic administration of EHL-FIX. When the effectiveness of ED wanes, patients switch treatment to prophylactic EHL-FIX. We do not alter the underlying model for the purposes of this case study but add two additional treatment options: on-demand treatment with FIX and ED followed by on-demand treatment if effectiveness wanes ("ED into on-demand"). These two treatments are not standard practice for the treatment of severe haemophilia B in Germany and hence were not included in the original analysis. However, we included them to compare more than two different treatment options as a demonstration of

---

[6] Lenovo ThinkPad with an Intel(R) Core (TM) i7-1165G7 @ 2.80 GHz processor and 16 GB of memory.

the methods used in this study. A WTP of EUR 50,000 per quality-adjusted life year (QALY) was assumed in this study.

The microsimulation contains three relevant and observable patient characteristics that impact the outcomes of treatments: the ABR, baseline age, and sex (which influences weight and thus FIX dosage). An ITR can therefore assign treatments based on these three patient characteristics.

In the haemophilia B case study [22], 2000 simulations are run with random draws of the probabilistic parameters. One hundred patients are generated per simulation, leading to 200,000 patients in total. The same 200,000 patients are used to compare each treatment strategy. The initial ITR is designed to optimize individualized treatment across all 200,000 patients and provides the structure that is used to subsequently assess the uncertainty of the ITR parameter values. We use grid searches [32] to determine which combination of ITR parameter values is optimal within the defined structure for each simulation, producing 2000 combinations of locally optimal ITR parameter values. We describe the 95% credible range of the individual ITR parameter values and depict the density of the ITR parameter value combinations graphically to understand the impact of parameter uncertainty on the ITRs.

In the original case study microsimulation model [22], there are two potentially relevant sources of stochastic uncertainty. First, for each patient treated with ED, there is a small probability of initial treatment failure. Second, in each cycle, there is a probability of death for every patient, based on ABR, age, and sex. However, in the base case of the case study, death is not modelled stochastically as a binary, random event in each cycle, as it was expected that stochastic modelling of death would increase the variability in patient outcomes, requiring a larger simulated population, without having a relevant impact on mean patient outcomes. Instead, survival status in each cycle is implemented as a survival probability based on the probabilities of death in all previous cycles. To show how stochastic uncertainty influences the ability of an ITR to assign optimal treatment to patients, we present two additional scenarios. In the first scenario, we implement death as a binary, random event to fully capture the stochastic uncertainty around the event of death. In the second scenario, we remove the randomly modelled treatment failure with ED from the microsimulation model, decreasing stochastic uncertainty. We compare the incremental value and regret of ITRs emerging from the model's base case and the scenarios to show how increasing or decreasing the stochastic uncertainty covered by the microsimulation model influences the ability of the ITRs to optimally assign treatments.

3 Results

3.1 Comparison of Empirical Models

All methods for deriving ITRs yield an expected NMB that is much closer to the optimal individualized treatment than uniform treatment strategies are (Table 1). The model produces a negative expected NMB for all treatment strategies, and the most cost-effective strategy exhibits the least negative expected NMB. The first four rows show that under uniform treatment, when all patients receive the same treatment, on-demand treatment is the most cost-effective treatment strategy, with an expected NMB of EUR − 1,657,944. However, making treatment decisions at the population level is not cost-effective when compared to individualized treatment. The highest expected NMB is obtained by assigning the optimal treatment to each patient individually, leading to an expected NMB of EUR − 1,392,709. The incremental value of optimal individualized treatment, EUR 265,235, is equivalent to the EVIC. The ITRs based on the three empirical models all produce very similar expected NMBs ranging from EUR − 1,450,973 (LASSO regression) to EUR − 1,447,264 (policy tree).

Table 2 presents a comparison of the three empirical models based on the three criteria we established. All models achieve approximately the same incremental value compared to uniform treatment, ranging from EUR 206,971 (LASSO regression) to EUR 210,681 (policy tree), with a higher value indicating a better individualization of care. These values are similar, and we consider all three models roughly equivalent in terms of incremental value.

