BMC
Medical Informatics & Decision Making

RESEARCH ARTICLE   **Open Access**

Stratification of coronary artery disease patients for revascularization procedure based on estimating adverse effects

Sebastian Pölsterl[1*], Maneesh Singh[2], Amin Katouzian[1], Nassir Navab[1], Adnan Kastrati[3], Lance Ladic[4] and Ali Kamen[2]

Abstract

Background: Percutaneous coronary intervention (PCI) is the most commonly performed treatment for coronary atherosclerosis. It is associated with a higher incidence of repeat revascularization procedures compared to coronary artery bypass grafting surgery. Recent results indicate that PCI is only cost-effective for a subset of patients. Estimating risks of treatment options would be an effort toward personalized treatment strategy for coronary atherosclerosis.

Methods: In this paper, we propose to model clinical knowledge about the treatment of coronary atherosclerosis to identify patient-subgroup-specific classifiers to predict the risk of adverse events of different treatment options. We constructed one model for each patient subgroup to account for subgroup-specific interpretation and availability of features and hierarchically aggregated these models to cover the entire data. In addition, we deviated from the current clinical workflow only for patients with high probability of benefiting from an alternative treatment, as suggested by this model. Consequently, we devised a two-stage test with optimized negative and positive predictive values as the main indicators of performance. Our analysis was based on 2,377 patients that underwent PCI. Performance was compared with a conventional classification model and the existing clinical practice by estimating effectiveness, safety, and costs for different endpoints (6 month angiographic restenosis, 12 and 36 month hazardous events).

Results: Compared to the current clinical practice, the proposed method achieved an estimated reduction in adverse effects by 25.0% (95% CI, 17.8 to 30.2) for hazardous events at 36 months and 31.2% (95% CI, 25.4 to 39.0) for hazardous events at 12 months. Estimated total savings per patient amounted to $693 and $794 at 12 and 36 months, respectively. The proposed subgroup-specific method outperformed conventional population wide regression: The median area under the receiver operating characteristic curve increased from 0.57 to 0.61 for prediction of angiographic restenosis and from 0.76 to 0.85 for prediction of hazardous events.

Conclusions: The results of this study demonstrated the efficacy of deployment of bare-metal stents and coronary artery bypass grafting surgery for subsets of patients. This is one effort towards development of personalized treatment strategies for patients with coronary atherosclerosis that could significantly impact associated treatment costs.

Keywords: Coronary artery disease, Decision making, Decision support systems, Restenosis, Stents, Coronary artery bypass

* Correspondence: sebastian.poelsterl@tum.de
[1]Computer Aided Medical Procedures, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, Germany
Full list of author information is available at the end of the article

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 2 of 16

Background

Treatment of coronary atherosclerosis due to accumulation of plaques throughout arteries often encompasses mainly two options: minimally invasive percutaneous coronary intervention (PCI) or open-chest coronary artery bypass grafting (CABG). The latter distinguishes itself from PCI by being a more elaborate procedure and by better long-term prognosis [1]. During PCI, the normal blood flow is established within the narrowed region (stenosis) by deploying either stents (small mesh tubes) or balloons that push away the atherosclerotic plaques toward the vessel wall. The two most commonly used types of stents are: bare-metal stents (BMS) and drug-eluting stents (DES).

Treatment of coronary occlusions with BMS implantation has shown to be associated with a high rate (20-30%) of restenosis (re-narrowing of the treated artery) [2-4], which is significantly lower (4-8%) with DES treatment [5,6]. However, at the same time a higher rate of late thrombosis (>1 year after initial treatment) is found to be associated with DES (5 events per 1000 patients compared to no events) [7,8]. Therefore, in contrast to bare-metal stents, drug-eluting stents have a lower risk of early restenosis, but require a prolonged dual antiplatelet therapy, which may lead to thrombosis in case of early discontinuation [9,10].

Overall, the results of multiple studies comparing outcomes by treatment type have shown that 1) the rate of myocardial infarction, death, stroke, and stent thrombosis is similar for BMS and DES, but that BMS has a higher rate of restenosis, and thus revascularization [11-14], and that 2) CABG is superior to DES with respect to major adverse cardiac or cerebrovascular events [1,14]. In addition to individual pros and cons, each treatment option is accompanied by different costs; initial treatment costs for CABG are higher than for PCI and initial DES treatment is more expensive than BMS.

Several studies showed that DES is superior to their BMS counterpart only for a subset of patients [15-18]. Amin et al. [19] quantified the benefits of both options. They showed that reducing the usage of DES by 50% for patients at low risk of target-vessel revascularization (TVR) could save up to $205 million per year in U.S. health care costs, while the overall rate of TVR was projected to increase by 0.5%. In response, Cavender and Ellis [20] pointed out several flaws in the analysis by Amin et al. [19]: 1) they used unbalanced data (the vast majority had been treated with DES), 2) they used TVR as a proxy for restenosis, and 3) that their model discriminated only modestly.

From a treatment cost perspective, multiple studies compared the cost-effectiveness of CABG and PCI for patients with multivessel coronary artery disease [21-25]. Varani et al. [22] concluded that the total treatment

costs during the first year were lower with PCI employing DES, whereas the remaining studies determined that CABG was more cost-effective than PCI.
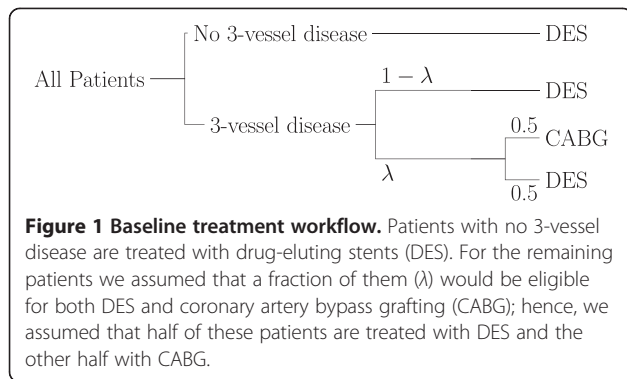
The SYNTAX[a] score II is a tool to aid decision making between CABG and PCI. It augments the anatomical SYNTAX score with seven clinical variables, which are strongly associated with mortality in either the PCI or CABG setting in patients with complex coronary artery disease [26]. The authors concluded that they could accurately predict 4-year mortality of patients. However, their model is only applicable to patients with complex coronary artery disease. Similarly, the authors of [27] developed a decision support tool for CABG and PCI with BMS or DES to predict long-term mortality (5 and 10 years) under each treatment option. Although, their models are based on a general patient population, they did not provide quantitative results with respect to the predictive power of their models or the economic impact.

Our aim was to develop a model that captures the current clinical knowledge about risks and benefits of treatment options for atherosclerosis to recommend a more personalized treatment regime for patients. If risks can be predicted reliably, it would be possible to deviate from the existing clinical practice and replace predominant DES treatment with either BMS, for patients with low risk of future restenosis, or CABG for patients at high risk of death, myocardial infarction, or thrombosis. To achieve this, we formed two hierarchical models: the first is to assess the risks of BMS treatment and the second is to evaluate the risks of DES treatment. The subgroup-specific models allowed us to overcome the lack of available features for all patients. Furthermore, by including in vitro diagnostic biomarkers, we demonstrated a potential improvement on both the patient outcome and the estimated overall treatment costs.

Clinical motivation

Our approach was motivated by two recent shifts in the clinical practice. First, although stenting predominately employs DES with particular focus on reducing restenosis rates, recent studies showed that about one third of patients in fact do not draw much benefit from DES [15]. Moreover, this specific patient subgroup might unnecessarily be exposed to a higher risk of thrombosis leading to major adverse events. Subsequently, for these patients the better treatment choice both clinically and financially could be BMS.

Second, it has been shown that patients, who were traditionally in a grey area for eligibility of both DES and CABG (see Figure 1), could benefit from DES treatment. These include diabetic patients with mild to less severe coronary artery disease affecting multiple vessels [1]. For these patients, who are not particularly at high risk of thrombosis and hence major adverse events with DES

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 3 of 16

Figure 1 Baseline treatment workflow. Patients with no 3-vessel disease are treated with drug-eluting stents (DES). For the remaining patients we assumed that a fraction of them (λ) would be eligible for both DES and coronary artery bypass grafting (CABG); hence, we assumed that half of these patients are treated with DES and the other half with CABG.

therapy, the better choice both clinically and financially could be DES.
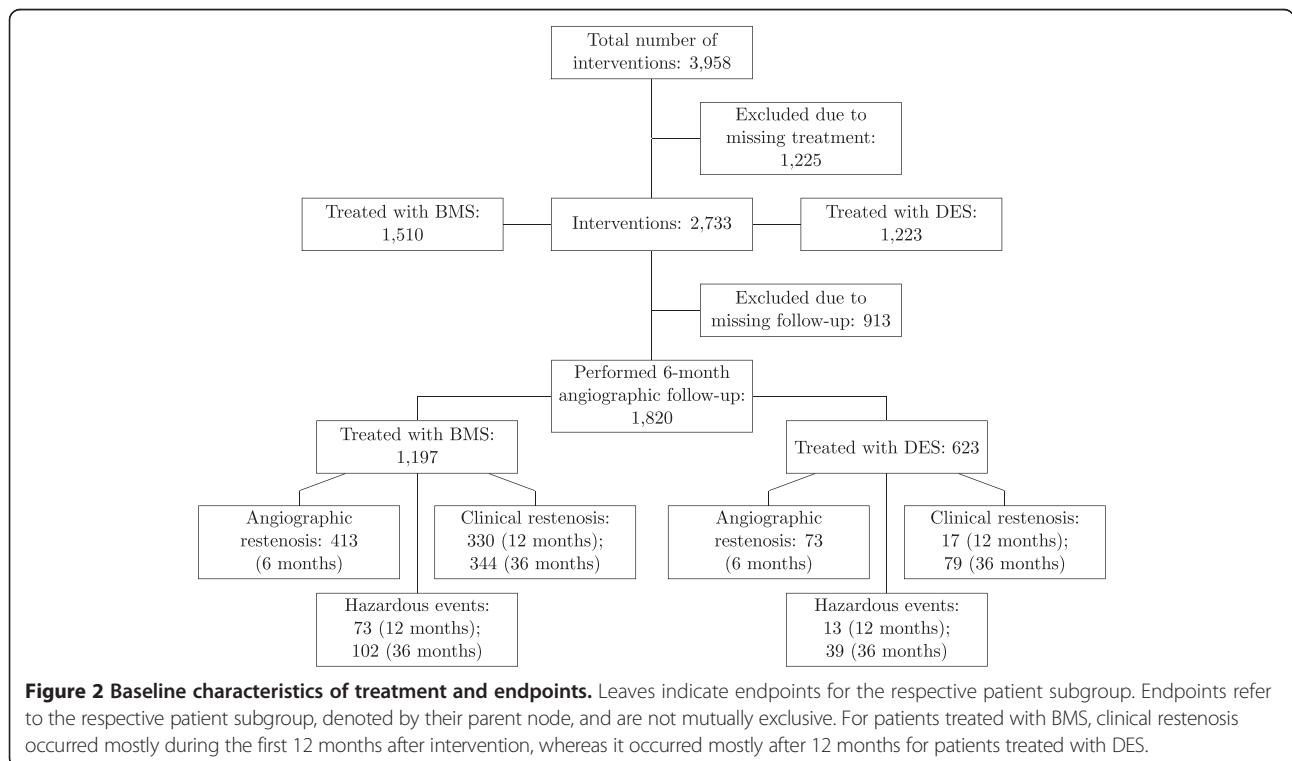
Methods

