UNIVERSITÀ DEGLI STUDI DELL'AQUILA
DIPARTIMENTO DI SCIENZE CLINICHE APPLICATE E BIOTECNOLOGICHE

Dottorato di Ricerca in MEDICINA SPERIMENTALE

Curriculum Scienze cliniche, mediche e chirurgiche

XXXVI ciclo

EVOLUTION OF SURGICAL TECHNIQUES IN PROCTOLOGY: PREOPERATIVE RISK ASSESSMENT WITH MACHINE LEARNING MODELS

SSD: MED/18 - Chirurgia Generale

Dottorando

LUCIA ROMANO

Lucia Romano

Coordinatrice del corso

Prof.ssa Mariagrazia Perilli

Mariagrazia Perilli

Tutor

Prof. Antonio Giuliani

Antonio Giuliani

Co-tutor

Prof. Fabio Vistoli

Fabio Vistoli

a.a. 2022/2023

ABSTRACT
Machine learning is a component of artificial intelligence; it relies on computer algorithms and data analysis to learn patterns that exceeds the capacity of the human mind to comprehend. It uses statistical methods to infer relationships between predictors and outcomes in large datasets, and it has been successfully applied to predict adverse events in health care settings. In the preoperative phase, for risk stratification of patients with surgical conditions, various types of supervised learning have been used with large clinical databases. In our work, the aim was to investigate the potential role of machine learning (ML) versus classical statistical methods (SM) for the preoperative risk assessment in proctological surgery. We used clinical data from a nationwide audit: the database consisted of 1510 patients affected by Goligher's grade III hemorrhoidal disease who underwent elective surgery. We collected anthropometric, clinical, and surgical data and we considered ten predictors to evaluate model-predictive performance. The clinical target was the complication rate evaluated at 30-days follow-up. Logistic regression and three ML techniques were compared. ML models included a Decision Tree, a Support Vector Machine, and a classification Extreme Gradient Boosting (XGB). These methodologies could be used to develop a surgical risk calculator, which is already used and widespread for other diseases, that could help clinicians to estimate the chance of an unfavorable outcome after surgery.

INDEX

1. INTRODUCTION

1.1.HEMORRHOIDAL DISEASE

Definition

Hemorrhoidal disease (HD) is defined as the symptomatic displacement and venous distention of the normal vascular cushions (hemorrhoids) [1], which are formed by loose connective tissue, smooth muscle, arterial and venous vessels [2]. Hemorrhoids are generally located in three main positions: left lateral, right anterior, and right posterior portions. They lie beneath the epithelial lining of the anal canal and consist of direct arteriovenous communications, mainly between the terminal branches of the superior rectal and superior hemorrhoidal arteries, and between branches originating from the inferior and middle hemorrhoidal arteries and the surrounding connective tissue. The vascular cushions participate in the venous drainage of the anal canal, and their presence is essential for continence: they contribute approximately 15% to 20% of the resting anal pressure, so they intensify the action of the anal sphincter mechanism and shield the anal canal and the anal sphincter during the act of evacuation by filling with blood and providing extra padding. They congest during Valsalva manoeuvre or when intraabdominal pressure is increased, enabling the anal canal to remain closed [3].

Epidemiology

HD is the most common proctological disease. Its true prevalence is unknown; studies evaluating the epidemiology of HD showed that 10 million people in the United States reported it, for a prevalence of 4.4%, with a peak in individuals between 45 and 65 years of age in both sexes [4]. Some reports in the 21st century from South Korea and Austria yielded a prevalence in adult population of 14.4% [5] and 38.9% [6], respectively. It has been estimated that 25% of British people and 75% of American citizens will experience hemorrhoidal symptoms at some time in their lives [7,8]. People with HD have a tendency to use self-medication rather than to seek proper medical attention [9], and this makes the epidemiological survey a bit difficult. According to the Google's annual roundup in 2012, HD was the top trending heath issue in the United States, ahead of gastroesophageal reflux disease and sexually transmitted disease.

Pathophysiology and risk factors

The exact pathophysiology of HD is poorly understood [10]. As a result of destructive changes in the supporting connective tissue and abnormal blood circulation within anal cushions, the sliding anal

cushions embrace abnormal dilation and distortion of hemorrhoidal plexus. A recent study revealed a hyperperfusion state the plexus in patients with HD [11], suggesting the dysregulation of vascular tone. Moreover, it is evident that hemorrhoidal tissue contains some inflammatory cells [12] and newly-formed microvessels [13]. Therefore, although the true pathophysiology of HD development is unknown, it is likely to be multifactorial - including sliding anal cushion, hyperperfusion, vascular abnormality, tissue inflammation and internal rectal prolapse (rectal redundancy) [14,15]. The role of mucosal prolapse in HD is in debate: some surgeons consider it as a completely different pathology; others consider that mucosal prolapse is an integral part of the HD [16].
The risk factors that could contribute to the development of HD are also multiple and little known. During evacuation, voluntary sphincter contraction returns any residual fecal matter from the anal canal to the rectum as part of the normal physiology of evacuation. Straining to attain complete evacuation serves only to congest the vascular cushions. So straining, prolonged lavatory sitting, constipation, diarrhea, and conditions such as pregnancy, ascites, and pelvic space-occupying lesions that are associated with elevated intraabdominal pressure have been suspected to contribute to the development of the disease. A family history of hemorrhoidal disease has also been suggested to contribute, although there is no evidence of a hereditary predisposition [17-20].