The relative differences in computational speed are considerable. The classification tree requires approximately 3 s, the LASSO regression requires 26 s, and the policy tree requires 292 s. These relative differences can translate into large absolute differences in the presence of more features, complex decision trees, or larger data.

In terms of the ability to generate decision rules that determine treatment assignment, we find regression approaches to be unsuitable. Regressions can be used to make individual predictions by multiplying patient characteristics with the estimated coefficients and thus determine the optimal treatment for any combination of patient characteristics. In theory, it would be possible to compute on the basis of the estimated coefficients at what threshold values the optimal treatment changes. These computations could be used to define the decision rules that determine which patient subgroups receive which treatments. However, when there are many coefficients, this becomes increasingly complex. Classification and policy trees, by design, split populations into subsamples, and represent decision rules as decision trees, which are easily interpretable.

Table 1 Comparison of the discounted mean outcomes depending on the treatment strategies

Treatment strategy	QALYs	Costs (EUR)	NMB (EUR)	Regret (EUR)	Inc. value (EUR)
Uniform treatment					
On-demand	11.62	2,238,940.48	– 1,657,944.49	265,235.39	–
ED into on-demand	14.26	2,499,796.58	– 1,786,915.54	394,206.44	–
Prophylaxis	16.51	6,427,659.72	– 5,601,952.15	4,209,243.05	–
ED into prophylaxis	17.13	5,308,144.04	– 4,451,697.16	3,058,988.06	–
Optimal individualized treatment	12.97	2,034,709.87	– 1,392,709.10	0.00	265,235.39
Individualized treatment rules					
LASSO regression	12.83	2,092,625.56	– 1,450,973.32	58,264.22	206,971.17
Classification tree	12.83	2,090,196.74	– 1,448,847.51	56,138.41	209,096.98
Policy tree	12.70	2,082,204.32	– 1,447,263.85	54,554.75	210,680.64

Rows 1–4 show the mean outcomes per patient when all patients in the population receive the same treatment. Rows 5–8 show the mean outcomes when patients are assigned treatments at the individual level

ED etranacogene dezaparvovec, *Inc.* incremental, *LASSO* least absolute shrinkage and selection operator, *NMB* net monetary benefit, *QALY* quality-adjusted life year

Table 2 Comparison of three methods for assigning treatments to individual patients on the basis of defined criteria

Criterion	LASSO regression	Classification tree	Policy tree
Criterion #1: incremental value	EUR 206,971	EUR 209,097	EUR 210,681
Criterion #2: computational speed	26 s	3 s	292 s
Criterion #3: interpretability	Complex	Simple	Simple

LASSO least absolute shrinkage and selection operator

In this comparison, we find that classification trees show comparable predictive performance, the shortest runtime, and are easy to interpret. Therefore, we only apply classification trees for the remaining illustration of the case study.

