# Modeling Re-Admission Risk after Endourological Surgeries

P. Crespo Darriba[1], S. Vera Carretero[1],
O. Barquero Pérez[2], M. Sánchez Encinas [3]

[1] Grado en Ingeniería Biomédica, Campus de Alcorcón, Universidad Rey Juan Carlos.
[2] Escuela Técnica Superior de Ingeniería en Telecomunicación, Universidad Rey Juan Carlos.
[3] Departamento de Urología, Hospital Universitario Rey Juan Carlos.

## Abstract

*Transurethral Resection of Bladder Tumor (TURBT) and Transurethral Resection of Prostate (TURP) are two of the most frequent endourological interventions nowadays. Their greatest disadvantage is the cause of several postoperative complications (being haematuria the most common) which lead to hospital readmission. Hence, in this work we propose a machine learning model defined by a Random Forest Classifier to predict the probability of readmission. The final results reveal that features such as the amount of creatinine, leucocytes and PSA are some of the most important measurements to avoid and control this type of readmissions. This leads the doctors with an additional mechanism to measure the severity of readmission a patient can have.*

## 1 Introduction

Endourology is the specific area in urology in which minimally invasive surgical techniques are used to perform surgery and treat diseases in the urinary tract. Two of the most common surgical procedures are Transurethral Resection of Bladder Tumor (TURBT) and Transurethral Resection of Prostate (TURP). TURBT is the leading surgical procedure used to diagnose, stage, and treat the presence of invasive bladder tumors. Three goals are pursued: determine number and size of tumors, obtain pathological specimens to determine the histology, and remove all visible invasive tumors [1,2]. TURP is considered the surgery option to treat men with moderate or severe urinary problems caused by benign prostatic hyperplasia (BPH), which is an enlargement of the prostate [3].

Although these conditions are eminently different, both surgical procedures share great part of the methodology. Transurethral resection (TUR) is performed with an endoscopic device called resectoscope, which includes a wide-angle microscope to allow visualization and a wire loop, which will cauterize the possible tumor area. The resectoscope is introduced by the urethral lumen and continues through the urinary tract until it reaches the surgical site, where the tumor or prostate section is cut and cauterized. The surgical area is constantly being irrigated with fluid (usually glycine 1.5%) to wash away blood and any possible debris, to improve visibility. This procedure requires either general, or spinal anesthesia, being the latter technique the most used. At the end of the process, a three-lumen catheter is inserted, and irrigation is continued for up to 24 hours after operation [4].

These surgeries can lead to potential intraoperative problems, such as myocardial ischemia (25% of TURP patients), myocardial infarction occurring in 2% of surgeries or unwanted perforations of urethra, bladder, or prostatic capsule [4].

Postoperative complications are also possible in the following hours or days after the procedure. Most of them, are graded as low risk complications including bleeding, urinary inconsistency, or gross hematuria. Although, these minor complications, may lead to the readmission of the patient [5, 6], which triggers an increase in patint recovery time and enormous growth on cost of care, which might be avoided reducing readmission rates.

Postoperative Hematuria is one of the most common causes for patient readmission, with incidences up to 7% TURP [7], and 9% in the case of TURBT.

## 2 Dataset and Exploratory Data Analysis

The study consists in using ensemble models to predict the readmission probability in patients who have been involved in these types of surgeries using data from the urological unit of the Hospital Universitario Rey Juan Carlos (HURJC) who were registered between 2015 and 2020.

These patients had to undergo surgery due to different urological disorders from the prostate and bladder. This criterion leads the data to be divided into two groups, bladder and prostate sets, each one with different features.

The dataset also contains a wide variety of variables that helps to assess the patient status both from the quantitative and qualitative point of view and before and after the surgery. It is composed by numeric variables, which corresponds to anthropometric (weight, size, IMC), physiological (leukocytes, creatinine, PSA), times (Surgery Duration, Diagnosis Date, Surgery Date) measurements; categorical variables for classification

systems (ASA, ECOG, Clavien complications); binary variables indicating the presence of urological problems (vesical lithiasis, hydronephrosis); and text variables which give extra information about the patients and their surgical interventions (Observations and incidents, Findings).

However, these variables, the ones purely coming from the source are not the final ones. They went through a deep analysis in order to select which ones were valid to work with. This procedure will be explained in subsequent sections.

The outcome variable corresponds to hospital readmission and it was not defined, it had to be created. It was assumed that the patients with a date of hospital readmission, had to be readmitted but the people who did not have this information, were considered as no readmitted. Hence, the outcome is converted into binary variable with 1 for readmission and 0 for no readmission.