Symptoms and diagnosis

The most common presentation of HD is painless rectal bleeding that occurs during or immediately after defecation, with or without prolapsing anal tissue. Usually, it is mild–moderate bright red bleeding which the patient observes on the feces or on the toilet paper [21-24]. Sometimes, HD may cause massive hemorrhage requiring urgent hospitalization and blood transfusions [25,26]. Other symptoms to consider are swelling, prolapse, soiling, perianal skin irritation, itching, and discomfort. Furthermore, large hemorrhoidal prolapse may cause sense of rectal filling and, rarely, difficult defecation. Pain is rare in case of uncomplicated HD. In fact, its presence may indicate other simultaneous painful conditions (fissure, abscess, pudendal neuropathy). Acute edema and thrombosis of external hemorrhoids are also responsible for acute anal pain [27,28].
Diagnosis of HD should start with the collection of medical history identifying symptoms suggestive of the disease and possible risk factors such as constipation, sedentary lifestyle, and pregnancy. The physical examination, including digital rectal examination and anoscopy, is imperative for the diagnosis. It should confirm the presence of HD ruling out other anorectal diseases. In any case, patients with rectal bleeding should be scheduled for colonoscopy [29-32]. Anorectal physiologic testing (such as manometry) and endorectal ultrasonography are important in evaluating patients with symptoms of soiling and incontinence [33].

Classification

Hemorrhoids are classified as internal or external based on the location from the dentate line. External hemorrhoids are located below, arising from the inferior hemorrhoidal plexus; they develop from ectoderm and are supplied by somatic nerves. Internal hemorrhoids lie above the dentate line, arising from the superior hemorrhoidal plexus; they develop from endoderm and are innervated by visceral nerve fibers [34].

The most widely used classification system for HD is the Goligher one, although concerns exist about the efficacy of this instrument to guide treatment. It ranks the presence and severity of prolapse into four grades: in first-degree the hemorrhoidal tissue protrudes into the lumen of the anal canal, but it does not prolapse outside the anal canal; second-degree hemorrhoids may prolapse beyond the external sphincter and be visible during evacuation but spontaneously return to lie within the anal canal; third-degree hemorrhoids protrude outside the anal canal and require manual reduction; fourth-degree hemorrhoids are irreducible and are constantly prolapsed [35]. Unfortunately, Goligher classification has several limitations, because it does not consider the associated symptoms and their impact on quality of life. To overcome these limitations, different grading systems have been developed [36-39], among which the one proposed in 2011 by Giordano et al [40-42].

Non-Surgical Treatments

It is important to document the grade of the hemorrhoids to determine appropriate treatment and to evaluate the efficacy of a particular treatment modality. In grade I and II HD conservative management usually can control symptoms. Conservative measures include dietary modifications with hydration and avoidance of straining. Symptomatic control using hot Sitz bath and topical treatments containing local anaesthetics, corticosteroids, or anti-inflammatory drugs are useful. Phlebotonics showed a significant effect on HD related symptoms if compared with control group [43-47]: flavonoids are the most common agents. The use of laxatives could be considered for symptom relief and to reduce bleeding.

Many office-based procedures (such as rubber band ligation, injection sclerotherapy, infrared coagulation, cryotherapy, radiofrequency ablation and laser therapy) are effectively performed for grade I - II HD and for some cases of grade III HD with or without local anaesthesia [48,49]. Choice of the outpatient procedure should be informed by shared decision-making, taking into account patient preferences, availability of procedures and fitness for further procedures.

Surgical Treatments

As stated in the revised practice parameters for the management of hemorrhoids [50,51], surgical treatment should be offered to patients in whom office procedures were unsuccessful, patients who are not capable of tolerating office procedures, patients with grades III to IV mixed (internal-external) HD, or when complications occurred [10]. An ideal operation should remove internal and external component of hemorrhoids, have minimal postoperative pain and complications, demonstrate less recurrence, and should be easy to learn and perform. Unfortunately, none of the currently available operations achieves all the ideal conditions.

Surgical treatment options for HD are grouped into non-excisional and excisional methods. The surgical excisional operations include various techniques, with or without closure of the anorectal mucosa or the anoderm [52,53]. The newer non-excisional technique includes hemorrhoidal artery ligation (HAL) and plication of hemorrhoids (or known as ligation anopexy or mucopexy) [54,55].

Treatment of Goligher's grade III HD is still debated: traditional excisional hemorrhoidectomy (EH) it is still associated with significant postoperative pain and with a high rate of complications [56,57]; HAL seems to be associated with decreased postoperative pain and faster recovery, but also with a higher recurrence rate [58]. Besides the type of intervention, many other pre- and intra-operative factors have been studied as possible predictors of adverse events, but to date no studies on large cohorts of patients are available.

1.2.MACHINE LEARNING

Advances in processing power and cloud storage have given surgeons access to increased amounts and types of data. These types of "big data" have facilitated the utilization of artificial intelligence (AI). Machine learning (ML) is a component of AI; it relies on computer algorithms and data analysis to learn patterns that exceeds the capacity of the human mind to comprehend [59]. It uses statistical methods to infer relationships between predictors and outcomes in large datasets and have been successfully applied to predict adverse events in health care settings [60-65].

ML algorithms are generally classified into (Figure 1):

a. supervised learning
b. unsupervised learning
c. reinforcement learning.

Supervised learning aims to solve diagnostic and prediction problems. "Supervised" refers to the existence of a training set: in supervised learning, ML approaches automatically learn the relationships between predictors from the data, enabling the development of a more flexible model than the conventional logistic regression model. In these algorithms, part of the data set (the “training” data set) is analyzed to build a model and another part of the data set (the “testing” data set) is used to validate the model. Supervised learning algorithms generally include lasso/ridge regression, support vector machine (SVM), decision tree, random forest, gradient boosting, and artificial/deep neural networks [66].