3.2 Individualized Treatment Rules via Classification Tree

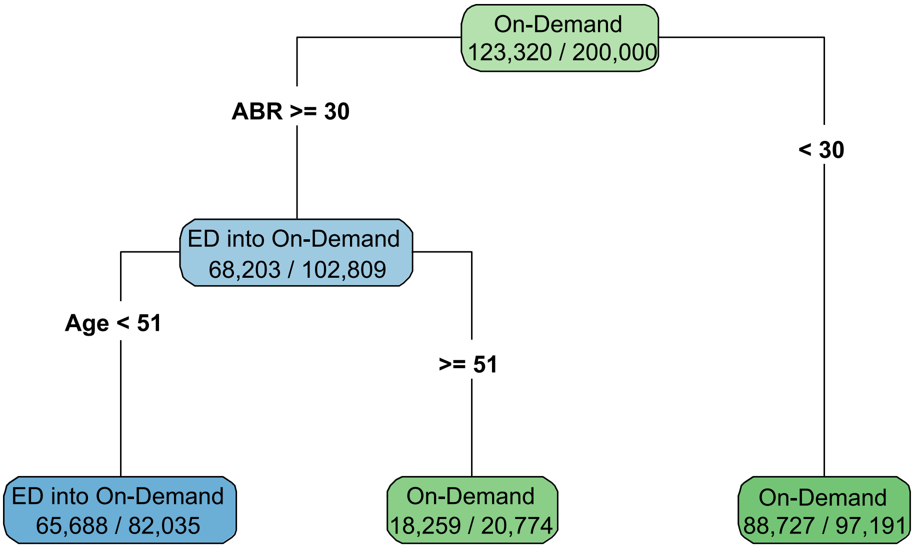
Figure 1 shows the classification tree that was generated via recursive partitioning. The iterative logic can be traced by following the branches. If only one treatment were assigned, then on-demand treatment would be cost-effective on average. However, this approach would only be the cost-effective treatment for 123,320 of 200,000 patients. As a first step, patients are split into two subsamples: patients with an ABR ≥ 30 receive ED followed by on-demand treatment, and patients with an ABR < 30 receive on-demand treatment. In the second step, patients who were assigned ED followed by on-demand treatment in the first step are again split into two subsamples: patients with an age < 51 years receive ED followed by on-demand treatment, and patients with an age ≥ 51 years receive on-demand treatment. Both of these steps are intuitively plausible. The expensive gene therapy ED is cost-effective for patients with a higher ABR and thus greater disease severity. However, among patients with high disease severity, high upfront costs are only worthwhile for patients below a certain age, who thus have a greater life expectancy. The recursive partitioning algorithm generating the classification tree did not select sex as a characteristic, as sex was less important than ABR or age for optimal treatment assignment. This tree can also be described in terms of three IF statements:

1. If ABR ≥ 30 and age < 51, patient is treated with ED into on-demand.
2. If ABR ≥ 30 and age ≥ 51, patient is treated on-demand.
3. If ABR < 30, patient is treated on-demand.

3.3 Parameter Uncertainty

The impact of the parameter uncertainty on the uncertainty of the threshold values is considerable. Across the 2000 simulations, we observe a 95% credible range of 21–46 for the optimal ABR threshold value and a range of 42–56 for the optimal age threshold value, holding the decision tree in Fig. 1 constant. The density graph of optimal threshold values shows that the greatest density of optimal threshold values is around an ABR of 30 and a patient age of 51 (Fig. 2). However, these two values are not independent: we find a slight negative correlation between the optimal ABR and the age threshold, indicating that the thresholds are not independent of each other. This means that changing one

Fig. 1 Individualized treatment rules generated as a classification tree via recursive partitioning. *ABR* annualized bleeding rate, *ED* etranacogene dezaparvovec



threshold value has an impact on the optimal value of the other threshold.

3.4 Stochastic Uncertainty

The random event scenarios are shown in Table 3. In the microsimulation model's base case, the incremental value of the classification tree ITR is EUR 209,097.

Scenario 1: If death is modelled randomly, the incremental value decreases to EUR 181,184 and regret increases to EUR 240,448. With optimal individualized treatment assignment, patients with random early deaths are not assigned ED, as it carries high upfront costs and is only cost-effective when patients gain sufficient life expectancy from the treatment. With knowledge of all counterfactuals, we could, in theory, avoid assigning ED to those patients who die early, but these early deaths cannot be predicted via the observable patient characteristics. Therefore, an ITR that assigns treatments on the basis of patient characteristics is less successful in individualizing treatment and worse at minimizing regret. However, the expected NMB of the classification tree ITR itself decreases by only EUR 13,667, and it remains cost-effective versus uniform treatment assignments.