The total number of patients involved was 862. For the bladder group, there were 523 patients of which 62 were readmitted to the hospital, while for the prostate group, 339 patients were implicated and 20 of them were readmitted. It should be highlighted that these numbers are translated into a 12% of readmitted people in the case of bladder and a 5% in the case of prostate patients. This implies a considerable imbalance in the information of both outcomes, which means in not much information related to readmission.

## 2.1 Pre-processing of the data

The pre-processing of the data represents one of the most important and critic steps of the whole project. This is because the data that come from its source of origin is usually not easy to handle and representative to get good results.

Consequently, the pre-processing process has been divided into three phases, Data Preparation, Treatment of Missing Values and Feature Selection.

### *2.1.1 Data Preparation*

Raw patient data was stored in an excel file with two sheets, one for the bladder and other for the urethra. All patients had different variables stored in different rows. It was transformed into two different excels files, one for the bladder and the other one for the prostate. Every row corresponded to each patient and the columns corresponded to all the possible variables that appeared along the different sheets, assigning NaN (Not a Number) values to the cells in which there was no information.

### *2.1.2 Treatment of Missing Values and Outliers*

Subsequently, the problematic with missing values was analysed as they correspond to incomplete information which will not allow the algorithm to work.

The percentage of missing values was computed for every feature of the bladder and the urethra sets and the following decisions were taken.

- Missing values from binary variables where transform to 0 as it has been assumed that they were not indicated because they do not proceed.

- Variables with more than 25% of missing values were removed from the bladder group while those with more than 40% were dropped from the urethra group.

- For the rest of features, when the division into training and test sets was done, missing values from numerical variables were replaced by the median value and those corresponding to the categorical type, were substitute by the mode. All of these values were obtained from the training set and are used to impute both training and test set.

On the other hand, the numerical features were analysed and those values which present an extreme or abnormal behaviour were eliminated. An example of this is the 'IMC' variable, which presented values higher than 100 which correspond to data errors.

### *2.1.3 Feature Selection*

In order to do a first selection of variables that will train the proposed model, the following was performed.

- Variables with unique values. Features with only one value were eliminated as they will not provide any differentiation from the two outcome groups.

- Variables with text values. Due to the complexity of working with variables which contain a whole description instead of numerical or defined categorical values, they have been removed to the final dataset. These features could be processed with Natural Language Processing (NLP) techniques to extract categorical knowledge. However, this technology is still developing and needs huge volume of data available to be precise.

- Variables with date information. The variables with date information were transformed into time measurements. One was used to define the time from the diagnosis until the surgery intervention. And the other, the time between the surgery date and the date the patient abandoned the hospital. Both were measured in days. Subsequently, the variables which correspond to these four dates were eliminated. Besides, two more dates that corresponded to inaccessible data for future patients were also removed. They reference to complications and readmission dates, which will be dependent of the possible outcome.

After this process, we kept 49 features for TURP database and 23 for the TURBT database. Their names and meanings can be found in Appendix A, at the end of the document.

## 2.2 Exploratory Data Analysis

The relevance of certain features and their relationship with the outcome have been firstly explored by their visualization.

For categorical variables, a graphical representation of the proportion of people that belong to certain values has been used due to their easy visual comprehension. ASA corresponds to a classification system used to estimate the physical status that can be observed in the patient and its values are I,II,IIIa,IIIb, IIIc and IV, being the I the healthiest and IV the worst scenario. It has been picked because it has been shown some kind of influence in the postoperative behaviour of the person.

Plots have been normalized for a better visualization of the proportion of the values in each of the outcome sets.

Conversely, histograms have been selected to describe numerical features. In this case, the selection of Creatinine follows a similar logic as before. It shows how well the kidneys are working.

These visualizations give us a good first approach to understand the data and their dependence on the outcome. However, it does not always reveal accurate information. This implies a future deeper analysis with statistical techniques.

In order to be more exact, the exploratory data analysis of both groups, bladder and prostate, were done separately. This system help to see what is happening in both cases individually.

In the first place, when analysing the TURBT, visualizations like the one in Figure 1 and Figure 2 are obtained.

In Figure 1, it can be appreciated that the lowest level of ASA, the classification system that estimates the physical status of the patient, are presented in people that has not been readmitted while the highest are more present in the readmitted group.