Study population

This is a retrospective observational study incorporating data from two studies at the Deutsches Herzzentrum (Munich, Germany) [28,29]. Data were originally collected between 1999 and 2006 to evaluate the prognostic value of in vitro diagnostic (IVD) biomarkers. Both studies conformed to principles of the Declaration of Helsinki and were approved by the institutional ethics committee at Deutsches Herzzentrum (Munich, Germany). All patients gave informed consent for angiographic examination and collection of in vitro diagnostic biomarkers. Based on these data sets, we performed secondary analysis

to estimate treatment proportions and associated adverse effects. All subjects underwent coronary angiography due to chest pain or other symptoms suggestive of coronary artery disease (CAD). Our analyses was based on 2,733 interventions with known treatment. This data set comprised 2,377 patients of which 913 were lost to follow-up, i.e., withdrew consent or were unreachable. For the remaining 1,820 interventions, DES were used for 34.2%, and the remainder were treated with BMS (see Figure 2).

Follow-up and endpoints

We investigated two main endpoints, restenosis and hazardous events. For restenosis, we considered angiographic restenosis or clinical restenosis. Angiographic restenosis was defined when the treated vessel developed a blockage with more than 50% diameter compared to the reference diameter proximal to the occlusion. Angiographic restenosis was evaluated at six months after the initial treatment as identified by a quantitative coronary angiography (QCA) exam. Restenosis most likely occurs in the first six months after BMS treatment [2-4]; thus, we selected this time point in our risk assessment of BMS. Clinical restenosis is based on whether the patient required an additional intervention, termed target lesion revascularization, and is typically evaluated one and three years after intervention. We included clinical restenosis to enable comparison with studies that used this specific endpoint as a proxy for angiographic restenosis.

Figure 2 Baseline characteristics of treatment and endpoints. Leaves indicate endpoints for the respective patient subgroup. Endpoints refer to the respective patient subgroup, denoted by their parent node, and are not mutually exclusive. For patients treated with BMS, clinical restenosis occurred mostly during the first 12 months after intervention, whereas it occurred mostly after 12 months for patients treated with DES.

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 4 of 16

Hazardous events were defined as a composite of three major adverse endpoints: death of any cause, myocardial infarction, and thrombosis. Particularly, hazardous events do not comprise target lesion revascularization, unlike the common clinical endpoint Major Adverse Cardiac Events (MACE). For DES we wanted to predict the risk of hazardous event, which has been shown to increase late (>1 year) after intervention [7,8]. We examined hazardous events at one and three years after intervention. All end points are summarized in Figure 2.

Predictive variables

Our predictive models were based on a variety of features. We recorded age, sex and diabetes indication for each record, along with 30 clinical variables (including angiographic measurements) and 11 in vitro diagnostic variables, which comprised lipid panel and cardiac markers. Since data were collected over a period of seven years, not all variables could be recorded for all patients in the entire time span. Therefore, we constructed three sets of features from the complete family of clinical variables and biomarkers, respectively. Clinical variables were divided into three sets of 28, 25 and 22 features, respectively (see Table 1). IVD variables were split into one set of eight and two sets of five measurements (see Table 2). All continuous features were normalized to have zero mean and unit variance.

Notations

Here, we consider families of features that represent different aspects of the patient's health and/or modalities. For instance, a patient's blood pressure, weight and disease history belongs to the family of clinical features, whereas measurements of in vitro diagnostic biomarkers comprise their own family. Each group of features derived from one family of features, is referred to as *feature set*. A classifier that considers two feature sets i and j from two different families is denoted by $T_k^{i,j}$, where $k \in \{R, H\}$ corresponds to prediction of $P(\text{Restenosis}|\text{BMS})$ and $P(\text{Hazard}|\text{DES})$, respectively. For simplicity, we drop the indices i and j if the feature sets have been fixed in previous steps. The predicted probability of restenosis by T_R is denoted as P_R and the predicted probability of hazardous events by T_H as P_H. Thresholds on the predicted probabilities are denoted as θ_R and θ_H, respectively, and allow the calculation of specificity (Spec), sensitivity (Sens), positive predictive value (PPV), and negative predictive value (NPV). When applying a threshold, a classifier's prediction can be positive ($P_k \geq \theta_k$) or negative ($P_k < \theta_k$), which we denote by $T_k^{i,j} = 0$ and $T_k^{i,j} = 1$, respectively. Probability estimates of treatments and complications are denoted as $\hat{P}(\cdot)$ or $\tilde{P}(\cdot)$, depending on whether estimates are based on classifiers'

Table 1 Predictive clinical variables

	C1	C2	C3
AHA/AAC lesion class	+	+	+
Angina	+	+	+
Canadian cardiovascular society (CCS) grading of angina pectoris	+	+	+
Angulation	+	-	+
Body mass index	+	+	+
Hypercholesterolemia	+	+	+
Dissection	+	-	+
Ejection fraction	+	+	+
Presence of eccentric lesion	+	+	+
Family history	+	+	+
History of bypass	+	+	+
History of Intervention	+	+	+
History of myocardial infarction	+	+	+
Hypertension	+	+	+
Lesion length	+	+	+
Left ventricular function	+	+	-
Minimum lumen diameter (pre)	+	+	+
Smoker	+	+	+
Number of stents	+	+	-
New York heart association class	+	+	+
Patient height	+	+	+
Patient weight	+	+	+
Plaques	+	+	-
Reference diameter (pre)	+	+	+
ST elevation	-	+	-
Stenosis type	+	+	-
Thrombus	+	-	-
Thrombolysis in myocardial infraction (TIMI) rating	+	+	+
Tortuosity	+	-	-
Vessels affected	+	+	+

List of clinical variables that were included in the different feature sets (C1, C2, C3). Plus signs indicate that a feature has been included, minus signs that it was not included in the respective feature set.

performances or on amounts retrieved directly from the data set itself or from literature. For instance, the quantity $\tilde{P}(\text{Restenosis}|s\text{DES})$ denotes the probability of restenosis for patients that have been assigned DES treatment by the proposed method. It is estimated by counting the actual number of restenosis events in this patient subgroup, as determined by clinical follow-up. In contrast, $\hat{P}(\text{Hazard}|s\text{DES})$ is the estimated probability of hazardous events for the same patient subgroup and is estimated by a classifier's ability to predict hazardous events accurately. Finally, $\text{Prev}(\text{Restenosis}|\text{BMS})$ denotes the prevalence of restenosis among BMS receivers in the overall population ("Treated with BMS"

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 5 of 16

Table 2 Predictive in vitro diagnostic biomarker variables

	B1	B2	B3
Cholesterol	+	-	-
Creatine kinase (all isoforms)	+	+	-
Creatine kinase MB isoenzyme	+	+	-
Creatinine	-	+	+
C-reactive protein	+	+	+
High density lipoprotein	+	-	-
Low density lipoprotein	+	-	-
N-terminal pro-brain natriuretic peptide	-	-	+
Triglycerides	+	-	-
Troponin T	+	+	+
high-sensitivity troponin T	-	-	+

List of in vitro diagnostic biomarkers that were included in the biomarker feature sets (B1, B2, B3). Plus signs indicate that a feature has been included, minus signs that it was not included in the respective feature set.

node in Figure 2), and Prev(Hazard|DES) represents the prevalence of hazardous events among all DES receivers in the overall population ("Treated with DES" node in Figure 2).

Predictive modeling

We propose a two-stage patient stratification scheme that is based on the current clinical understanding of the risks of BMS, DES and CABG treatment (outlined above). We first estimated the risk of restenosis when treated with BMS, followed by estimating the risk of hazardous events when treated with DES (see Figure 3).

If the rate of hazardous events is similar for BMS and DES [11-14], it is sufficient to estimate the risk of restenosis, when treated with BMS, and to suggest BMS only for low-restenosis-risk patients (see T_R in Figure 3). Furthermore, if DES and CABG treatment mostly differ in the rate of hazardous events [1,14], only the risk of these events needs to be estimated when treated with DES, to suggest one or the other (see T_H in Figure 3). We assumed that the risk of restenosis under BMS and the risk of hazardous events under DES are independent (P_R and P_H in Figure 3). This assumption stemmed from the fact that pathophysiologies of in-stent restenosis and in-stent thrombosis are different [30,31]. In other words, these two are competing risks and considered mutually exclusive.

Therefore, we trained two classifiers: one to predict the probability of restenosis when treated with BMS and one to predict the probability of hazardous events when treated with DES. Each of these classifiers was actually a composite of multiple patient subgroup-specific classifiers in the form of a decision tree by recursively partitioning the data set and training a logistic regression model in the leaf nodes [32]. This addresses two aspects: 1) different effects or interpretations of variables on the

predicted outcome in-between these patient subgroups, and 2) a highly heterogeneous data set, which may only have a subset of features available for all patients.

Recursive partitioning

The driving hypothesis was that by stratifying the entire population into subgroups based on some clinically motivated features, we could build a better overall classifier. First, the clinical interpretation of various biomarkers could drastically differ among age groups and genders [33,34], and second, diabetes is already known to be one of the main indicators in deciding PCI treatment option [35]. Therefore, we considered these variables as the main features to recursively partition the data set into patient subgroups as shown in Figure 4. We first performed a binary split according to one feature and then performed additional splits on the resulting patient subgroups with respect to other features. The result was a decision tree whose structure depends on the order and number of binary splits. The maximum depth of a decision tree was limited to four levels, because we only split by age, gender and diabetes indication (see Table 3). We compensated for the limit in tree depth, by training regularized logistic regression models in the leave nodes. We constructed all possible trees (47 in total), with respect to the features age, sex, and diabetes indication, and selected the optimal one as explained below. This is in contrast to greedy algorithms for constructing decision trees, such as CART [32], which do not guarantee an optimal decision tree for a given set of splits.