Unsupervised learning aims to explore latent subgroups in a particular condition, allowing the model to find undetected features from the data set; that is, having “no answer”. Unsupervised learning has the potential to classify heterogeneous diseases into homogeneous groups. While many unsupervised learning algorisms have been suggested, k-means and partitioning around Medoids are the major methods used in clinical research.

Reinforcement learning aims to learn the best strategies. Reinforcement learning solves problems through the simulation of trial-and-error experiments to gain the best outcomes. Reinforcement learning is trained using a set of states (environment), a set of strategies (actions) and outcomes (reward) [67]. A Markov decision process is one of the major tools employed in reinforcement learning.

In clinical settings, for risk stratification of patients, various types of supervised learning have been used with large clinical databases.

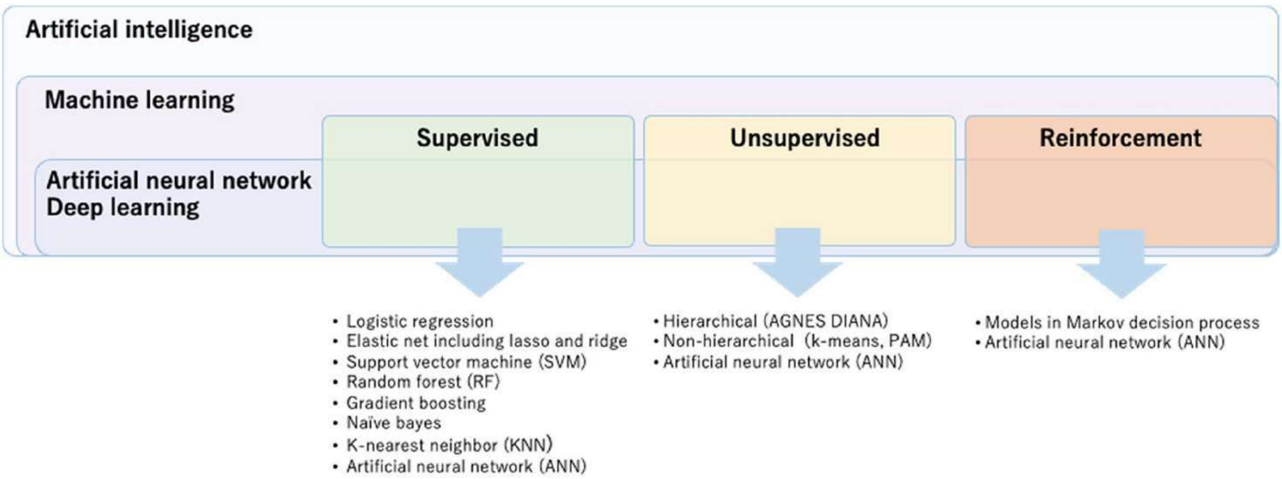


Figure 1. Hierarchy of artificial intelligence and machine learning algorithms [65]

2. RESEARCH AIMS AND RATIONALE

Clinical risk prediction models are ubiquitous in many medical domains. Model development studies aim to derive a prediction model by selecting predictors and combining them into a multivariable model. The traditional approach to develop these models involves the use of regression models, for example, logistic regression (LR) to predict disease presence (diagnosis) or disease outcomes (prognosis). These models are relatively easy to use and interpret but have some drawbacks: first, the usual assumption for logistic regression is that there is a linear relation between the independent and dependent variables; second, predictors are usually chosen using backward selection.

Machine learning (ML) algorithms are gaining in popularity as an alternative approach for prediction and classification problems, as they are less prone to the above-mentioned problems. ML algorithms can detect non-linear relationships between independent and dependent variables and incorporate many of them. Moreover, they do not require data to conform to statistical assumptions, such as independence of observations and the avoidance of multicollinearity of independent variables. These models are, however, more susceptible to overfitting (too many predictors and too much complexity relative to few outcome events) and have the so-called black-box phenomenon (high accuracy, but low transparency and interpretability for humans) [68,69].

A useful definition of ML is that it focuses on models that directly and automatically learn from data [70]. By contrast, regression models are based on theory and assumptions, and benefit from human intervention and subject knowledge for model specification [71,72]. Although both techniques have been used to develop risk models for postoperative complications, it is unclear if machine learning is superior to logistic regression when using structured data.

This study used clinical data from a nationwide audit to compare machine learning to logistic regression in predicting complications after planned surgery for hemorrhoidal disease. We investigated the application of three ML algorithms (Decision Tree, Support Vector Machine, XGBoost) to provide a comprehensive evaluation of the ML models for the problem at hand. The predictive value of different models was compared using AUC (Area Under the Curve) balanced accuracy, sensitivity, and specificity.

3. METHODS

Source of data

We performed a multicenter retrospective observational study including patients affected by Goligher's grade III hemorrhoidal disease (HD) who underwent hemorrhoidal artery ligation (HAL) with mucopexy or excisional hemorrhoidectomy (EH) between January 2016 and February 2020. Any centre belonging to the Italian Society of Colorectal Surgery (SICCR) in which at least 30 surgical procedures per year for hemorrhoidal disease were performed was able to join the study. Given the effects of pandemic SARS-CoV-2 on surgical and outpatient activities, data regarding surgery or follow-up after March 2020 have not been included in the analyses.

The study was conducted in accordance with the Declaration of Helsinki (1996) and International Conference on Harmonization-Good Clinical Practice (ICHGCP) guidelines, and it obtained approval from local ethics committee (Prot. n. 51380, 22.04.2021). It was registered on ClinicalTrials.gov (Identifier: NCT04863963). Patients selected for the study gave informed consent to participate.