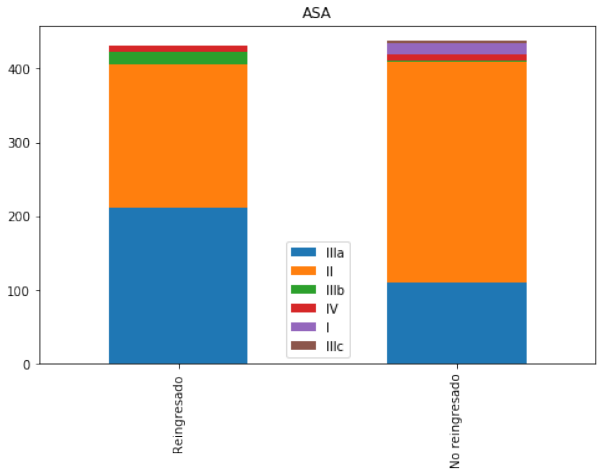


***Figure 1.*** *Graphical representation of the proportion of people from the bladder set that belongs to the different values of the ASA feature.*

In Figure 2, it appears that the higher values of creatinine are related to the readmission of patients to the hospital. What makes sense as an elevated amount of creatinine may be a sign of poor kidney function.

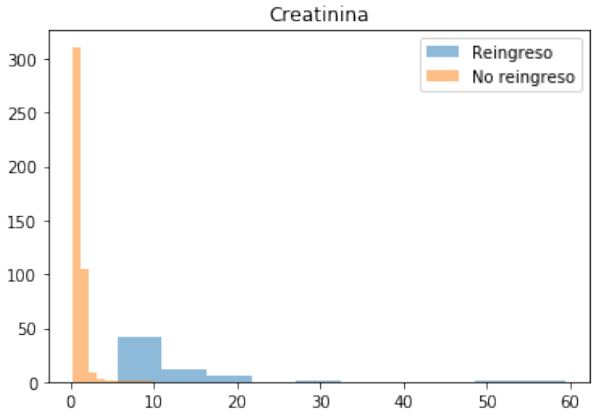


***Figure 2.*** *Creatinine histogram for the bladder set.*

Both figures seems to represent variable that contain information for the predictor and their decisions.

Then, in the case of TURP, the same procedure as before has been done.

In Figure 3, the differentiation between levels of ASA is not as evident as in the TURBT case. However, it can be perceived that level I is only present for no readmitted patients while level IIIa is only present in readmitted patients.

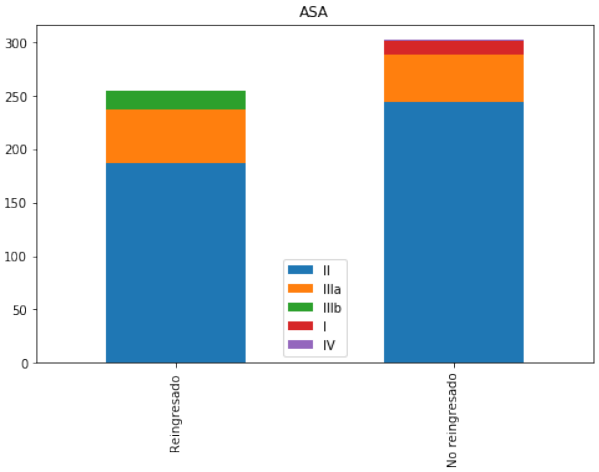


***Figure 3.*** *Graphical representation of the proportion of people from the prostate set that belongs to the different values of the ASA feature*

Figure 4 shows that for this set of data, the creatinine feature can be useful and informative when deciding who could be readmitted to the hospital. Higher values lead the positive outcome.

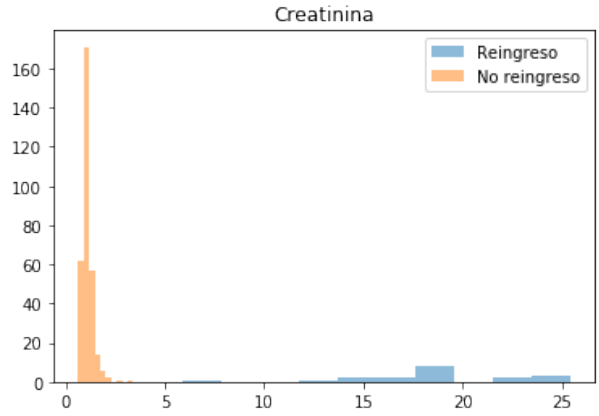


***Figure 4.*** *Creatinine histogram of the prostate set.*

The same conclusion as before can be drawn. Both features looks like representative for the study.

# 3 Methods and Models

## 3.1 Statistical Analysis

Machine learning methods are the most powerful techniques for prediction and data analysis. However, the first step in these problems must be to analyze and understand the dataset. Hence, the study of feature interactions and significances is crucial to approach the problem with the correct perspective.