Each partitioning of the data set resulted in a decision tree with the respective patient subgroups as leaves (see Figure 4, left). At each leaf node, we consider two patient subgroups with respect to treatment options of BMS and DES, corresponding to T_R and T_H in Figure 3. Here, we split according to three features and therefore obtained 27 different patient subgroups at the leaf level across all decision trees (including trees where one or more splits have not been performed).

Leaf node logistic regression model

For each patient subgroup, i.e., leaf node in Figure 4, we trained multiple models to predict P(Restenosis|BMS) and P(Hazard|DES), respectively. As stated in section "Predictive Variables", training of multiple models was necessary to account for varying feature availability as well as different effects of features on the treatment options. Thus, the models trained in each subgroup differed by the sets of clinical and biomarker features they considered. Furthermore, for some combination of patient subgroup and feature sets, it might not be possible to construct a classifier due to missing data.

Let us consider N different sets of clinical features, M different sets of in vitro diagnostic biomarkers, as well as

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 6 of 16
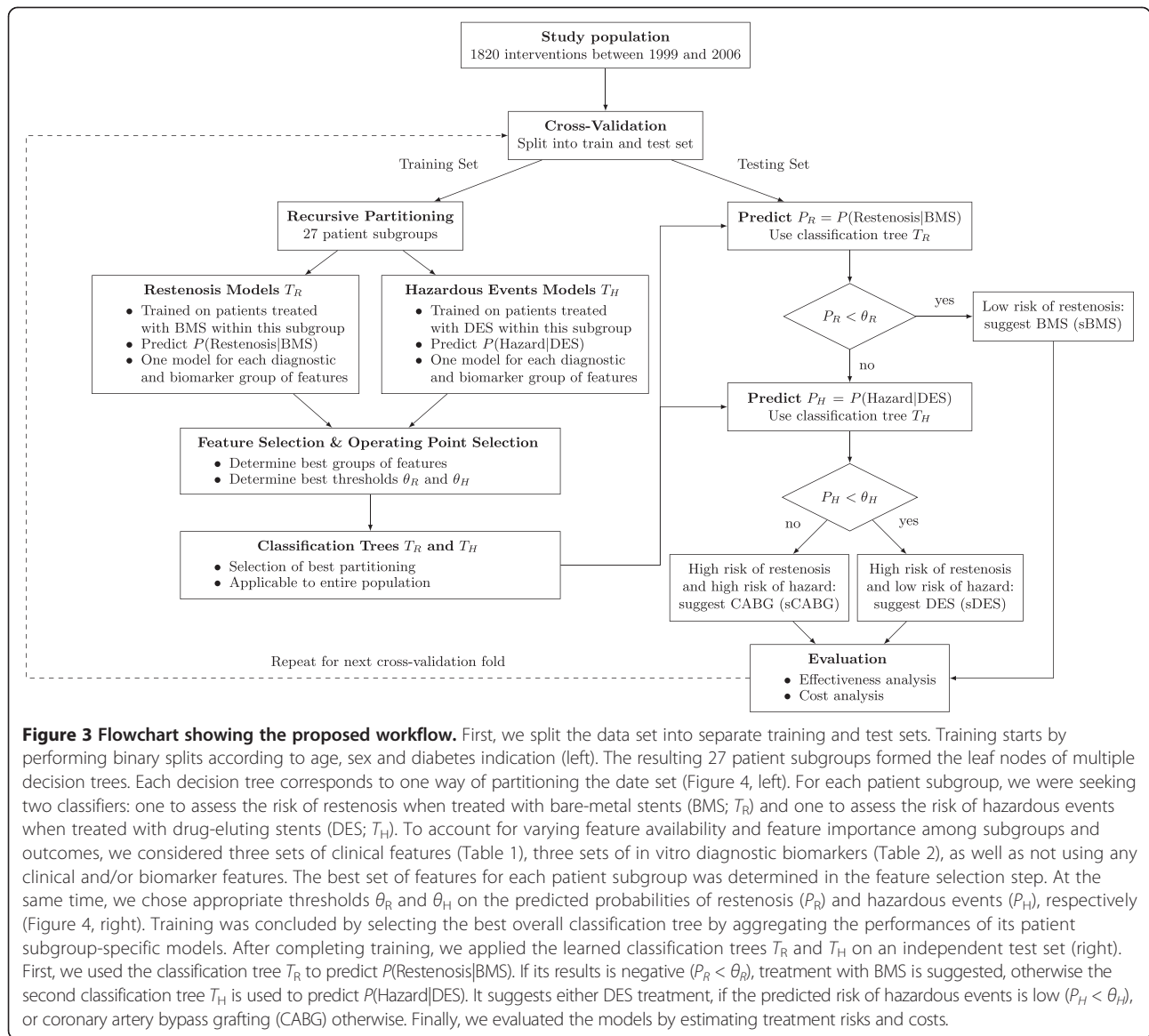
Figure 3 Flowchart showing the proposed workflow. First, we split the data set into separate training and test sets. Training starts by performing binary splits according to age, sex and diabetes indication (left). The resulting 27 patient subgroups formed the leaf nodes of multiple decision trees. Each decision tree corresponds to one way of partitioning the date set (Figure 4, left). For each patient subgroup, we were seeking two classifiers: one to assess the risk of restenosis when treated with bare-metal stents (BMS; T_R) and one to assess the risk of hazardous events when treated with drug-eluting stents (DES; T_H). To account for varying feature availability and feature importance among subgroups and outcomes, we considered three sets of clinical features (Table 1), three sets of in vitro diagnostic biomarkers (Table 2), as well as not using any clinical and/or biomarker features. The best set of features for each patient subgroup was determined in the feature selection step. At the same time, we chose appropriate thresholds θ_R and θ_H on the predicted probabilities of restenosis (P_R) and hazardous events (P_H), respectively (Figure 4, right). Training was concluded by selecting the best overall classification tree by aggregating the performances of its patient subgroup-specific models. After completing training, we applied the learned classification trees T_R and T_H on an independent test set (right). First, we used the classification tree T_R to predict P(Restenosis|BMS). If its results is negative ($P_R < \theta_R$), treatment with BMS is suggested, otherwise the second classification tree T_H is used to predict P(Hazard|DES). It suggests either DES treatment, if the predicted risk of hazardous events is low ($P_H < \theta_H$), or coronary artery bypass grafting (CABG) otherwise. Finally, we evaluated the models by estimating treatment risks and costs.
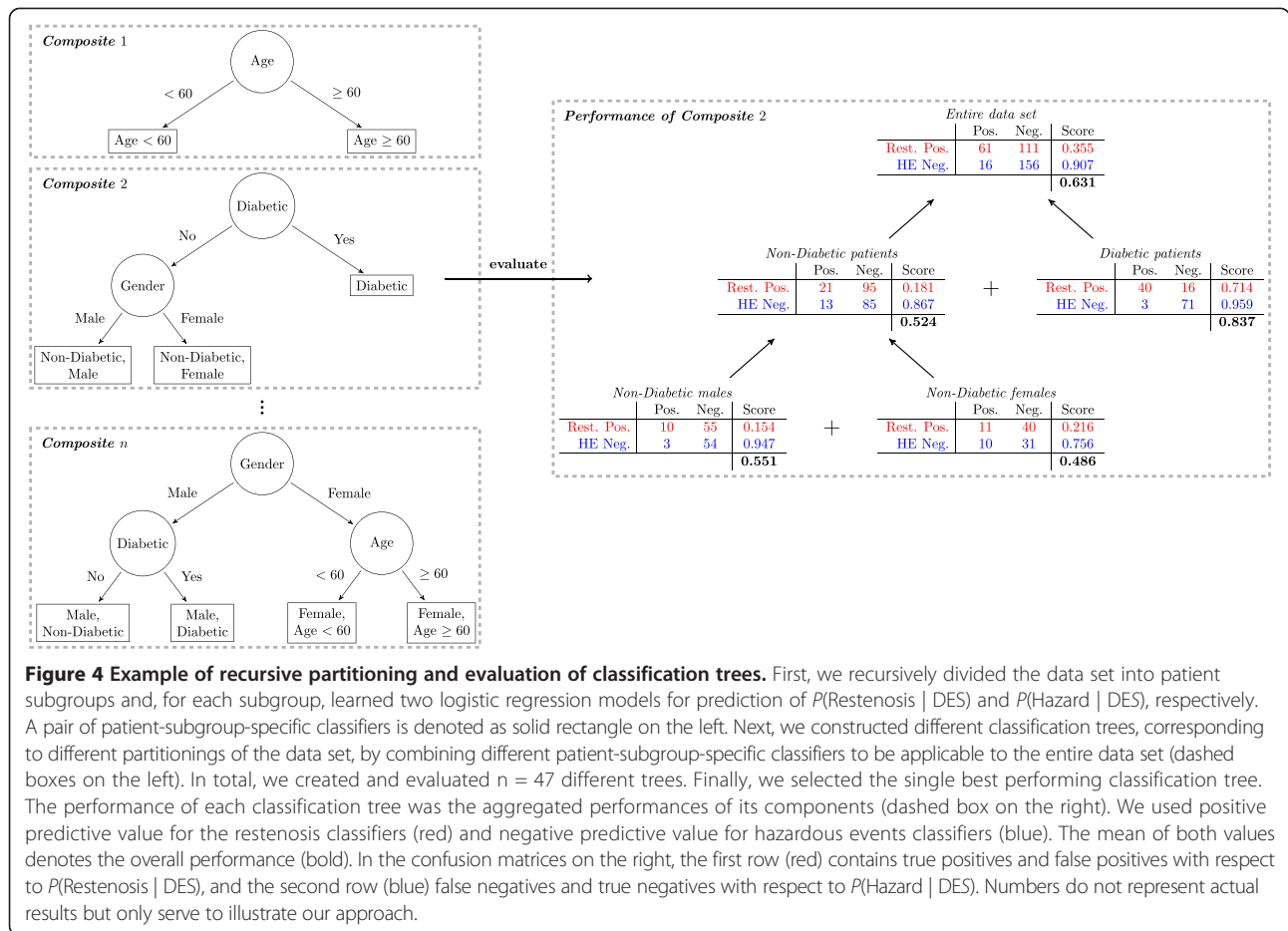
the option of not using any clinical and/or biomarkers information. Therefore, for each of 27 patient subgroups at the leaf level, we need to consider at most *NM* different combinations of clinical and biomarker feature sets – depending on data availability. For instance, with $N = M = 4$, the maximum number of classifiers that have to be trained is 432.