Participants

We included patients aged 18 years or older, with Goligher's grade III HD, who underwent elective conventional excisional hemorrhoidectomy (EH) or transanal dearterialization (HAL) with or without use of the Doppler transducer and with mucopexy, for whom a 30-day follow-up was available. Exclusion criteria were as follows: recurrent disease; presence of Crohn's disease or ulcerative colitis; coagulopathies. We also excluded cases in which combined surgical procedures were performed and cases in which the procedure did not correspond to those described below. Diagnosis of primary HD was established by clinical examination and anoscopy or proctoscopy.

Surgical techniques

Excisional hemorrhoidectomy: it was performed using an Eisenhammer retractor according to the Milligan–Morgan technique with the patient in the lithotomy position. The standardized 3-quadrant open hemorrhoidectomy was performed. It consisted of a skin incision on the mucocutaneous border, retraction of the pile mass, dissection and excision of the hemorrhoids to the anorectal junction. The wounds were not closed [73].

Transanal Hemorrhoidal Artery Ligation with Doppler (DG-HAL): the surgical procedure described by Ratto et al. [74] was followed. Patients were placed in a lithotomy position. The proctoscope was inserted through the anal canal reaching the lower rectum. Tilting the proctoscope, the best Doppler

signals were sought corresponding to all 6 main trunks of the hemorrhoidal arteries. Hemorrhoidal artery ligation was performed with absorbable suture. Mucopexy was also performed in all patients.
Transanal Hemorrhoidal Artery Ligation without Doppler guidance: patients were placed in a lithotomy position. The proctoscope was inserted through the anal canal reaching the lower rectum. Arterial ligation was performed at standard zones, identified by intraoperative palpation. Mucopexy was then performed.

Data collection

Information was stored in an online database. Collected data included patients' demographics and clinical information; operative details such as procedure and anaesthesia; post-operative complications. Patients were asked to report their symptoms in the preoperative period, using a questionnaire [40] (Figure 2) specifically designed to assess the severity of hemorrhoidal symptoms using five different parameters (bleeding, prolapse, manual reduction, discomfort/pain, impact on QoL), and based on a five points scale from 0 (never) to 4 (with every bowel movement). An overall score of 0 corresponds to a total absence of symptoms, while an overall score of 20 indicates the worst possible symptomatology.
Complications were evaluated according to the Clavien-Dindo classification [75]:

- Grade 0 means no complications;
- Grade I means any deviation from the normal postoperative course without the need for pharmacological treatment or surgical, endoscopic, and radiological interventions. Allowed therapeutic regimens are: antiemetics, antipyretics, analgetics, diuretics, electrolytes, and physiotherapy;
- Grade II means any event requiring pharmacological treatment with drugs other than such allowed for grade I;
- Grade III means any event requiring surgical, endoscopic or radiological intervention;
- Grade IV means life-threatening complication requiring IC (intensive care) management;
- Grade V means death of a patient.

	Never	At least once per year	At least once per months	At least once per week	With every bowel movement
Bleeding	0	1	2	3	4
Prolapse	0	1	2	3	4
Manual reduction	0	1	2	3	4
Discomfort/pain	0	1	2	3	4
Impact on QoL	Not at all 0	Minimal 1	Moderate 2	Severe 3	Very severe 4

Figure 2. Symptom questionnaire [40]

3.1.STATISTICAL ANALYSIS

The dataset contained ten prognostic factors. Eight were qualitative: sex (female or male), presence of rectal mucosal prolapse (yes or no), preoperative prolapse (yes or no), preoperative bleeding (yes or no), preoperative manual reduction (yes or no), preoperative pain/discomfort (yes or no), surgical treatment performed (EH or HAL), type of anaesthesia (general, spinal or local). Two were quantitative: age (in years) and preoperative score (from 0 to 24).

Qualitative data were expressed as numerical values and percentages, and comparison between groups was performed using the χ^2 test or Fisher's exact test, as appropriate. Quantitative variables were described as mean and standard deviation (SD) and median and range, and the comparison between groups was conducted using Two-sample Wilcoxon rank-sum (Mann-Whitney) test. Statistical significance was set at $p<0.05$.

Dataset was firstly randomly splitted into a "train" group (80% of the whole dataset), for the development and validation of the model, and a "test" group (20% of the whole dataset) used for making predictions. This data split was fixed and used for all models.

For evaluating model performance, we considered:

- true positives (TP)
- true negatives (TN)
- test accuracy = (TP+TN)/all patients
- sensitivity (true positive rate, TPR) = TP / (TP+FN)
- specificity (true negative rate, TNR) = TN/(TN+FP)
- balanced accuracy = average between the sensitivity and the specificity
- precision (positive predictive value, PPV) = TP/(TP+FP)
- negative predictive value (NPV) = TN/(TN+FN)

- the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC).

A graphical visualization of the performance of the models was represented by confusion matrices.

Logistic regression model

Analyses were conducted using R software (version 4.1.2). The associations between the predictors and outcomes were displayed as Odds Ratios (ORs), 95% Confidence Intervals (Cis), and *p*-values. Backward stepwise selection with the Akaike and Baesyan information criterion (AIC and BIC) was used to choose the best model.

The logistic function for p was given by:

$$p = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}$$

Here, "p" denoted the probability for a patient that a post-operative complication occurs; "$x_1,\ldots x_n$" were the included variables; "e" was the base of the natural logarithm (2.718); "β_0" was the intercept; "$\beta_1,\ldots\beta_n$" were the coefficients for variables "$x_1,\ldots x_n$".