As we discussed in Section 2, datasets contained a high proportion of features which needed to be discarded due to the high number of missing values. Remaining features will be studied and analyzed using statistical methods in order to obtain a clearer vision of the problem. Furthermore, these results might improve the interpretability for the prediction model.

The aim of this analysis is to test whether the classes form a specific feature, have significant differences in the Outcome result or not. In other words, we will observe the contribution of each feature for the classification. Can we ensure that the class distribution in Figure 2 give us some information about the future development of hematuria? Visual approach may suggest it does, but we will use hypothesis testing methods to proof the level of significance.

In this section we will accept variables as random distributions and divide them in 2 types, Numerical and Categorical.

Numerical features are first tested for normality, which is how well a gaussian distribution fits data values. Among the wide range of statistical methods available, we selected Shapiro-Wilk (SW) test, which is gold standard for normality testing. Features which failed to be fitted by a gaussian distribution will go through non-parametric tests (Wilcoxon test). Meanwhile, normal distributed features, will be analyzed with parametrical tests (T-test hypothesis).

Then, values of the feature are split into normal or abnormal population based on values determined by physicians [9].

Categorical features are in turn, split in populations by their classes, and tested both, between all classes (chi-squared test) and in pairwise testing (Z-statistic)

Hypothesis testing is then used to compare both populations. Null hypothesis ($H_0$) is defined as populations came from the same data distribution, otherwise, Alternative hypothesis ($H_1$) is the assumption that populations are significantly different. To conduct the test, we observe p-value, which is the probability that the difference between population means is at least as large as observed given $H_0$. If p-value is smaller than a significance level (0.05), we reject $H_0$, and assume that populations are not obtained from the same distribution.

## 3.2 Machine Learning Algorithms

Once our datasets are cleaned, we are ready create the prediction models. Due to the large variety of available artificial intelligence methods, we must recall the original problem, which is the prediction of future patient's readmission caused by a specific complication. These issues are called classification problems, and they are normally addressed with supervised machine learning techniques. The strength of these algorithms is that they use the data imputed in the model, to understand how the different feature values explain an outcome feature (Readmission caused by Hematuria). If the model is correctly trained, it should generalize to unseen situations and give an accurate prediction for each new case.

The standard methodology in these algorithms is to split data in 2 subsets, training set and testing set. The first one will be used to train the model, afterwards, the testing set serves to measure the reliability of the model. However, this procedure is not possible with unbalanced databases, as ours. Available training data is quite limited, so splitting our datasets, would hinder the generalization of the model. Thus, we must follow cross-validation testing and other techniques, which allow us to train the model and improve generalization, while using the largest possible amount of training data and provide reliable performance results.

The reliability of the classifier depends on the quality of datasets and the feature extraction (discussed in section 2), in addition to the choice of the prediction method and its hyperparameters, as we will discuss now.

There is a wide range of algorithms, each built with its own learning strategy. Therefore, it is crucial to select 'good' classifiers for approaching the problem with the correct perspective and achieve an accurate and precise prediction. In addition, interpretability and other factors must be considered for the selection of the machine learning algorithm.

### 3.2.1 *Logistic Regression*

Regression is a statistical method, in which the outcome or dependent feature, is explained based on a function linearly combining other variables which are called independent features or predictors. The simplest machine learning algorithm is a linear regression between the predictors and the outcome. This algorithm predicts a continuous variable as the outcome as it is seen in Figure 5.

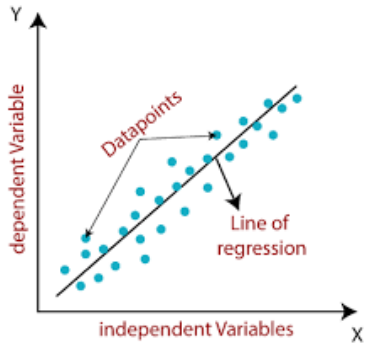


***Figure 5.*** *Linear regression plot (Medium.com)*

Nevertheless, what if the model outcome is not a continuous variable, but a binary feature? In this case, prediction takes the value 0 or 1, and so do the possible errors[10]. In logistic regression, the model predicts for each observation, the probability odds (log-odds) of belonging to category '1' by a linear combination of the predictors. Then, a logistic function creates a map from Log-odds $(-\infty, +\infty)$ to probabilities (0,1).

$$\sigma(\omega^T x) = P(y = 1 \mid \omega, x) = \frac{1}{1 + e^{-\omega^T x}}$$

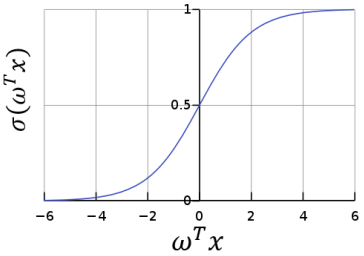


***Figure 6.*** *Logistic regression function (Wikipedia)*

Then, the algorithm rounds the probability either to 0 or 1, which will be the output of the model.