We used logistic regression with ℓ_2 regularization to effectively deal with multicollinearity [36,37]. The classifiers $T_R^{i,j}$ were trained on patients treated with BMS, and classifiers $T_H^{i,j}$ on patients treated with DES, where $i \in \{0, ..., N\}$ and $j \in \{0, ..., M\}$ indicate the set of clinical and biomarkers features used by the classifier or that no features from that group were used (zero value). For building the logistic regression models with ℓ_2 regularization, we used Weka

version 3.6.9, which uses a quasi-Newton algorithm to search for the optimal coefficients [38].

Feature selection and operating point selection

In the previous step, the classifiers for restenosis ($T_R^{i,j}$) and hazardous events ($T_H^{i,j}$) were trained on the training set portion based on maximizing the log-likelihood of the logistic regression model [36,37] for all possible patient subgroups and feature sets independently. The next step was to select the best clinical and biomarker feature sets as well as to find the best thresholds θ_R and θ_H on the predicted probabilities of restenosis and hazardous events, respectively. We obtained classifiers T_R and T_H for each of the 27 patient subgroups (leaves in Figure 4), from which we formed the cascade depicted in Figure 3 (right).

Figure 4 Example of recursive partitioning and evaluation of classification trees. First, we recursively divided the data set into patient subgroups and, for each subgroup, learned two logistic regression models for prediction of *P*(Restenosis | DES) and *P*(Hazard | DES), respectively. A pair of patient-subgroup-specific classifiers is denoted as solid rectangle on the left. Next, we constructed different classification trees, corresponding to different partitionings of the data set, by combining different patient-subgroup-specific classifiers to be applicable to the entire data set (dashed boxes on the left). In total, we created and evaluated n = 47 different trees. Finally, we selected the single best performing classification tree. The performance of each classification tree was the aggregated performances of its components (dashed box on the right). We used positive predictive value for the restenosis classifiers (red) and negative predictive value for hazardous events classifiers (blue). The mean of both values denotes the overall performance (bold). In the confusion matrices on the right, the first row (red) contains true positives and false positives with respect to *P*(Restenosis | DES), and the second row (blue) false negatives and true negatives with respect to *P*(Hazard | DES). Numbers do not represent actual results but only serve to illustrate our approach.

The clinical objective was to have the least disruptions to the current trend, which is primarily toward DES utilization for all patients. T_R captures the risk of restenosis for BMS patients. By maximizing the NPV of T_R, we minimized the number of patients falsely suggested to get BMS as a treatment option. The second paired test T_H captures the risk of hazardous events for patients receiving DES treatment. Maximizing the PPV of T_H, in fact minimizes the number of patients falsely prescribed to be treated with DES, while they are at high risk of adverse events (under such treatment). Consequently, our clinically motivated objective function for selecting the thresholds and features for classifiers T_R and T_H is

$$\max_{i,j,\theta_R,p,q,\theta_H} \text{NPV}\left(T_R^{i,j}, \theta_R\right) + \text{PPV}(T_H^{p,q}, \theta_H) \qquad (1)$$

Feature selection and operating point selection were carried out concurrently by doing an exhaustive search over all possible combinations of feature sets (i, j, p and q) and thresholds (θ_R and θ_H). For the thresholds, all possible values that correspond to a classifier's specificity in the range from 5% to 95% in 5% increments were considered; excluding the trivial classifiers with 0% and 100% specificity. Also, the pair of classifiers $T_R^{i,j}$ and $T_H^{p,q}$ always considered the same patient subgroup (i.e., age group, gender and diabetes indication).

Referring to the above example with clinical and biomarkers composed of four feature sets each, at least all combinations of thresholds of both classifiers need to be evaluated per patient subgroup ($2 \cdot 19 = 38$ comparisons). At the other end of the spectrum, if all feature sets are available, all possible combinations of feature sets must additionally be compared ($2 \cdot 19 \cdot 16 = 608$ combinations). During training, the exact number of comparisons may vary from subgroup to subgroup and depends on the data availability with respect to particular patient

Table 3 Different combinations of features used for creating patient subgroups

	Patient subgroups		
Age	All	≤60	>60
Sex	All	Female	Male
Diabetes	All	No diabetes	With diabetes

In total, the complete data set was partitioned into 27 different patient subgroups.

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 8 of 16

subgroups. Note that feature selection, operating point selection and pairing were only performed on the training set of each cross-validation iteration and were fixed during evaluation and predictions.

Classification tree selection

Returning to the partitioning of the data set into patient subgroups described above, the last step of the training procedure consisted of aggregating the patient-subgroup-specific classifiers T_R and T_H to form a classification tree that is applicable to the entire patient population and choosing the best performing classification tree (Figure 4, right). Hence, the result is a classification tree that is composed of a set of regularized logistic regression models applicable to different population subsets, respectively.

To determine the best partitioning into patient subgroups, we devised a greedy algorithm that maximizes the objective criterion of equation (1) and selects the best performing decomposition of the training data set into subgroup-specific classifiers (Figure 4, right). First, the predictions of all subgroup-specific classifiers on the training set were collected; next, the resulting confusion matrix was used to calculate the performance of the classification tree according to the objective function in equation (1).

Since multiple classification trees, which correspond to alternative ways of partitioning the data set, are considered, each of them is competing for the best predictive performance. For instance, a decision tree with leaves "Male", "Female with Age ≥ 60" and "Female with Age < 60" is competing with a decision tree that only has the split according to sex (or any other combination of splits). From all competing decompositions into different patient subgroups, we picked the decomposition that maximized the objective in equation (1), which was extracted from the root nodes of the classification trees. As a result, decision trees for T_R and T_H and their respective performances were obtained.

Evaluation

Effectiveness analysis

Based on the aggregated performance results of the classification tree for restenosis and hazardous events prediction, the probability of adverse effects of the two test setup in Figure 3 was estimated. We considered the composite of restenosis and hazardous events as adverse effects, thus $P(\text{Adverse Effect})$ was defined as

$$P(\text{Adverse Effect}) = P(\text{Restenosis}) + P(\text{Hazard}) \quad (2)$$

Let sBMS, sDES, sCABG denote that BMS, DES, and CABG was suggested by the two test workflow as shown on the right of Figure 3, respectively. The

probabilities on the right hand side of eq. (2) can be decomposed according to the three possible treatments $\mathcal{T} = \{\text{sBMS}, \text{sDES}, \text{sCABG}\}$ considered here

$$P(\text{Restenosis}) = \sum_{t \in T} P(t)P(\text{Restenosis}|t)$$

$$P(\text{Hazard}) = \sum_{t \in T} P(t)P(\text{Hazard}|t)$$

In both equations, we estimated the probabilities of individual treatments by taking the predictive performance of classification trees T_R and T_H into account. The first test predicts the probability P_R of restenosis when treated with BMS, and is negative if P_R is lower than threshold θ_R, in which case BMS would be suggested. Thus, the estimate $\hat{P}(\text{sBMS})$ can be calculated by

$$\begin{aligned} \hat{P}(\text{sBMS}) &= P(P_R < \theta_R) = P(T_R = 0) \\ &= \text{Prev}(\text{Restenosis}|\text{BMS})\big(1 - \text{Sens}(T_R)\big) \\ &\quad + \big(1 - \text{Prev}(\text{Restenosis}|\text{BMS})\big)\text{Spec}(T_R). \end{aligned}$$

The second test T_H predicts the probability of hazardous events, if treated with DES. Assuming independence between the first and second test, we obtain $\hat{P}(\text{sDES})$ as

$$\begin{aligned} \hat{P}(\text{sDES}) &= P(P_R \geq \theta_R, P_H < \theta_H) = P(T_R = 1)P(T_H = 0) \\ &= \big(1 - \hat{P}(\text{sBMS})\big)\text{Prev}(\text{Hazard}|\text{DES})\big(1 - \text{Sens}(T_H)\big) \\ &\quad + \big(1 - \hat{P}(\text{sBMS})\big)\big(1 - \text{Prev}(\text{Hazard}|\text{DES})\big)\text{Spec}(T_H). \end{aligned}$$

Finally, the probability of the remaining treatment CABG was simply calculated based on the previous two probabilities.

$$\begin{aligned} \hat{P}(\text{sCABG}) &= P(P_R \geq \theta_R, P_H \geq \theta_H) \\ &= 1 - P(T_R = 0) - P(T_R = 1)P(T_H = 0) \\ &= 1 - \hat{P}(\text{sBMS}) - \hat{P}(\text{sDES}) \end{aligned}$$

Next, we estimated the conditional probabilities $P(\text{Restenosis}|\text{sBMS})$ and $P(\text{Hazard}|\text{sDES})$. These probabilities depend on the performances of T_R and T_H, respectively, and were estimated by

$$\begin{aligned} \hat{P}(\text{Restenosis}|\text{sBMS}) &= P(\text{Restenosis}|T_R = 0) \\ &= 1 - \text{NPV}(T_R) \\ \hat{P}(\text{Hazard}|\text{sDES}) &= P(\text{Hazard}|T_R = 1, T_H = 0) \\ &= 1 - \text{NPV}(T_H), \end{aligned}$$

where we assumed independence between the two tests in the second equation.

$P(\text{Hazard}|\text{sBMS})$ and $P(\text{Restenosis}|\text{sDES})$ depend on stratification by the first and second test, respectively. We estimated these probabilities by retrieving the corresponding label from samples classified as negative by T_R, which yields $\tilde{P}(\text{Hazard}|\text{sBMS})$, and by T_H, to obtain $\tilde{P}(\text{Restenosis}|\text{sDES})$.

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 9 of 16

Because our data set did not contain any information about CABG treatments, we collected estimates for $P(\text{Hazard}|\text{sCABG})$ and $P(\text{Restenosis}|\text{sCABG})$ from literature [39-50], and used 8.56% and 1.9%, respectively (see Additional file 1 for details).