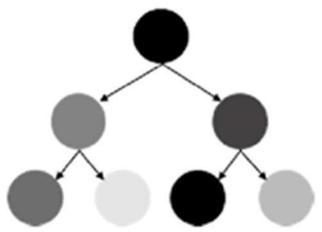
The parameters characterizing the predictive performance of the test based on the logistic discriminant function, as well as the corresponding standard errors, were estimated by bootstrapping with 100 replications. The optimal cut-off value of *p* was calculated using a receiver operating characteristic (ROC) curve analysis and the Youden Index.

Machine learning techniques

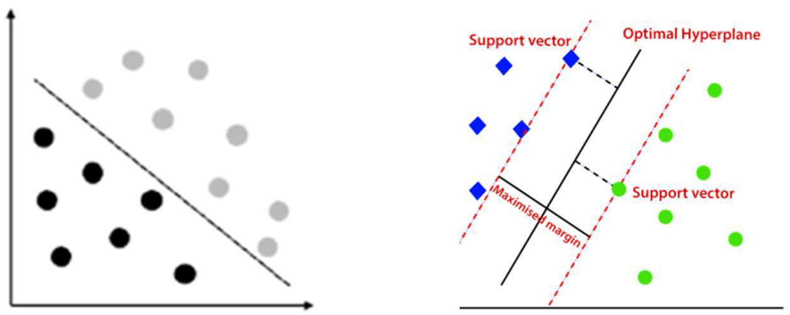
Decision tree (DT), Support vector machine (SVM), and Extreme Gradient Boosting (XGB) algorithms [76-79] were used to assess the predictive weight of prognostic factors on the occurrence of complications. DT and SVM were implemented using the Python sklearn package, while XGB was taken from https://xgboost.readthedocs.io/en/stable. All the experiments have been carried out on a laptop Intel(R) Core(TM) i7-6700HQ CPU @ 2.60G with 8 GB of RAM.

- Decision trees (DTs) create a series of decision rules based on continuous and/or categorical input variables to predict an outcome. To derive a decision tree, the algorithm applies a splitting rule on successively smaller partitions of data, with each partition being a node on the tree. DTs are generally easy to understand, making their output ideal for a range of target audiences. They are also flexible to non-linear covariate effects and can incorporate higher-

order interactions between covariates. Trees may lose information by dichotomizing or categorizing variables where associations are continuous, and they can be unstable to even small data changes. Moreover, they are prone to overfitting, and their ultimate utility depends heavily on appropriately implemented stopping criteria.
DTs provide feature importance as part of their output. This importance is based on how much each feature reduces the impurity in the tree.

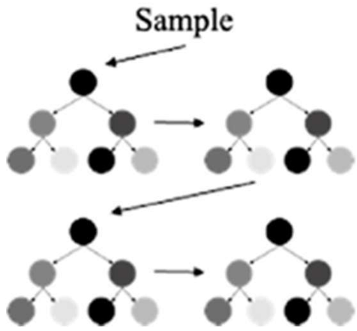


- Support vector machines (SVMs) are a popular Machine Learning method originally designed for binary classification tasks. They construct a hyperplane, which is the optimal boundary, that separates observations according to their class membership. Such hyperplane divides the data space into two half-spaces so that new observations can be classified based on which of the two half-spaces they lie in. Does not work well for classifying more than 2 phases. Requires features to be manually defined for use as inputs, causing potential loss of important features.
SVMs generally demonstrate low misclassification error and scale well to high-dimensional data. They do not provide feature importance as part of their output.



- Ensemble methods utilize information from multiple models to improve predictive performance compared to a single model. The idea is that even though any individual model within an ensemble is not adequate to capture the characteristics of the entire phenomenon, so long as they perform better than at random, once combined they can borrow strength from each other and achieve high predictive accuracy.

  Extreme Gradient Boosting (XGBoost, XGB) is a classification ensemble method. that is trained by building many decision trees. The final prediction is determined by combining the predictions of the single DTs by different mechanism like, for example, the majority voting. This is done to improve the performance of a single DT, by increasing the generalization capability of the model. GB is designed to iteratively add new DTs to the whole model so as to reduce the errors made by already inserted trees. XGB has successfully addressed many ML challenges; for this reason, it has recently been applied to solving many practical problems.



4. RESULTS

Twenty-eight centers belonging to the Italian Society of Colorectal Surgery (SICCR) joined the study [80]. The mean number of patients per center was 60. The database consisted of 1731 total patients who underwent elective surgery between January 2016 and February 2020, of which 1681 met the defined criteria. The dataset contained 171 (10%) missing data overall for the variables, with 1510 complete cases (90%) that were included in the analysis. The demographic and clinical characteristics of the whole sample are shown in Table 1. The most adopted technique was excisional hemorrhoidectomy (1128/1510 patients, 74.7%). The mean age of all patients (59.7% males and 40.3% females) was 53 years (SD: 13.0).

Based on the Clavien-Dindo classification, we considered grades 0 and I as “no complications” and grades II, III, IV, and V as “presence of complications”. According to this definition, ten per cent (10%) of patients reported complications (148 patients out of 1510). Characteristics of the sample stratified by complications are reported in Table 2.