This method is the simplest machine learning parametric classifier. Therefore, it is usually used for building a first simple model and see how precise the results are.

### 3.2.2 *Random Forest*

Classification problems in clinical domain are normally solved by Decision Tree Classifiers for their closely resemble to human reasoning and ease to be interpretated. These models use a supervised nonparametric perspective to perform the classification between 2 or more classes, In our case 'Readmitted' (0) or 'Not Readmitted' (1) [11].

They are based on a sequential workflow of simple threshold tests. All samples start in the Root Node, where a test compares a feature value against a threshold predefined in the training phase. This threshold, branches the Root Node into 2 sub-nodes, where other tests will be made to continue with the recursive divisions until all the samples of a node belong to the same class.

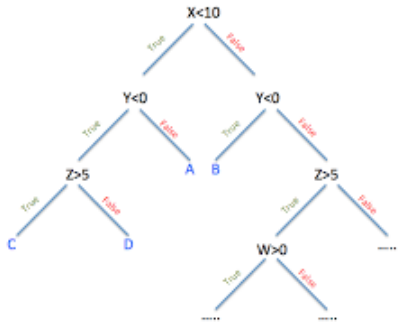


***Figure 7.*** *Decision Tree Classifier Diagram (Medium.com)*

Each test chooses the variable which performs 'better' the split . There are two main splitting methods: Gini Impurity, which measures the probability of mislabeling an observation; and Entropy, which is related to uncertainty between the feature values distribution and the outcome.

As a consequence of this learning procedure, we can score the importance of each feature on the classification task. These scores help to interpret the model results, and they can also be used as a feature selection technique. We want our prediction model to work with as much data as we can obtain, but sometimes this data is not informative, which we call noise leakage. Feature importances show this information, so the model is only imputed with 'good' data, deleting features which do not help in the classification task.

One drawback of these methods is the tendency to overfit very quickly. Overfitting happens when the model fits so well to training data that cannot generalize well to new cases. The embedded overfitting solution in Decision Trees is 'Pruning'. This technique changes some of the tree hyperparameters, so final nodes (leaves) are not too specialized on the training data, and the model can perform well with new observations.

The search for the optimal combination of hyperparameters can be challenging and requires high computational cost. We use Grid Search, which works defining the set of hyperparamters as a n-dimensional space, where each hyperparamter is a dimension and the optimal combination is the settings with the best performance [12].

We showed the configuration of decision trees, but the method we are discussing is Random Forest. This

algorithm consists on a large set of decision trees, in which the final output depends on the classification output of the whole set. Each tree shows its prediction, and the most voted class, becomes the model prediction.

### 3.3 Prediction Model

The proposed solution for the classification task is a Random Forest Classifier, with the following procedure

1. Definition of specific hyperparameters. Not all hyperparameters are equally important, some of them will work on 'default' setting, such as the splitting method. In our case, we search the optimal values for: number of Decision Trees, maximum tree depth, and minimum number of samples per leaf.

2. Train the model with 10 fold cross-validation, using stratified random sampling. On each iteration, missing values are replaced with statistical methods using the median of non-missing values.

3. In order to deal with feature imbalance, we perform data oversampling. The least represented class is randomly re-sampled until equal class distribution is reached.

4. Train the model with the described configuration. Feature importances are computed, selecting the variables which add most information to the model and discarding the least relevant ones. As the number of features is reduced, there is a reduction in curse of dimensionality, in addition to lower computational cost, and lower levels of noise.

5. Re-Train the classifier only with remaining features, and validate the performance.

The model was developed in Python 3.7.3 software, using the library Scikit-Learn for analysis and machine learning purposes, in addititon to other elemental libraries such as NumPy, Pandas or MatplotLib.

## 4 Results

For this section, the results of both databases are analysed separetely in order to deepen in their meaning.