Scenario 2: If treatment failure is removed from the model, i.e. there is no stochastic uncertainty in the model, the incremental value increases to EUR 256,253. The regret

Fig. 2 A density graph showing the distribution of optimal threshold value combinations across the 2000 simulations. The dark red colour indicates a high density of observations, whereas a bright yellow colour represents a low density. The outliers not contained within these densities are depicted by black dots. *ABR* annualized bleeding rate

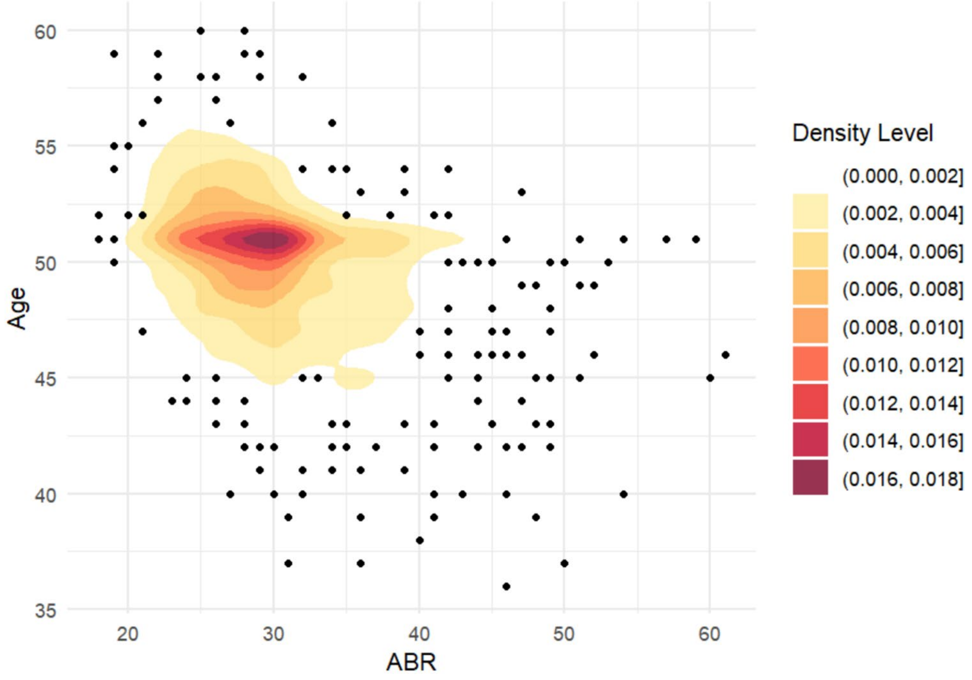


Table 3 Comparison between the base case and two scenarios in terms of the net monetary benefit (EUR)

	Base case	Scenario 1: random death	Scenario 2: no treatment failure
Optimal individualized treatment strategy (a)	− 1,392,709	− 1,222,067	− 1,365,318
Optimal uniform treatment strategy (b)	− 1,657,944	− 1,643,699	− 1,643,699
ITR via classification tree (c)	− 1,448,848	− 1,462,515	− 1,387,446
Regret (a–c)	56,138	240,448	22,128
Incremental value (b–c)	209,097	181,184	256,253

ITR individualized treatment rule

is reduced to EUR 22,128, as the ITR is able to almost perfectly assign individualized treatments. However, as treatment failure is a realistic modelling assumption and carries a random element with it, this implies an overestimation of the incremental value and an underestimation of the regret of this ITR.

4 Discussion

This study examines the use of ITRs in cost-effectiveness models, and three empirical approaches are compared using haemophilia B microsimulation as a case study [22]. We find classification trees to be a suitable method for three reasons. First, it produces ITRs comparable to those of the other tested empirical models in terms of incremental value and regret. Second, it has low computational demands. Third, its decision tree structure allows ITRs to be expressed in an easily interpretable manner [29]. We believe that the viability of classification trees is generalizable to other microsimulation models: its computational efficiency holds even for large datasets, and the interpretation of a classification tree always follows the same principle, no matter what microsimulation model it is applied to. The incremental value of the derived ITR depends on the complexity of the classification tree, but this can be controlled by the modeller: a deeper classification tree allows for better ITRs at the cost of greater complexity. These findings can and should be confirmed via other microsimulation models.