Table 1. Demographic and clinical characteristics of the whole sample

VARIABLES	**PATIENTS (n=1510)**
AGE, years	
mean (SD)	53 (13.0)
median (range)	52 (20-90)
SEX, n (%)	
male	901 (59.7)
female	609 (40.3)
MUCOSAL PROLAPSE, n (%)	
yes	1195 (79.1)
no	315 (20.9)
PRE-OP. PROLAPSE, n (%)	
yes	1258 (16.7)
no	252 (83.3)
PRE-OP. BLEEDING, n (%)	
yes	1358 (89.9)
no	152 (10.1)
PRE-OP. MANUAL REDUCTION, n (%)	
yes	977 (64.7)
no	533 (35.3)
PRE-OP. PAIN/DISCOMFORT, n (%)	
yes	1401 (92.8)
no	109 (7.2)
PRE-OP. SCORE	
mean (SD)	10.8 (4.5)
median (range)	11 (0-24)
TREATMENT, n (%)	
EH*	1128 (74.7)
HAL*	382 (25.3)
ANAESTHESIA, n (%)	
general	120 (7.9)
spinal	1077 (71.3)
local	313 (20.7)
COMPLICATIONS, n (%)	
yes	148 (9.8)
no	1362 (90.2)

*EH= Excisional Hemorrhoidectomy; HAL= Hemorrhoidal Artery Ligation

Table 2. Demographic and clinical characteristics of the sample stratified by complications yes/no

VARIABLES	NO COMPLICATIONS (n=1362)	COMPLICATIONS (n=148)	*p*-value
AGE, years			**0.006****
mean (SD)	53 (13.1)	50 (12.2)	
median (range)	52 (20-90)	49 (24-81)	
mean (SD)	53 (13.1)	50 (12.2)	
median (range)	52 (20-90)	49 (24-81)	
SEX, n (%)			**<0.001***
male	838 (61.5)	63 (42.6)	
female	524 (38.5)	85 (57.4)	
MUCOSAL PROLAPSE, n (%)			**<0.001***
yes	1056 (77.5)	139 (93.9)	
no	306 (22.5)	9 (6.1)	
PRE-OP. PROLAPSE, n (%)			**0.024***
yes	1119 (82.2)	139 (93.9)	
no	243 (17.8)	9 (6.1)	
PRE-OP. BLEEDING, n (%)			**<0.001***
yes	1212 (89.0)	146 (98.6)	
no	150 (11.0)	2 (1.4)	
PRE-OP. MANUAL REDUCTION, n (%)			**<0.001***
yes	841 (61.7)	136 (91.9)	
no	521 (38.3)	12 (8.1)	
PRE-OP. PAIN/DISCOMFORT, n (%)			**0.004***
yes	1255 (92.1)	146 (98.6)	
no	107 (7.9)	2 (1.4)	
PRE-OP. SCORE			**<0.001***
mean (SD)	10.4 (4.4)	14.5 (3.8)	
median (range)	11 (0-24)	15 (2-20)	
TREATMENT, n (%)			**<0.001***
EH*	996 (73.1)	132 (89.2)	
HAL*	366 (26.9)	16 (10.8)	
ANAESTHESIA, n (%)			**<0.001***
general	116 (8.5)	4 (2.7)	
spinal	1048 (76.9)	29 (19.6)	
local	198 (14.5)	115 (77.7)	

*EH= Excisional Hemorrhoidectomy; HAL= Hemorrhoidal Artery Ligation
*χ2 test or Fisher's exact test
**Two-sample Wilcoxon rank-sum (Mann-Whitney) test

We can see from Table 2 that all the collected variables were statistically significant among groups “complications” and “no-complications” at the univariate analysis, so they could be all included in the regression analyses. However, based on the structure of the administered questionnaire, we reported the occurrence of collinearity among the variable “preoperative score” and the variables “preoperative prolapse”, “preoperative bleeding”, “preoperative manual reduction”, and “preoperative pain/discomfort”. Backward stepwise selection showed that the choice of variable “preoperative score” reduced both parameters AIC and BIC, so we decided for a model that included this variable and excluded the other four.
Figure 3 shows the dataset split used for all the models analysed below, and the characteristics of the sample stratified by group (training/testing) are reported in Table 3.
It is common to divide the training set into a part dedicated to training the algorithm, properly called the “training” set, and a part dedicated to verifying the goodness of the training, called the “validation” set.

Figure 3. Dataset split (train, validation, and test groups)

Table 3. Demographic and clinical characteristics of the sample stratified by group (train or test)

VARIABLES	TRAIN (n=1208)	TEST (n=302)	*p*-value
AGE, years			0.304**
mean (SD)	53 (13.1)	52 (12.7)	
median (range)	52 (20-90)	52 (27-88)	
SEX, n (%)			0.416*
male	727 (60.2)	174 (57.6)	
female	481 (39.8)	128 (42.4)	
MUCOSAL PROLAPSE, n (%)			0.874*
yes	955 (79.1)	240 (79.5)	
no	253 (20.9)	62 (20.5)	
PRE-OP. SCORE			0.526**
mean (SD)	10.8 (4.4)	11.0 (4.7)	
median (range)	11.0 (1-24)	11.0 (0-20)	
TREATMENT, n (%)			0.723*
EH*	900 (74.5)	228 (75.5)	
HAL*	308 (25.5)	74 (24.5)	
ANAESTHESIA, n (%)			0.656*
general	90 (7.5)	30 (9.9)	
spinal	870 (72.0)	207 (68.5)	
local	248 (20.5)	65 (21.5)	
COMPLICATIONS, n (%)			0.931*
yes	118 (9.8)	30 (9.9)	
no	1090 (90.2)	272 (90.1)	

*EH= Excisional Hemorrhoidectomy; HAL= Hemorrhoidal Artery Ligation
*$\chi 2$ test or Fisher's exact test
**Two-sample Wilcoxon rank-sum (Mann-Whitney) test

In this case, differences among variables were not statistically significant between the two groups, which resulted therefore to be comparable.
Logistic regression analysis performed on the "train" group is showed in Table 4.