In the case of TURBT, Statistical Hypothesis Analysis was computed in clinical features. P-Values lower than significance level (0.05) were considered as statistically significant.

| Feature | p-Values (k=0.05) |
|---|---|
| Número Tm | 0.01 |
| Localización | 0.09 |
| RTU Vesical | 0.00 |
| Posición | 0.92 |
| Tamaño | 0.00 |
| ASA | 0.01 |
| Leucocitos | 0.14 |
| Creatinina | 0.00 |

***Figure 8.*** *Statistical Hypothesis Analysis for TURBT*

After the first model was trained, most relevant features are selected using cross-validated feature importances. The 7 features with highest importance, were proved to keep the 89% of model certainty, reducing in 70% the dimensionality. These features are 'Creatinina', 'Leucocitos', 'Tiempo de Cirugía' 'Tamaño (mm)' 'tiempo-ingr', 'RTU Vesical' and 'ASA-II'.

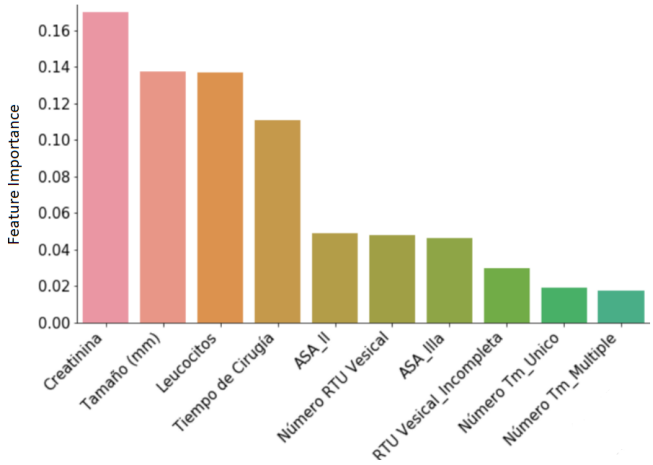


***Figure 9.*** *Highest Feature Importances for TURBT*

We find similarities between feature importances shown in Figure 9, and the hypothesis tests we observe in Figure 8. All of selected features were reported as being significant in the statistical analysis, except for 'Leucocitos', which we assume that could have a non-linear dependency in relation to 'Outcome' variable. We can observe this relationship in figure 11. The relation is clearly non-linear. Since the applied statistical analysis only determines linear dependencies, groups were defined as not different, but the algorithm was strong enough to find the relationship.

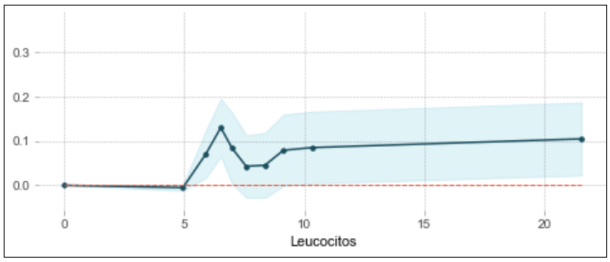


***Figure 10.*** *'Leucocitos' dependence plot. It shows the relevance variation of the feature to make the prediction by changing its value*

After the training and validation phases on the different models, we achieved the results observed in Table 1.

| Model | F1 Score | Accuracy | ROC | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Random Forest | 0.278 | 0.816 | 0.602 | 0.904 | 0.26 |
| Feature Selection Random Forest | 0.323 | 0.829 | 0.624 | 0.913 | 0.322 |
| Feature Selection Log. Regression | 0.306 | 0.739 | 0.637 | 0.917 | 0.231 |

***Table 1.*** *Metrics for TURBT*

On the other hand, for the TURP cases, although the

procedure has been the same as in TURBP, some other different conclusions has been observed.

Figure 11 shows a table with the results from the Statistical Hypothesis Analysis. As before, P-Values lower than 0.05 were considered as statistically significant which means that classes are significantly different, being considerably relevant for the study. Hence, at first it seems that 'PSA libre', 'Creatinina' and 'IMC' are worthy to consider.

| Feature | p-Values |
|---|---|
| PSA libre | 0.04 |
| IMC | 0.04 |
| Litiasis Vesical | 0.60 |
| Posición | 0.89 |
| Creatinina | 0.04 |
| PSA Sérico | 0.10 |
| Bulbar | 0.84 |

***Figure 11.*** *Statistical Hypothesis Analysis for TURP*

Subsequently, the first Random Forest Classificator was trained obtaining as most relevant features the ones that appear in Figure 12. These are 'PSA libre', 'Creatinina', 'PSA Sérico', 'IMC', 'Hb. Post (a 24h)', 'Vol. Próstata ', 'ASA', 't1', 't2', 'TR', 'Tiempo de Cirugía', 'Sistema de Lavado Vesical', 'Lóbulo Medio'.