We consider the ability to express an ITR in an easily interpretable manner important for the translation of the model results into a format that can be considered for real-world decision-making. Our case study demonstrates this facet: on the basis of a decision tree consisting of only two threshold values and three IF statements, treatment can be assigned to all simulated patients, leading to a mean NMB higher than that of any uniform treatment strategy. While a regression model may allow for the calculation of an expected NMB based on patient characteristics, it would be difficult to formulate simple cost-effectiveness treatment recommendations on this basis. We further consider it advantageous for the empirical model to derive an optimal ITR structure in a data-driven manner, as the manual specification of an ITR structure including many patient characteristics may be prone to bias [33].

When real patient data are used, classification trees can be unsuitable because they adhere too closely to the data, which can lead to overfitting, particularly with deeper trees, and may exhibit biased variable selection [34]. However, with simulated patient data, this is not a drawback. The correlations between patient characteristics and outcomes stem from simplifying yet realistic modelling assumptions about disease and treatment effects. It is therefore desirable for any ITR to fit itself to the data closely and capture the impact of these modelling assumptions as completely as possible. When microsimulation model results lack real-world generalizability, the ITR will also lack real-world generalizability.

In our case study, the classification tree-based ITR assigns either on-demand or ED followed by on-demand treatment if effectiveness wanes, based on the ABR and age threshold values. These ITR parameter values are subject to considerable parameter uncertainty. We also find it important to consider the role of stochastic uncertainty in model events, as highlighted by two scenarios that consider random death events and random treatment failures. These scenarios show that if stochastic uncertainty is not sufficiently considered in the underlying microsimulation model, the incremental value of ITRs may be overestimated. Certain modelling simplifications that reduce stochastic uncertainty, which may be acceptable when only considering mean outcomes [35], become an issue when a microsimulation is used to derive ITRs. With such a purpose, making microsimulation models representational of individual patient pathways and avoiding abstract expected values of outcomes becomes even more crucial. While our case study is limited to three patient characteristics, our approaches allow us to derive ITRs with more complex patient data when available.

The approach described in this paper has various limitations. While we compare ITRs based on three empirical models, there may be other suitable approaches. However, such approaches can be compared based on our suggested criteria, and a gold standard may eventually be established. As determining the uncertainty of the ITRs is based on probabilistic analysis, we inherit the limitations of this approach: only parameter uncertainty is captured by this [21]. For the parameter uncertainty, we assume that the underlying effect estimates are unbiased, which may not be the case. In the case study, as the input parameters are derived from various secondary sources, we do not capture the correlations between these parameters [36].

Stochastic uncertainty in our microsimulation case study was represented by random events that varied in their occurrence or timing between patients. Alternatively, it can be approached in a different manner: Glynn et al. [12] simulated each individual patient 10,000 times for 8746 real-world patient profiles to optimize the ITR, leading to 87.46 million simulations, and then took the average of each patient's simulations. The equivalent simulated population of 200,000 in this case study, multiplied by the four evaluated treatments to generate all counterfactuals, would require 8 billion simulations. To reduce the simulated population size, probabilistic uncertainty can also be considered at the individual level, as Glynn et al. did. Considering both stochastic and probabilistic uncertainty at the individual level in microsimulations could be an interesting avenue for further research.