Table 4. Multivariate logistic regression analysis of predictors for complications

			95% Confidence Interval		
Predictor	**Estimate**	**Odds ratio**	**Lower**	**Upper**	***p*-value**
Intercept	-6.86	0.001	2.42e-4	0.005	< .001
AGE	-0.03	0.975	0.9	0.993	0.006
SEX	0.33	1.397	0.892	2.187	0.144
MUCOSAL PROLAPSE	1.33	3.768	1.609	8.820	0.002
TREATMENT	-0.55	0.579	0.311	1.076	0.084
ANAESTHESIA	2.31	10.103	6.326	16.133	< .001
Spinal anaesthesia	-0.30	0.738	-1.555	0.948	0.635
Local anaesthesia	2.27	9.710	1.039	3.507	< .001
PREOPERATIVE SCORE	0.12	1.128	1.056	1.204	< .001

The model elaborated by the logistic regression was as follow:

$$\mathrm{p} = \frac{e^{(-6.86-0.03\ \text{age}\ +\ 0.33\ \text{sex}+\ 1.33\ \text{mucosal prolapse}\ -\ 0.55\ \text{treatment}\ +2.31\ \text{anaesthesia}\ +\ 0.12\ \text{preoperative score})}}{1+e^{(-6.86-0.03\ \text{age}\ +\ 0.33\ \text{sex}+\ 1.33\ \text{mucosal prolapse}\ -\ 0.55\ \text{treatment}\ +2.31\ \text{anaesthesia}\ +\ 0.12\ \text{preoperative score})}}$$

considering "age" and "preoperative score" as quantitative variables ranging from 20 to 90 and from 0 to 24, respectively; "sex", "mucosal prolapse", and "treatment" as categorical variables coded 0 or 1; "anaesthesia" as categorical variable coded 0 (general anaesthesia), 1 (spinal anaesthesia), or 2 (local anaesthesia).
The AUC was 0.83 (Figure 3) and the optimal cut-off value of p defined by the Youden Index was 0.24. This model was then performed on the "test" database (302 patients, of which 30 reported complications), with the performance metrics reported in Table 5.

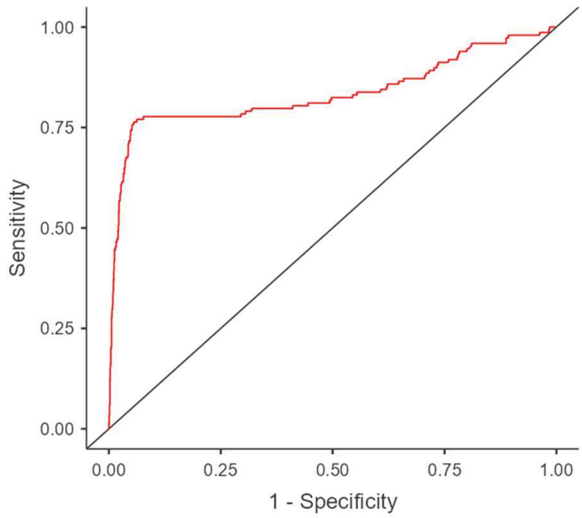


Figure 3. ROC curve of logistic regression model

Decision Tree (DT), Support Vector Machine (SVM), and XGBoost (XGB) algorithms were then performed. Confusion matrices of each model is reported in Figure 4, and discrimination performance of different models was finally compared (Table 5).

Logistic Regression

		Predicted 0	**Predicted** 1
Observed	0	250	22
	1	5	25

Decision Tree

		Predicted 0	**Predicted** 1
Observed	0	241	31
	1	6	24

Support Vector Machine

		Predicted 0	**Predicted** 1
Observed	0	222	50
	1	4	26

XGBoost

		Predicted 0	**Predicted** 1
Observed	0	254	18
	1	5	25

Figure 4. Confusion matrices of the included models

Table 5. Comparison among model performances

	LOGISTIC REGRESSION MODEL*	DECISION TREE	SUPPORT VECTOR MACHINE	XGBOOST
P	30	30	30	30
N	272	272	272	272
TP	25	24	26	25
TN	250	241	222	254
Test Accuracy	91%	88%	82%	92%
Balanced Test Accuracy	88%	84%	84%	88%
ROC AUC score	0.83	0.84	0.84	0.88
TPR (sensitivity)	83%	80%	87%	80%
TNR (specificity)	92%	89%	82%	93%
PPV (precision)	53%	44%	34%	58%
NPV	98%	98%	98%	98%

P=positive; N=negative; TP=true positives; TN= true negatives; TPR=true positives rate; TNR= true negatives rate; PPV= positive predictive value; NPV=negative predictive value
*Cut-off value of p= 0.24

The AUC of the multivariate logistic regression model was 0.83, and the balanced test accuracy was 88%, with sensitivity and specificity of 0.83 and 0.92, respectively.
The AUC of the DT model was 0.84, and the balanced test accuracy was 84%, with sensitivity and specificity of 0.80 and 0.89, respectively.
The AUC of the SVM model was 0.84, and the balanced test accuracy was 84%, with sensitivity and specificity of 0.87 and 0.82, respectively.
The AUC of the XGBoost model was 0.88, and the balanced test accuracy was 88%, with sensitivity and specificity of 0.80 and 0.93, respectively.

The models also reported the input features importance to establish the importance of each feature in the competing risk assessment, except for the SVM algorithm which could not provide this information.
For logit analysis, feature importance was reported in terms of Odds Ratio (OR) and *p*-value for significance; for ML models it was reported in terms of relative importance (the higher the score for a feature, the larger effect it has on the model to predict a certain variable).