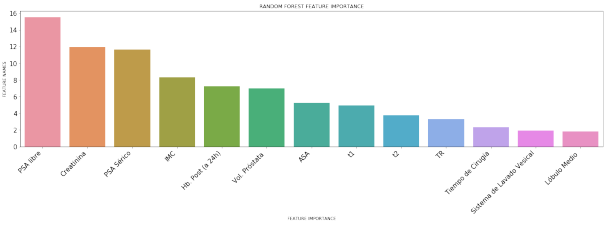


***Figure 12.*** *Highest Feature Importances for TURP*

The results previously obtained with the Statistical Hypothesis Analysis largely coincide with the information of the feature importance technique of the Random Forest model.

With this last method 'PSA Sérico' feature is considered as important, but in the Statistical Hypothesis Analysis not. This could imply as before that this variable does not have a linear dependence with respect to the outcome. However, it does not make sense as it also depend on the 'PSA libre' which seems to be linear dependent. It may happen due to the great amount of missing values. This could have altered the importance with in either of the methods.

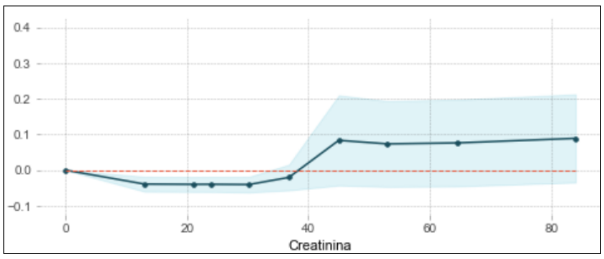


***Figure 13.*** *'Creatinina' dependence plot. It shows the relevance of the feature to make the prediction depending on its value*

In Figure 13, it can be appreciate how the importance of the variable 'Creatinina' changes depending on the values of it.

Finally, in Table 2 there is a comparisson between the results obtained with a Logistic Regression model, the proposed Random Forest classificator without feature selection by means of feature importance and the other one with it.

| Model | F1 Score | Accuracy | ROC | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Random Forest | 0.126 | 0.862 | 0.529 | 0.914 | 0.165 |
| Feature Selection Random Forest | 0.214 | 0.885 | 0.585 | 0.931 | 0.325 |
| Feature Selection Log. Regression | 0.096 | 0.373 | 0.526 | 0.968 | 0.054 |

***Table 2.*** *Metrics for TURP*

## 5 Conclusions

Endourological surgeries are increasingly common due to their minimally invasive technique which avoids the greatest number of complications during and after the surgery. However, unforeseen events can still appear, being hemorrhage the most usual cause. This is the reason why health care professionals continue studying how to reduce these issues until reaching the best possible scenario.

In addition, statistical analysis and artificial intelligence are currently used procedures that try to achieve the most accurate predictions in future events. Consequently, they can be applied to problems like the one exposed so that doctors can have an instrument that helps them how to take decisions depending on the probability of what will happen in the future.

After comparing different methods, it can be concluded the features mentioned in the results are decisive for our goal, being 'PSA libre' and 'Creatinina' highly informative to determine the behaviour of the patient after surgical intervention. And although the metrics of the prediction does not appear with unsatisfactory results, they are not completely reliable. The percentage of readmission in the dataset is considerably low which leads to a bias in our model towards no readmission. This means that the model usually predicts no readmitted and guess it correctly but because it is also usually tested with more no readmission patients. The f1 metric is what informs about this certainty.

It should be highlighted that the features previously mentioned characterized in an optimal manner which implies a possible application in the daily work.

Nevertheless, the results could be improved adding more features to the dataset like the age, also more patients to the dataset, especially of those who have been readmitted. Also, more complex techniques could be applied for the pre-processing like introducing data with Natural Language Processing (NLP) techniques instead of assuming it as noise.

Finally, it is also noteworthy that these analytical techniques could be scalable to other type of problems inside this area of the hospital or even, to other healthcare areas.

## References


[1] Gregg, Justin R., et al. "Short term complications from transurethral resection of bladder tumor." The Canadian journal of urology 23.2 (2016): 8198-8203.

[2] Herr, Harry W., and S. Machele Donat. "Quality control in transurethral resection of bladder tumours." BJU international 102.9b (2008): 1242-1246.

[3] Rassweiler, Jens, et al. "Complications of transurethral resection of the prostate (TURP)—incidence, management, and prevention." European urology 50.5 (2006): 969-980.