Our ITRs focus purely on expected NMB, which has some drawbacks. Calculating the expected NMB requires an explicit WTP threshold, which does not exist in many countries. However, the described methods also allow for the optimization of other outcome parameters, e.g. QALYs or costs, when they are relevant for decision-making. In future studies, the methods could be extended, e.g. to consider aspects of equity. For example, an ITR could be designed to maximize NMB within constraints to avoid sacrificing too many QALYs at the individual patient level. Alternatively, the ITR could consider the Gini index [37] and keep inequality below a target value. ITRs based on such extensions would allow for more balanced optimization approaches and might be more suitable for informing real-world decision-making than the examples given in this article can. We believe that there are many additional avenues to explore, ranging from different empirical models to types of analysis, that would allow ITRs to become a common method for interpreting the results of microsimulation models.

Author contributions All authors conceived and designed the study. NM implemented the described methods in R and drafted the manuscript. MP made important contributions to the underlying methods. ACQG, MP, and MS performed detailed reviews of the initial manuscript drafts and approved the final version.

Funding Open access funding provided by University of Basel. No funding was received for conducting this study.

Declarations

Conflict of interest Mark Pletscher is an employee of Novartis Pharma AG and holds stock of Novartis Pharma AG. The other authors report no conflicts of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Availability of data and material All data generated or analysed during this study are included in this published article or are available from the corresponding author (NM) on reasonable request for non-commercial purposes.

Code availability The microsimulation model and the methods described in this article are implemented in R, and the full code is available at https://github.com/NiklausMeier/HB_CE_microsimulation.

References

1. Basu A, Meltzer D. Value of information on preference heterogeneity and individualized care. Med Decis Mak. 2007;27:112–27.
2. Van Gestel A, Grutters J, Schouten J, Webers C, Beckers H, Joore M, et al. The role of the expected value of individualized care in cost-effectiveness analyses and decision making. Value Health. 2012;15:13–21.
3. Espinoza MA, Manca A, Claxton K, Sculpher MJ. The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. Med Decis Mak. 2014;34:951–64.
4. Shields GE, Clarkson P, Bullement A, Stevens W, Wilberforce M, Farragher T, et al. Advances in addressing patient heterogeneity in economic evaluation: a review of the methods literature. Pharmacoeconomics. 2024;42:737–49.
5. Chakraborty B, Moodie EEM. Statistical methods for dynamic treatment regimes: reinforcement learning, causal inference, and personalized medicine [Internet]. New York: Springer New York; 2013 [cited 2025 Feb 5]. https://doi.org/10.1007/978-1-4614-7428-9.
6. He Q, Zhang S, LeBlanc ML, Zhao Y-Q. Estimating individualized treatment rules by optimizing the adjusted probability of a longer survival. Stat Methods Med Res. 2024;33:1517–30.