Cost analysis

To evaluate the economic impact of our method, we performed a cost analysis of the proposed two test setup. The average treatment costs of the proposed setup depends on the predicted outcome of both tests and was defined as

$$\hat{P}(\text{sBMS})(C_{\text{sBMS}} + C_{\text{DAPT}}) + \hat{P}(\text{sDES}) \qquad (3)$$
$$(C_{\text{sDES}} + C_{\text{DAPT}}) + \hat{P}(\text{sCABG})C_{\text{sCABG}},$$

where C_t denotes the costs for suggested treatment $t \in \mathcal{T}$ and C_{DAPT} the costs for one year dual anti-platelet therapy. Finally, we considered average costs per patient for any corrective procedure in the case of restenosis or hazardous events.

$$\hat{P}(\text{Hazard})\left(\frac{1}{2}(C_{\text{MI}} + C_{\text{Stroke}}) + C_{\text{Corrective}}\right)$$
$$+ \hat{P}(\text{Restenosis})C_{\text{Corrective}}, \qquad (4)$$

where C_{MI} and C_{Stroke} are the treatment costs for myocardial infarction and stroke, respectively, and $C_{\text{Corrective}}$ the costs for any corrective procedure.

Baseline effectiveness and costs

We compared our proposed method with the baseline workflow, where most patients are treated with DES, by estimating adverse effects and costs similar to outlined above. In the baseline workflow, all patients are treated with DES except patients with 3-vessel disease, for whom the SYNTAX trial concluded that CABG would be an appropriate alternative to PCI [1]. We analyzed a PCI population, hence we did not consider CABG exclusively; instead we considered CABG and DES by equal amounts.

The overall relative amount of patients suitable for both CABG and DES treatments was calculated as $\omega = \lambda P(3 \text{ - vessel disease})$, where $\lambda \in [0, 1]$ denotes the percentage of patients with 3-vessel disease that fall into this group (see Figure 1). The estimate $\tilde{P}(3\text{-vessel disease})$ was retrieved from data, based on selected subgroups. Accordingly, the baseline initial treatment costs were calculated as follows:

$$(1-\omega)(C_{\text{sDES}} + C_{\text{DAPT}})$$
$$+ \frac{1}{2}\omega(C_{\text{CABG}} + C_{\text{sDES}} + C_{\text{DAPT}}).$$

For corrective procedure costs, the probability of restenosis was estimated as the weighted sum of $P(\text{Restenosis}|\text{DES})$ and $P(\text{Restenosis}|\text{CABG})$ as

$$(1-\omega)\tilde{P}(\text{Restenosis}|\text{DES}) + \frac{1}{2}\omega\Big(\tilde{P}(\text{Restenosis}|\text{DES})$$
$$+ \tilde{P}(\text{Restenosis}|\text{CABG})\Big).$$

A similar formula can be obtained for the costs of hazardous events.

Finally, we estimated treatment portions and associated adverse effects by retrieving estimates for $P(\text{Adverse Effect})$ and its components directly from data, i.e., it did not depend on the performance of a classifier. We used $\lambda = 0.21$, which is the percentage of low complexity lesions (SYNTAX scores ≤ 22) determined by the 3-year results of the SYNTAX trial [51]. For these patients, no significant difference between CABG and PCI treatment was observed.

Results

Cross-validation scheme

For our analyses, we used different outcomes for the restenosis label and the hazardous events label, respectively: angiographic restenosis at 6 months after intervention, early (1-year) clinical restenosis, and late (3-year) clinical restenosis for restenosis prediction, as well as early (1-year) and late (3-year) hazardous events prediction. We used these outcomes in four different analyses: 1) we analyzed angiographic restenosis and early hazardous events, 2) we considered late hazardous events instead of the 1-year outcome, 3) we considered target lesion revascularization (clinical restenosis) as a proxy for restenosis combined with early hazardous events, and 4) we combined clinical restenosis with late hazardous events.

We required that each patient subgroup using the aforementioned sets of features to have no missing values and at least 100 samples, as well as ten samples where the class variable is positive (restenosis or hazardous events). Subgroups that did not satisfy these constraints were disregarded. Consequently, we constructed between 133 and 185 classifiers, depending on the choice of the endpoint (see Additional file 1: Table S3).

To get an estimate for the expected prediction error of our proposed method, 10-fold cross-validation was applied on our data set. Furthermore, this process was repeated three hundred times with different randomly selected cross-validation splits [52]. Referring to the number of classifiers trained on each training set from above, we trained in total between 133·10·300 and 185·10·300 different classifiers during evaluation of individual endpoints. In each iteration, performance evaluation was solely carried out on the test set, which was not used during training. In addition, this allowed us to assess the variance of the expected prediction error by means of confidence intervals, which have been derived from the percentiles of the empirical distribution of the

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 10 of 16

prediction error. Instead of selecting the ℓ_2 penalty of individual logistic regression models based on nested cross-validation, which would have been prohibitively expensive, we fixed the ℓ_2 penalty of all models at the beginning to 0.01.

Classification trees
Figure 5 compares the performance of subgroup-specific classifiers (gray bars) with conventional population wide regression (white bars). It demonstrates that prediction of hazardous events was highly discriminative with a median area under the receiver operating characteristic curve (AUC) of approximately 0.85 for the 1-year outcome and 0.75 for the 3-years outcome. Models used to predict restenosis performed modestly with a median AUC of 0.61 and 0.60. No significant differences in performance with respect to angiographic and clinical restenosis were observed. We compared the performance of complementary subgroup-specific classifiers with a single global classifier that was constructed using feature selection and operating point selection as described above, but without partitioning the data set. When predicting hazardous events, the global classifier achieved a median

AUC of 0.76 and 0.71 for the 1-year and 3-years outcome. When targeting angiographic restenosis at 6 months, it resulted in an AUC of 0.57.

Effectiveness
Table 4 summarizes the probability of adverse effects with respect to angiographic restenosis and 1-year and 3-year hazardous events, respectively. Using the two-test setup, DES was suggested for the majority of patients (81.8% at 1 year; 85.3% at 3 years). The overall probability of restenosis was 6.4% at 1 year (6.2% at 3 years) and that of hazardous events was 3.3% (7.9%). This yielded a total rate of adverse effects of 9.7% for 1-year hazardous events and 14.1% for 3-year hazardous events. Compared to the baseline treatment, where the rate of adverse effects was 14.1% and 18.8%, respectively, this resulted in a relative reduction of 31.2% (95% CI, 25.4 to 39.0) and 25.0% (95% CI, 17.8 to 30.2).

Similar results were obtained when analyzing clinical restenosis (see Table 5). The absolute decrease in adverse effects was highly significant with 3.2% and 4.9% (both $P < 0.001$) for the 1-year and 3-year endpoints, respectively, corresponding to 33.2% (95% CI,
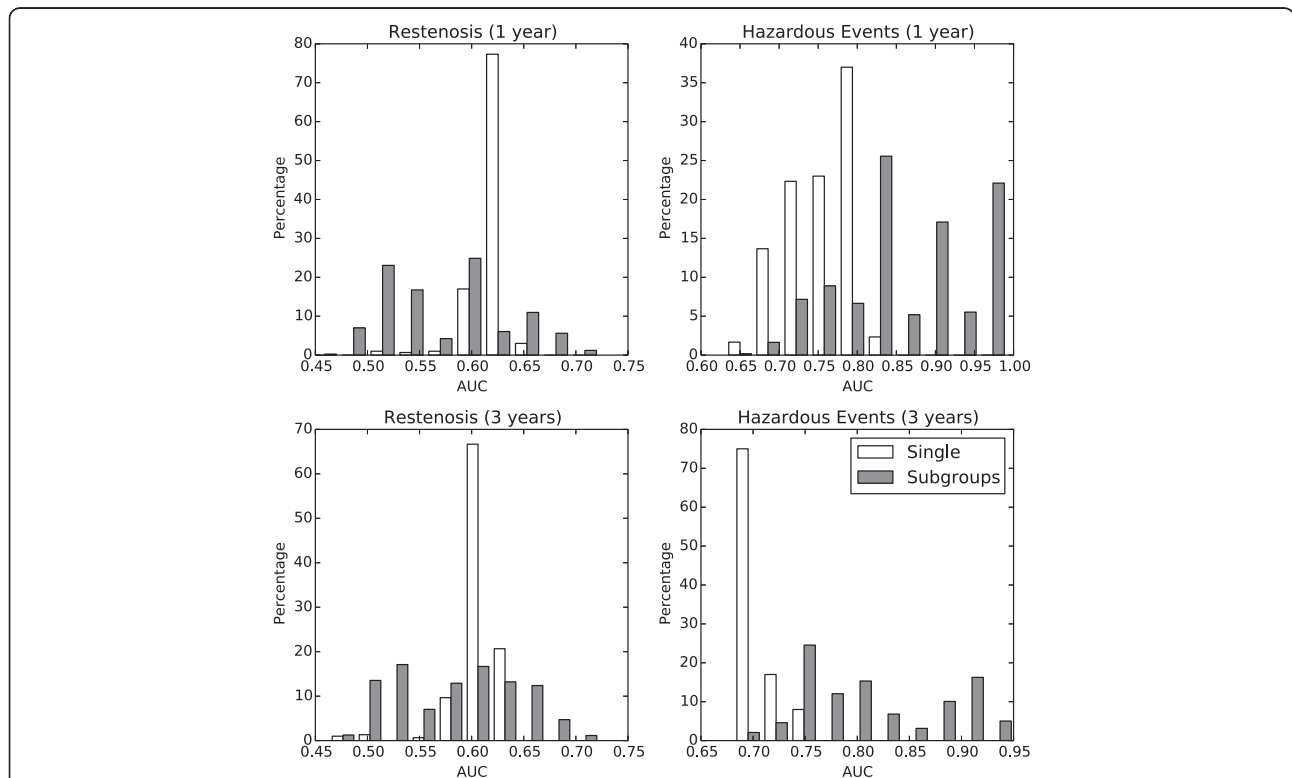
Figure 5 Histograms of area under receiver operating characteristics curve for subgroup-specific and global classifiers. Histograms show the area under receiver operating characteristics curve (AUC) for angiographic restenosis and hazardous events classifiers after feature selection and operating point selection. White bars indicate the performance of ℓ_2 regularized logistic regression models trained on the whole population using only feature selection and operating point selection, but no partitioning. Gray bars indicate the performance of subgroup-specific classifiers.