[4] Aidan M. O'Donnell, BSc MB ChB FRCA, Irwin T.H. Foo, MB BChir MD MRCP(UK) FRCA, Anaesthesia for transurethral resection of the prostate, Continuing Education in Anaesthesia Critical Care & Pain, Volume 9, Issue 3, June 2009, Pages 92–96

[5] Bansal, Ankur et al. "Grading of complications of transurethral resection of bladder tumor using Clavien-Dindo classification system." Indian journal of urology : IJU : journal of the Urological Society of India vol. 32,3 (2016): 232-7. doi:10.4103/0970-1591.185104

[6] Bolat, Deniz et al. "Comparing the short - term outcomes and complications of monopolar and bipolar transurethral resection of bladder tumors in patients with coronary artery disese: a prospective, randomized, controlled study." International braz j urol : official journal of the Brazilian Society of Urology vol. 44,4 (2018): 717-725. doi:10.1590/S1677-5538.IBJU.2017.0309

[7] Shamout, S., Carlson, K., Brotherhood, H. L., Crump, T., & Baverstock, R. (2020). Incidence and Predictors of Early and Late Hospital Readmission Following Transurethral Resection of The Prostate: A Population-Based Cohort Study.

[8] Ghali, F., Moses, R. A., Raffin, E., & Hyams, E. S. (2016). What factors are associated with unplanned return following transurethral resection of bladder tumor? An analysis of a large single institution's experience. Scandinavian Journal of Urology,50(5),370–373.

[9] Creatinine Blood Test (Normal, Low, High Levels) Medical Author: Charles Patrick Davis, MD, PhDMedical Editor: William C. Shiel Jr., MD, FACP, FACR. MedicineNet.

[10] Hosmer Jr, David W, Stanley Lemeshow, and Rodney X. Sturdivant. Applied logistic regression. Vol. 398. John Wiley & Sons, 2013.

[11] Kotsiantis, S.B. Decision trees: a recent overview. Artif Intell Rev 39, 261–283 (2013). https://doi.org/10.1007/s10462-011-9272-4

[12] James Bergstra and Yoshua Bengio. 2012. Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, null (3/1/2012), 281–305.

APPENDIX A. VARIABLES NAMES AND THEIR MEANING

| **Variable names** | **Meaning** |
|---|---|
| Indicación | Failure of medical treatment, hypocontractile bladder or acute urinary retention |
| ASA | Classification system to estimate the physical status |
| IMC | Corporal index mass |
| Bulbar | Bulbar urethra strictures |
| Posición | Surgical position |
| RTU-P Bipolar | Bipolar Transurethral Resection of Prostate |
| Tipo RTU Vejiga | Transurethral Resection of Bladder |
| RTU Vesical | Transurethral Resection of Bladder Tumor |
| Hb. Post (a 24h) | Postoperative hemoglobin levels |
| TR | Rectal Examination |
| Anticoagulación / Antiagregación | Anticoagulation drugs |
| PSA Sérico | Protein produced exclusively by prostate cells |
| PSA libre | Free-PSA test measures the percentage of unbound PSA |
| Leucocitos | Amount of leucocytes |
| Tipo Cirugía | Type of endourlogical surgery |
| Alta Sonda Vesical | Urinary catheter |
| Sonda Vesical Previa | Previoius urinary catheter |
| Membranosa | Membranous urethra |
| Complicaciones Clavien | Postoperative urological complications |
| Muestras Recogidas (Microbiológica y A. Pato.) | Microbiology/Pathological Anatomy Samples |
| Vol. Próstata | Prostate volume |
| Tiempo de Cirugía | Surgery duration |
| Glande | Glans |
| Divertículos vesicales | Presence of vesical diverticulums |
| Litiasis Vesical | Vesical Lithiasis |
| Hidronefrosis | Hydronephrosis |
| Navicular | Fossa Navicularis |
| Lóbulo Medio | Median lobe |
| Residuo | Wastes |
| Diagnóstico | Diagnosis |
| Creatinina | Creatinine |
| Via Acceso | Surgical Access Route |
| Cirugía Previa Abdominal | Previous abdominal surgery |
| Sistema de Lavado Vesical | Vesical Wash System |
| Tipo Complicaciones | Complications |
| Sangrado ml | ml of bleeding |
| Tiempo vaporización | Duration of the vaporization |
| Fármacos In Situ | Patients drugs |
| Prótesis y/o Implantes | Prosthesis/Implants |
| Drenajes | Surgical drainage |
| Número de transfusiones | Amount of blood transfusion |
| t1 | Time from diagnosis until the surgery |
| t2 | Time from surgery until discharge from the hospital |
| Reingreso | Outome of patient readmission |