7. Brothers KB, Rothstein MA. Ethical, legal and social implications of incorporating personalized medicine into healthcare. Pers Med. 2015;12:43–51.
8. Asada Y. A framework for measuring health inequity. J Epidemiol Community Health. 2005;59:700–5.
9. São José JMS, Amado CAF, Ilinca S, Buttigieg SC, Taghizadeh Larsson A. Ageism in health care: a systematic review of operational definitions and inductive conceptualizations. Gerontologist. 2019;59:e98-108.
10. Bonander C, Svensson M. Using causal forests to assess heterogeneity in cost-effectiveness analysis. Health Econ. 2021;30:1818–32.
11. Hattab Z, Moler-Zapata S, Doherty E, Sadique Z, Ramnarayan P, O'Neill S. Exploring heterogeneity in the cost-effectiveness of high-flow nasal cannula therapy in acutely ill children—insights from the step-up first-line support for assistance in breathing in children trial using a machine learning method. Value Health. 2025;28:60–9.
12. Glynn D, Giardina J, Hatamyar J, Pandya A, Soares M, Kreif N. Integrating decision modeling and machine learning to inform treatment stratification. Health Econ. 2024;33:1772–92.
13. Zhang Y, Kreif N, Gc VS, Manca A. Machine learning methods to estimate individualized treatment effects for use in health technology assessment. Med Decis Mak. 2024;44:756–69.
14. Padula WV, Kreif N, Vanness DJ, Adamson B, Rueda J-D, Felizzi F, et al. Machine learning methods in health economics and outcomes research—the PALISADE checklist: a good practices report of an ISPOR task force. Value Health. 2022;25:1063–80.
15. Athey S, Wager S. Estimating treatment effects with causal forests: an application [Internet]. arXiv; 2019 [cited 2025 Feb 6]. Available from: https://arxiv.org/abs/1902.07409.
16. Athey S, Wager S. Policy learning with observational data. Econometrica. 2021;89:133–61.
17. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Econ. 2006;15:677–87.
18. Graves J, Garbett S, Zhou Z, Schildcrout JS, Peterson J. Comparison of decision modeling approaches for health technology and policy evaluation. Med Decis Mak. 2021;41:453–64.
19. Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: a review. Med Decis Mak. 2011;31:10–8.
20. Zucchelli E, Jones AM, Rice N. The evaluation of health policies through dynamicmicrosimulation methods. Int J Microsimul. 2011;5:2–20.
21. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM modeling good research practices task force-6. Value Health. 2012;15:835–42.
22. Meier N, Fuchs H, Galactionova K, Hermans C, Pletscher M, Schwenkglenks M. Cost-effectiveness analysis of etranacogene dezaparvovec versus extended half-life prophylaxis for moderate-to-severe haemophilia B in Germany. PharmacoEconomics - Open. 2024;8:373–87.
23. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18:341–64.
24. Briggs AH. Decision modelling for health economic evaluation. In: Claxton K, Sculpher M, editors. Oxford: Oxford University Press; 2011.
25. Huber M. Causal analysis: impact evaluation and causal machine learning with applications in R. Cambridge, London: The MIT Press; 2023.
26. Ranstam J, Cook JA. LASSO regression. Br J Surg. 2018;105:1348–1348.
27. Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Ser B Stat Methodol. 1974;36:111–33.
28. Friedman J, Hastie T, Tibshirani R, Narasimhan B, Tay K, Simon N, et al. glmnet: lasso and elastic-net regularized generalized linear models [Internet]. 2008 [cited 2025 Jul 31]. p. 4.1–10. Available from: https://CRAN.R-project.org/package=glmnet.
29. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees [Internet], 1st edn. Routledge; 2017 [cited 2025 Feb 6]. Available from: https://www.taylorfrancis.com/books/9781351460491.
30. Therneau T, Atkinson B. rpart: Recursive partitioning and regression trees [Internet]. 1999 [cited 2025 Feb 6]. p. 4.1.24. Available from: https://CRAN.R-project.org/package=rpart.
31. Sverdrup E, Kanodia A, Zhou Z, Athey S, Wager S. policytree: policy learning via doubly robust empirical welfare maximization over trees [Internet]. 2020 [cited 2025 Feb 6]. p. 1.2.3. Available from: https://CRAN.R-project.org/package=policytree.
32. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res. 2012;13:281–305.
33. Gu J, Oelke D. Understanding bias in machine learning. 2019 [cited 2025 Mar 6]. Available from: https://arxiv.org/abs/1909.01866.
34. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. J Comput Graph Stat. 2006;15:651–74.
35. Flaxman AD, Deason AW, Dolgert AJ, Mumford JE, Sorensen RJ, Eldrenkamp E, et al. Untangling uncertainty with common random numbers: a simulation study. Summer Comput Simul Conf SCSC [Internet]. Bellevue: Society for Modeling and Simulation International (SCS); 2017 [cited 2025 Mar 6]. Available from: http://dl.acm.org/citation.cfm?id=3140096.
36. Naveršnik K, Rojnik K. Handling input correlations in pharmacoeconomic models. Value Health. 2012;15:540–9.
37. Gastwirth JL. The estimation of the Lorenz curve and Gini index. Rev Econ Stat. 1972;54:306.