Table 4 Estimates of key quantities characterizing the performance of the proposed two test workflow when 1-year or 3-year occurrences of hazardous events (HE) are taken into account along with 6-month angiographic restenosis results

Estimated quantity	1-year HE		3-years HE	
	Mean	95% CI	Mean	95% CI
$\hat{P}$(sBMS)	9.7	3.6 – 18.0	5.0	2.6 – 9.5
$\hat{P}$(sDES)	81.8	74.3 – 87.5	85.3	81.4 – 87.7
$\hat{P}$(sCABG)	8.5	7.5 – 9.0	9.6	7.8 – 10.6
$\hat{P}$(Restenosis|sBMS)	0.1	0.0 – 0.0	0.0	0.0 – 0.0
$\tilde{P}$(Restenosis|sDES)	7.6	6.7 – 7.9	7.0	6.1 – 7.5
$\hat{P}$(Restenosis)	6.4	5.5 – 6.9	6.2	5.3 – 6.6
$\tilde{P}$(Hazard|sBMS)	0.0	0.0 – 0.0	1.5	0.0 – 16.7
$\hat{P}$(Hazard|sDES)	3.2	2.1 – 3.5	8.2	7.2 – 9.5
$\hat{P}$(Hazard)	3.3	2.5 – 3.8	7.9	6.9 – 9.5
Baseline estimate for P(Adverse effect)	14.1	12.6 – 14.5	18.8	17.6 – 19.6
Proposed estimate for P(Adverse effect)	9.7	8.3 – 10.6	14.1	12.4 – 15.6
ΔP(Adverse effect)	4.4	3.5 – 5.4	4.7	3.4 – 5.5

The acronyms sBMS, sDES and sCABG denote that the proposed model suggested treatment with bare-metal stents, drug-eluting stents or coronary artery bypass grafting, respectively. Baseline refers to predominant treatment with drug-eluting stents as described in section "Baseline Effectiveness and Costs".

26.6 to 41.9) and 25.0% (95% CI, 20.4 to 31.4) of relative reduction.

Furthermore, we performed the same analyses without partitioning and used a single global classifier instead. With respect to angiographic restenosis, the resulting probability of adverse effects was 10.6% (95% CI, 9.1 to 12.0) and 15.8% (95% CI, 14.3 to 17.7), which is significantly higher ($P < 0.001$) compared to results above (see Additional file 1: Table S7).

Costs

Based on the estimates $\hat{P}$(sBMS) , $\hat{P}$(sDES) and $\hat{P}$(sCABG) and Medicare reimbursements in U.S. dollars of the fiscal year 2013, we calculated average costs of initial treatments and corrective procedures per patient. Thus, costs of procedures are based on the average reimbursement rates across all hospitals in the U.S., which are covered by the Medicare health insurance program (see Additional file 1 for details).

Table 5 Estimates of key quantities characterizing the performance of the proposed two test workflow when 1-year or 3-year occurrences of hazardous events (HE) and clinical restenosis are taken into account

Estimated quantity	1-year		3-years	
	Mean	95% CI	Mean	95% CI
$\hat{P}$(sBMS)	5.6	3.4 – 9.8	6.4	3.1 – 13.0
$\hat{P}$(sDES)	85.5	81.9 – 87.6	84.1	78.4 – 87.2
$\hat{P}$(sCABG)	8.9	8.2 – 9.2	9.5	8.5 – 10.4
$\hat{P}$(Restenosis|sBMS)	0.0	0.0 – 0.0	0.0	0.0 – 0.0
$\tilde{P}$(Restenosis|sDES)	2.9	1.9 – 3.4	8.2	7.6 – 9.0
$\hat{P}$(Restenosis)	2.7	1.8 – 3.1	7.1	6.4 – 7.9
$\tilde{P}$(Hazard|sBMS)	7.2	0.0 – 16.1	9.9	0.0 – 21.4
$\hat{P}$(Hazard|sDES)	3.2	2.0 – 3.5	7.4	5.6 – 8.9
$\hat{P}$(Hazard)	3.9	2.7 – 4.6	7.6	6.2 – 8.9
Baseline estimate for P(Adverse effect)	9.7	8.4 – 10.3	19.7	18.2 – 20.7
Proposed estimate for P(Adverse effect)	6.5	5.0 – 7.6	14.7	13.1 – 16.1
ΔP(Adverse effect)	3.2	2.7 – 3.8	4.9	4.1 – 6.2

The acronyms sBMS, sDES and sCABG denote that the proposed model suggested treatment with bare-metal stents, drug-eluting stents or coronary artery bypass grafting, respectively. Baseline refers to predominant treatment with drug-eluting stents as described in section "Baseline Effectiveness and Costs".

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 12 of 16

From Table 6 it is evident that the proposed setup resulted in slightly increased estimated expenses for the initial treatment ($285 to $499), but savings for any corrective procedures due to lower probability of adverse effects ($834 to $1,229). This resulted in estimated overall savings of 4.7% (95% CI, 3.1 to 7.0) and 2.7% (95% CI, 1.8 to 4.2) per patient at 1 year, when considering angiographic and clinical restenosis, respectively. Estimated total savings were equivalent to $693 and $441 per patient. Increasing the time frame to three years resulted in a modest increase of initial treatment costs. The overall costs decreased by 3.8% (95% CI, 2.7 to 5.3) and 4.0% (95% CI, 2.9 to 6.4) per patient, compared to the baseline treatment workflow. With respect to total savings, this amounted to $693 and $739, respectively.

Finally, we investigated the relationship between BMS and DES procedure costs on the estimated total savings by increasing and decreasing the costs by 25% in steps of 1.25%. Therefore, BMS costs ranged between $8,900 and $14,833, and DES costs between $9,342 and $15,571. Figure 6 shows that procedure costs of BMS and DES are dependent on each other. Thus, changing the costs of one of them suggested changing the other one as well in order to sustain the same amount of savings per patient. No setting resulted in a total loss, indicating robustness to actual stent costs. Furthermore, we observed that for the 3 year time frame the cost savings decreased slightly and that the relationship between BMS and DES treatment costs was weaker (illustrated by higher slope of contour lines in Figure 6).

Discussion

We developed a two-stage stratification procedure for patients with coronary atherosclerosis that is based on: 1) the current clinical knowledge about three common treatment options (BMS, DES, and CABG), 2) differences in feature availability and importance among patient subgroups, and 3) the objective to only deviate

Table 6 Estimated mean costs per patient in U.S. dollar for initial treatment and corrective procedures according to equations (3) and (4), respectively

Restenosis	Angiographic restenosis		Clinical restenosis	
	1-year	3-years	1-year	3-years
Hazardous events				
Initial Costs (Baseline)	13,821	13,822	13,801	13,799
Initial Costs (Proposed)	14,106	14,321	14,193	14,289
Corrective procedure (Baseline)	3,022	4,325	2,297	4,439
Corrective procedure (Proposed)	1,943	3,131	1,463	3,210
Total savings	794	693	441	739
Total savings (%)	4.71	3.82	2.74	4.05

Baseline refers to predominant treatment with drug-eluting stents as described in section "Baseline Effectiveness and Costs".

from the currently preferred treatment with DES for patients that would benefit – in the case of CABG – or be at no increased risk – in the case of BMS – with an alternative treatment.

We found that the proposed two-test setup resulted in an increased estimated effectiveness as well as lower estimated costs when compared to the baseline workflow in all settings. By optimizing negative and positive predictive values of patient subgroup-specific classifiers, we accounted for current clinical evidence that DES are dominating BMS only for a subset of patients [15] and that DES are a viable alternative to CABG for certain patient subgroups [1].

In the analyses of clinical restenosis, the rate of restenosis differed considerably between the 1-year and 3-year analysis. This increase is due to the fact that the time point for restenosis varied (1 or 3 years) whereas it was fixed previously (6 months). Therefore, an increase in the restenosis rate was expected due to the longer time period considered.

The current clinical practice for coronary revascularization focuses to a large extent only on PCI with DES and CABG as its alternative, but ignores treatment with BMS. However, patients at low risk of restenosis could still benefit from BMS [15]. We identified these patients and were able to minimize the estimated rate of restenosis for these patients. For a preponderating portion of validation runs the estimate $\hat{P}(\text{Restenosis}|\text{sBMS})$ was zero. We attributed this result to the fact that both the estimate $\hat{P}(\text{Restenosis}|\text{sBMS})$ as well as the objective in eq. (1) incorporate $\text{NPV}(T_R, \theta_R)$ and thus classifiers with a negative predictive value of 1 were selected preferably, which resulted in a restenosis rate of zero.
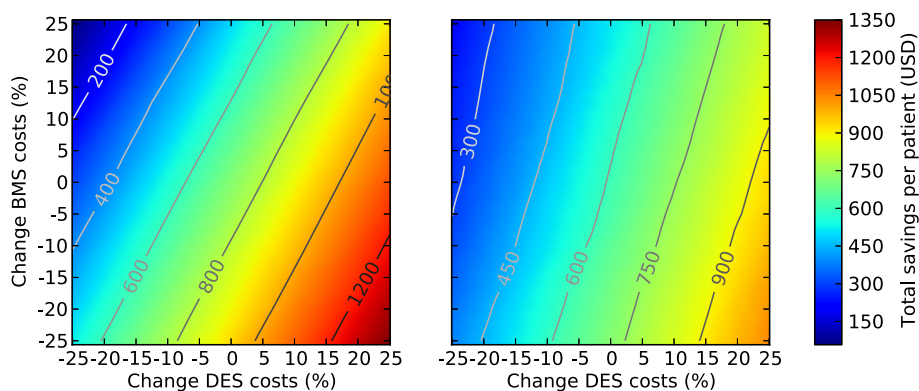
Furthermore, the SYNTAX trial [1] demonstrated that there exists a grey area where the outcome of DES and CABG treatment does not differ significantly. Hence, for some patients, PCI with DES would be as effective as CABG, but significantly less perilous. We incorporated this knowledge in our method and observed that the estimated proportion of hazardous events for patients treated with DES was low (3.2 to 8.2%).

The proposed two stage setup deviated from the baseline workflow only for a small set of patients. This is where there is no increase in risks, as in the case of BMS versus DES, or where it is very likely that they would benefit from an alternative treatment, as in the case of CABG versus DES. When evaluating treatment risks, only those factors that are most decisive – based on results of clinical studies [1,11-14] – were considered, instead of trying to model all possible risks and their interrelationships. For instance, when assessing the risks of BMS treatment we do not suggest ignoring hazardous events in this setting, but based on previous results

Figure 6 Comparison of costs of bare-metal and drug-eluting stents with respect to total saving. Solid lines indicate contour of total savings. Estimation is based on angiographic restenosis together with hazardous effects at 1-year (left) and at 3-years (right).

[11-13], it is known that the rate of hazardous events differs only slightly among BMS and DES treatments. Therefore, predicting hazardous events in this setting would add little additional information.

With respect to economic implications, we noticed that average costs per patient for index procedures slightly increased, whereas average costs per patient for corrective procedures decreased. The former can be explained by the small cost difference between BMS and DES treatment ($590), together with high expenses associated with CABG ($28,683). Therefore, performing stenting twice, once for the index procedure and once for the corrective procedure, instead of one surgery would still be less expensive at the population level than treating the patient by CABG first. This severely diminished the amount of total savings, especially in early stages.

Regarding angiographic restenosis, we observed that the total amount of savings decreased from the 1-year to the 3-year hazardous events analysis and that the percentage of suggested BMS treatment dropped from 8.8% to 4.4%, favoring DES and CABG. Treating a large number of patients with procedures that were more expensive than BMS increased initial treatment costs and ultimately lowered the total savings.

It is important to note that the method proposed here and the SYNTAX score [53] – an algorithm to grade the complexity of CAD based on angiographic data – address different aspects related to CAD. Sianos et al. [53] state the SYNTAX score "is focusing on anatomy of coronary vasculature and not on treatment plan" (p. 226). The SYNTAX score II aimed to overcome this limitation by considering clinical variables in addition to the anatomical SYNTAX score. Although this model performs well, it was derived from a data set composed of patients with left main coronary disease or three-vessel disease. Thus, the SYNTAX score II can facilitate decision making only for a subset of patients that satisfy these criteria.

In addition, only lesions with more than 50% reduction in luminal diameter should be included in the SYNTAX score. In contrast, our approach does not have any exclusion criteria and is based on all available information for one patient – including those mentioned above – to assess the risk of particular treatment options. In fact, the SYNTAX score could complement our approach as an additional source of information in helping to estimate risks of therapeutic options.

These results showed that considering BMS and CABG as alternatives to DES is one effort to a more personalized treatment regime that could also positively impact treatment costs. Further research, in particular in the form of a prospective study, would provide additional insight into the magnitude of these effects.

Limitations

One of the major limitations, subject to all retrospective observational analyses, is that treatment was determined by physicians and was not random. This inherent and unavoidable bias might lead to identifying false estimators. Because data were collected over a period of eight years, changes in technology during this time span, such as changes in data collection protocol, might have influenced obtained measurements. Therefore, temporal latent effects could possibly affect our models.

Due to the requirement of angiographic follow-up at 6 months after intervention and the collection of biomarkers, the patient population used was relatively small. As a consequence, some subgroups could not be analyzed. In our analyses, we treated multiple consecutive interventions per patient as independent and did not take the outcome of previous procedures into account. Risk prediction was formulated as a binary classification problem instead of predicting time-to-event. Furthermore, our data did not include medications and their influence on biomarker measurements could not

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 14 of 16

be quantified. Finally, our estimates with respect CABG are solely based on data we retrieved from literature and models used for restenosis prediction only showed modest performance, which was observed by Amin et al. [19], too.

Conclusions

We demonstrated that by considering multiple treatment options for coronary atherosclerosis and modeling their respective risks, the estimated overall adverse effects and costs could be reduced. This could lead to more efficacious and personalized treatment for patients, and significant economic impact at the population level. When projecting our results on the estimated 1,085,357 coronary revascularization procedures performed on adults in the U.S. in 2007 [54], this would positively affect approximately 47,756 and 51,012 patients for angiographic restenosis at one and three years, respectively, which accumulates to estimated cost savings ranging from $475.4 to $842.2 million dollars, depending on the respective outcome (angiographic or clinical restenosis) and time frame (one or three years).

Endnotes

ᵃDerived from the study of the same name (SYNergy between Percutaneous Coronary Intervention with TAXus and Cardiac Surgery).

Additional file

Additional file 1: Supplementary material. PDF document containing supplementary materials.

Abbreviations

AHA/ACC: American Heart Association/American College of Cardiology; AUC: Area under the receiver operating characteristic curve; B1-3: Feature set 1, 2, or 3 comprising in vitro diagnostic biomarkers (see Table 2); BMS: Bare-metal stent treatment; C1-3: Feature set 1, 2 or 3 comprising diagnostic variables (see Table 1); CABG: Coronary artery bypass grafting surgery; CAD: Coronary artery disease; CCS: Canadian Cardiovascular Society; DAPT: Dual anti-platelet therapy; DES: Drug-eluting stent treatment; HE: Hazardous events; IVD: In vitro diagnostic; MACE: Major adverse cardiac events; MI: Myocardial infarction; NPV: Negative predictive value; PCI: Percutaneous coronary intervention; $\tilde{P}(\cdot)$: Probability estimate of treatment or event given in brackets derived from classifiers' performances; $\hat{P}(\cdot)$: Probability estimate of treatment or event given in brackets derived from the data set itself; P_H: Predicted probability $P(\text{Hazard} \mid \text{DES})$ by classifier trained on patients treated with drug-eluting stents; P_R: Predicted probability $P(\text{Restenosis} \mid \text{DES})$ by classifier trained on patients treated with bare-metal stents; Prev(·): Prevalence of event given in brackets; PPV: Positive predictive value; QCA: Quantitative coronary angiography ROC, Receiver operating characteristic; sBMS: Bare-metal stent treatment suggested by the proposed workflow; sCABG: Coronary artery bypass grafting surgery suggested by the proposed workflow; sDES: Drug-eluting stent treatment suggested by the proposed workflow; Sens: Sensitivity; Spec: Specificity; SYNTAX: SYNergy between Percutaneous Coronary Intervention with TAXus and Cardiac Surgery; θ_H: Threshold on predicted probabilities P_H of hazardous events; θ_R: Threshold on predicted probabilities P_R of restenosis; T_H: Classifier trained on patients treated with drug-eluting stents to predict hazardous events; T_R: Classifier trained on patients treated with bare-metal stents to predict restenosis;

TIMI: Thrombolysis In Myocardial Infraction; TVR: Target-vessel revascularization; U.S: United States.

Authors' contributions

SP served as primary investigator, did the implementation and evaluation of the proposed method, and drafted the paper. MS assisted in data preparation, analysis and evaluation. AmK participated in the design of the study and helped to draft the manuscript. NN participated in the design and coordination of the study at Technische Universität München. AdK organized the study, collected and prepared the data, and assisted in the clinical aspects of this work. AIK conceived of the study and coordinated it at Siemens Corporate Technology and helped to draft the manuscript. LL participated in the design of the study at Siemens Healthcare Diagnostics and helped to draft the manuscript. All authors read and approved the final manuscript.

Author details

¹Computer Aided Medical Procedures, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, Germany. ²Siemens Corporation, Corporate Technology, Imaging and Computer Vision, 755 College Rd E, Princeton, NJ, USA. ³Deutsches Herzzentrum and 1. Medizinische Klinik rechts der Isar, Technische Universität München, Lazarettstr. 36, 80636 München, Germany. ⁴Siemens Healthcare Diagnostics, Strategic Innovation Group, 511 Benedict Ave, Tarrytown, NY, USA.

References

1. Mohr FW, Morice M-C, Kappetein AP, Feldman TE, Ståhle E, Colombo A, et al. Coronary artery bypass graft surgery versus percutaneous coronary intervention in patients with three-vessel disease and left main coronary disease: 5-year follow-up of the randomised, clinical SYNTAX trial. Lancet. 2013;381:629–38.
2. Nayak AK, Kawamura A, Nesto RW, Davis G, Jarbeau J, Pyne CT, et al. Myocardial infarction as a presentation of clinical in-stent restenosis. Circ J. 2006;70:1026–9.
3. Mehran R, Dangas GD, Abizaid AS, Mintz GS, Lansky AJ, Satler LF, et al. Angiographic patterns of in-stent restenosis: classification and implications for long-term outcome. Circulation. 1999;100:1872–8.
4. Farb A, Sangiorgi G, Carter AJ, Walley VM, Edwards WD, Schwartz RS, et al. Pathology of acute and chronic coronary stenting in humans. Circulation. 1999;99:44–52.
5. Morice M-C, Serruys PW, Sousa JE, Fajadet J, Ban Hayashi E, Perin M, et al. A randomized comparison of a sirolimus-eluting stent with a standard stent for coronary revascularization. N Engl J Med. 2002;346:1773–80.
6. Moses JW, Leon MB, Popma JJ, Fitzgerald PJ, Holmes DR, O'Shaughnessy C, et al. Sirolimus-eluting stents versus standard stents in patients with stenosis in a native coronary artery. N Engl J Med. 2003;349:1315–23.
7. Ong ATL, McFadden EP, Regar E, de Jaegere P, van Domburg RT, Serruys PW. Late Angiographic Stent Thrombosis (LAST) events with drug-eluting stents. J Am Coll Cardiol. 2005;45:2088–92.
8. Bavry AA, Kumbhani DJ, Helton TJ, Borek PP, Mood GR, Bhatt DL. Late thrombosis of drug-eluting stents: a meta-analysis of randomized clinical trials. Am J Med. 2006;119:1056–61.

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 15 of 16

9. Spertus JA, Kettelkamp R, Vance C, Decker C, Jones PG, Rumsfeld JS, et al. Prevalence, predictors, and outcomes of premature discontinuation of thienopyridine therapy after drug-eluting stent placement: results from the PREMIER registry. Circulation. 2006;113:2803–9.

10. Pfisterer ME, Brunner-La Rocca HP, Buser PT, Rickenbacher P, Hunziker P, Mueller C, et al. Late clinical events after clopidogrel discontinuation may limit the benefit of drug-eluting stentsan observational study of drug-eluting versus bare-metal stents. J Am Coll Cardiol. 2006;48:2584–91.

11. Kastrati A, Dibra A, Spaulding C, Laarman GJ, Menichelli M, Valgimigli M, et al. Meta-analysis of randomized trials on drug-eluting stents vs. bare-metal stents in patients with acute myocardial infarction. Eur Heart J. 2007;28:2706–13.

12. Brar SS, Leon MB, Stone GW, Mehran R, Moses JW, Brar SK, et al. Use of drug-eluting stents in acute myocardial infarction: a systematic review and meta-analysis. J Am Coll Cardiol. 2009;53:1677–89.

13. Weisz G, Leon MB, Holmes DR, Kereiakes DJ, Popma JJ, Teirstein PS, et al. Five-year follow-up after sirolimus-eluting stent implantation results of the SIRIUS (Sirolimus-Eluting Stent in De-Novo Native Coronary Lesions) trial. J Am Coll Cardiol. 2009;53:1488–97.

14. Serruys PW, Onuma Y, Garg S, Vranckx P, De Bruyne B, Morice M-C, et al. 5-year clinical outcomes of the ARTS II (Arterial Revascularization Therapies Study II) of the sirolimus-eluting stent in the treatment of patients with multivessel de novo coronary artery lesions. J Am Coll Cardiol. 2010;55:1093–101.

15. Stone GW, Parise H, Witzenbichler B, Kirtane A, Guagliumi G, Peruga JZ, et al. Selection criteria for drug-eluting versus bare-metal stents and the impact of routine angiographic follow-up: 2-year insights from the HORIZONS-AMI (Harmonizing Outcomes With Revascularization and Stents in Acute Myocardial Infarction) trial. J Am Coll Cardiol. 2010;56:1597–604.

16. Brophy JM, Erickson LJ. Cost-effectiveness of drug-eluting coronary stents in Quebec, Canada. Int J Technol Assess Health Care. 2005;21:326–33.

17. Hill RA, Boland A, Dickson R, Dündar Y, Haycox A, McLeod C, et al. Drug-eluting stents: a systematic review and economic evaluation. Health Technol Assess. 2007;11:iii. xi–221.

18. Wijeysundera HC, Tomlinson G, Ko DT, Dzavik V, Krahn MD. Medical therapy v. PCI in stable coronary artery disease: a cost-effectiveness analysis. Med Decis Making. 2013;33:891–905.

19. Amin AP, Spertus JA, Cohen DJ, Chhatriwalla A, Kennedy KF, Vilain K, et al. Use of drug-eluting stents as a function of predicted benefit: clinical and economic implications of current practice. Arch Intern Med. 2012;172:1145–52.

20. Cavender MA, Ellis SG. Stent choice and the hidden consequences of cost savings. Nat Rev Cardiol. 2012;9:559–60.

21. Poulin F, Rinfret S, Gobeil F. Potential shift from coronary bypass surgery to percutaneous coronary intervention for multivessel disease and its economic impact in the drug-eluting stent era. Can J Cardiol. 2007;23:1139–45.

22. Varani E, Balducelli M, Vecchi G, Aquilina M, Maresta A. Comparison of multiple drug-eluting stent percutaneous coronary intervention and surgical revascularization in patients with multivessel coronary artery disease: one-year clinical results and total treatment costs. J Invasive Cardiol. 2007;19:469–75.

23. Vieira RD, Hueb WA, Hlatky MA, Favarato D, Rezende PC, Garzillo CL, et al. Cost-effectiveness analysis for surgical, angioplasty, or medical therapeutics for coronary artery disease: 5-year follow-up of medicine, angioplasty, or surgery study (MASS) II trial. Circulation. 2012;126 Suppl 1:S145–50.

24. Cohen DJ, Lavelle TA, Van Hout B, Li H, Lei Y, Robertus K, et al. Economic outcomes of percutaneous coronary intervention with drug-eluting stents versus bypass surgery for patients with left main or three-vessel coronary artery disease: one-year results from the SYNTAX trial. Catheter Cardiovasc Interv. 2012;79:198–209.

25. Magnuson EA, Farkouh ME, Fuster V, Wang K, Vilain K, Li H, et al. Cost-effectiveness of percutaneous coronary intervention with drug eluting stents versus bypass surgery for patients with diabetes and multivessel coronary artery disease: results from the FREEDOM trial. Circulation. 2013;127:820–31.

26. Farooq V, van Klaveren D, Steyerberg EW, Meliga E, Vergouwe Y, Chieffo A, et al. Anatomical and clinical characteristics to guide decision making between coronary artery bypass surgery and percutaneous coronary intervention for individual patients: development and validation of SYNTAX score II. Lancet. 2013;381:639–50.

27. Raza S, Sabik JF, Ellis SG, Houghtaling PL, Rodgers KC, Stockins A, et al. Survival prediction models for coronary intervention: strategic decision support. Ann Thorac Surg. 2014;97:522–8.

28. Ndrepepa G, Braun S, Schömig A, Kastrati A. Accuracy of N-terminal pro-brain natriuretic peptide to predict mortality in various subsets of patients with coronary artery disease. Am J Cardiol. 2007;100:575–8.

29. Ndrepepa G, Braun S, Mehilli J, Birkmeier KA, Byrne RA, Ott I, et al. Prognostic value of sensitive troponin T in patients with stable and unstable angina and undetectable conventional troponin. Am Heart J. 2011;161:68–75.

30. Dangas GD, Claessen BE, Caixeta A, Sanidas EA, Mintz GS, Mehran R. In-stent restenosis in the drug-eluting stent era. J Am Coll Cardiol. 2010;56:1897–907.

31. Kibos A, Campeanu A, Tintoiu I. Pathophysiology of coronary artery in-stent restenosis. Acute Card Care. 2007;9:111–9.

32. Breiman L. Classification and Regression Trees. New York, NY, USA: Chapman and Hall; 1984.

33. Wiviott SD, Cannon CP, Morrow DA, Murphy SA, Gibson CM, McCabe CH, et al. Differential expression of cardiac biomarkers by gender in patients with unstable angina/non-ST-elevation myocardial infarction: a TACTICS-TIMI 18 (Treat Angina with Aggrastat and determine Cost of Therapy with an Invasive or Conservative Strategy-Thrombo. Circulation. 2004;109:580–6.

34. Fortmann SP, Ford ES, Criqui MH, Folsom AR, Harris TB, Hong Y, et al. CDC/ AHA workshop on markers of inflammation and cardiovascular disease: application to clinical and public health practice: report from the population science discussion group. Circulation. 2004;110:e554–9.

35. Gilbert J, Raboud J, Zinman B. Meta-analysis of the effect of diabetes on restenosis rates among patients receiving coronary angioplasty stenting. Diabetes Care. 2004;27:990–4.

36. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics. 1970;12:55–67.

37. Le Cessie S, Van Houwelingen JC. Ridge estimators in logistic regression. Ser C (Applied Stat). 1992;41:191–201.

38. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. SIGKDD Explor Newsl. 2009;11:10–8.

39. Serruys PW, Ong ATL, van Herwerden LA, Sousa JE, Jatene A, Bonnier JJRM, et al. Five-year outcomes after coronary stenting versus bypass surgery for the treatment of multivessel diseasethe final analysis of the arterial revascularization therapies study (ARTS) randomized trial. J Am Coll Cardiol. 2005;46:575–81.

40. Rodriguez A, Bernardi V, Navia J, Baldi J, Grinfeld L, Martinez J, et al. Argentine randomized study: coronary angioplasty with stenting versus coronary bypass surgery in patients with multiple-vessel disease (ERACI II): 30-day and one-year follow-up results. J Am Coll Cardiol. 2001;37:51–8.

41. Morrison DA, Sethi G, Sacks J, Henderson W, Grover F, Sedlis S, et al. Percutaneous coronary intervention versus coronary artery bypass graft surgery for patients with medically refractory myocardial ischemia and risk factors for adverse outcomes with bypass: a multicenter, randomized trial. J Am Coll Cardiol. 2001;38:143–9.

42. Investigators TS. Coronary artery bypass surgery versus percutaneous coronary intervention with stent implantation in patients with multivessel coronary artery disease (the Stent or Surgery trial): a randomised controlled trial. Lancet. 2002;360:965–70.

43. Hueb WA, Soares PR, Gersh BJ, César LAM, Luz PL, Puig LB, et al. The medicine, angioplasty, or surgery study (MASS-II): a randomized, controlled clinical trial of three therapeutic strategies for multivessel coronary artery diseaseOne-year results. J Am Coll Cardiol. 2004;43:1743–51.

44. Pohl T, Giehrl W, Reichart B, Kupatt C, Raake P, Paul S, et al. Retroinfusion-supported stenting in high-risk patients for percutaneous intervention and bypass surgery: results of the prospective randomized myoprotect I study. Catheter Cardiovasc Interv. 2004;62:323–30.

45. Eefting F, Nathoe H, van Dijk D, Jansen E, Lahpor J, Stella P, et al. Randomized comparison between stenting and off-pump bypass surgery in patients referred for angioplasty. Circulation. 2003;108:2870–6.

46. Reeves BC, Angelini G, Bryan AJ, Taylor FC, Cripps T, Spyt TJ, et al. A multi-centre randomised controlled trial of minimally invasive direct coronary bypass grafting versus percutaneous transluminal coronary angioplasty with stenting for proximal stenosis of the left anterior descending coronary artery. Health Technol Assess. 2004;8:1–43.

47. Cisowski M, Drzewiecki J, Drzewiecka-Gerber A, Jaklik A, Kruczak W, Szczeklik M, et al. Primary stenting versus MIDCAB: preliminary report–comparison of two methods of revascularization in single left anterior descending coronary artery stenosis. Ann Thorac Surg. 2002;74:1334–9.

48. Kim JW, Lim D-S, Sun K, Shim WJ, Rho YM. Stenting or MIDCAB using ministernotomy for revascularization of proximal left anterior descending artery? Int J Cardiol. 2005;99:437–41.

Pölsterl *et al. BMC Medical Informatics and Decision Making*

Page 16 of 16

49. Serruys PW, Morice M-C, Kappetein AP, Colombo A, Holmes DR, Mack MJ, et al. Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. N Engl J Med. 2009;360:961–72.

50. Hong SJ, Lim D-S, Seo HS, Kim Y-H, Shim WJ, Park CG, et al. Percutaneous coronary intervention with drug-eluting stent implantation vs. minimally invasive direct coronary artery bypass (MIDCAB) in patients with left anterior descending coronary artery stenosis. Catheter Cardiovasc Interv. 2005;64:75–81.

51. Kappetein AP, Feldman TE, Mack MJ, Morice M-C, Holmes DR, Ståhle E, et al. Comparison of coronary bypass surgery with drug-eluting stenting for the treatment of left main and/or three-vessel disease: 3-year follow-up of the SYNTAX trial. Eur Heart J. 2011;32:2125–34.

52. Burman P. A comparative study of ordinary cross-validation, v -fold cross-validation and the repeated learning-testing methods. Biometrika. 1989;76:503–14.

53. Sianos G, Morel M-A, Kappetein AP, Morice M-C, Colombo A, Dawkins KD, et al. The SYNTAX score: an angiographic tool grading the complexity of coronary artery disease. EuroIntervention. 2005;1:219–27.

54. Epstein AJ, Polsky D, Yang F, Yang L, Groeneveld PW. Coronary revascularization trends in the United States, 2001–2008. JAMA. 2011;305:1769–76.