According to logistic regression model, the three most important and significant variables that led to increased surgical risk were:

- anaesthesia (OR 10.1, $p<0.001$)
- mucosal prolapse (OR 3.8, $p=0.002$)
- preoperative score (OR 1.1, $p<0.001$)

According to DT, the three most important variables that led to increased surgical risk were:

- anaesthesia (42.74)
- preoperative score (27.59)
- treatment (10.22)

According to XGB, the three most important variables that led to increased surgical risk were:

- anaesthesia (35.87)
- treatment (12.47)
- mucosal prolapse (9.02)

5. DISCUSSION

To the best of our knowledge, this is the first research comparing classical statistical methods (SM) with ML techniques for predicting complications after surgery for hemorrhoidal disease. Among ML techniques we decided to use DT and SVM because they are the most employed models for clinical risk prediction [70], and we also performed XGB because it is an ensemble method with improved predictive performance than single ones.

Our study found no superiority of the ML using structured pre- and intra-operative variables. In detail, the results showed that among the ML techniques, XGB was the most complex and accurate; however, it was overlapping with SM in terms of balanced accuracy, specificity, and negative predictive value (NPV). The AUC and precision were slightly better in XGB than SM, but on the other hand SM had a higher sensitivity, which means that it has the ability to better predict high risk patients. In the others ML models (DT and SVM) the performance metrics were on average weaker with respect to both SM and XGB. These results are consistent with other studies and reviews [69,70,72,81], concluding that ML do not seem to outperform logistic regression model in predicting postoperative complications.

Regarding the relative importance of the input features, all models agreed in identifying the most important factor, which resulted to be anaesthesia. In particular, its OR meant that local anaesthesia was about 10-times more at risk than general anaesthesia ($p< .001$), while spinal anaesthesia was associated with a lower risk, although this last data was not statistically significant (p=0.635). Among the other variables, regression model identified as significant risk factors the presence of mucosal prolapse, in agreement with XGB model, and a high preoperative score, in agreement with DT. Regarding the treatment, only ML models identified it among the three most important features for preoperative competing risk. According to logistic regression model, the treatment had an OR of 0.579, suggesting that HAL has a lower risk of complications than EH, but this data was not statistically significant (p=0.084).

Clinical risk prediction models are ubiquitous in many medical domains, in which risk calculators are used and widespread to estimate the chance of an unfavorable outcome after surgery. ML potentially provides a powerful tool for this purpose, because it has the ability to demonstrate correlations that may be missed by traditional methods; however, there are some risks to utilizing it incorrectly.

In the future, a surgeon will likely see ML analysis of population data augmenting each phase of care; automated analysis of all preoperative data could provide a more patient-specific risk score and valuable predictors for postoperative care. Surgeons could also improve their decision-making intraoperatively, based on real-time analysis that integrates operative video, vital signs, and instrument/hand tracking (unstructured data), and integration of pre-, intra-, and post-operative data could help to monitor recovery and predict complications [82].

In our case we still did not find evidence of superior performance of ML over SM. There are several reasons why machine learning did not outperform logistic regression. A first limitation could be that most of the variables in this study were discrete values, although the potency of machine learning models lies in the analysis of the above-mentioned unstructured data (text, video, images,…). A second limitation could be that although the etiology and the potential risk factors are partly understood, there might still be unknown and unrecognized factors (both in the pre and perioperative phase) playing roles of importance. Moreover, the number of variables and patients in this database may have been too small for machine learning to show its benefit.

In fact, to successfully train and use ML models, the dataset on which they are trained needs to be sufficiently large [83]. Medical settings are often characterized by low-medium sample size and limited number of predictors, and in these instances application of ML algorithms should only be motivated for exploration of the collected data. In these settings, a classical regression approach may still have good prognostic performance in predicting risk, although with some drawbacks: for example, the usual assumption for logistic regression is that there is a linear relation between the independent and dependent variables; second, predictors are usually chosen using backward selection. ML algorithms are less prone to these problems, but they are also more susceptible to overfitting and to the black-box phenomenon [68,69]. Black-box phenomenon refers to systems with internal workings that are invisible to the user: you can give them input and get output, but you cannot examine the system's code or the logic that produced the output. In fact, ML techniques are typically opaque [84]: they provide algorithms for prediction or recommendation but do not explain or justify those results, raising challenges in validation, regulation, and integration into practice.

The question of algorithmic interpretability is the subject of an expansive literature; it is necessary to report prediction models powered by artificial intelligence completely and transparently to allow critical appraisal, reproducibility of the modelling steps and results by a wider audience [81]. By contrast, regression models are relatively easy to use and interpret, are based on theory and assumptions, and benefit from human intervention and subject knowledge for model specification [71,72].

Another current drawback of ML is that it cannot yet determine causal relationships in data at a level necessary for clinical implementation, nor can it provide an automated clinical interpretation of its analyses. While big data can be rich with variables, it is poor in providing the appropriate clinical context with which to interpret them.

As we can see from our results, for example, ML models understood and underlined that anaesthesia and treatment were important variables for risk prediction, but to the best of our knowledge, they do not specify which treatment among two or which anaesthesiologic technique among three is most at risk. This information is instead provided by classic logistic regression through the Odds Ratio values. Human physicians, therefore, must critically evaluate the predictions generated by ML and interpret the data in clinically meaningful ways.

6. CONCLUSION

In this research, we discussed ML models (DT, SVM, XBG) alternative to SM to predict complications after planned surgery for hemorrhoidal disease. We used a medium sample size, a limited number of predictors (simple setting), and mostly discrete variables. Methods were compared in terms of AUC, balanced accuracy, sensitivity, and specificity. In this setting, ML and SM models reached an equivalent predictive performance, but the conventional regression model had the advantage to be transparent and easily interpretable. In our opinion, for non-complex real-life data such as these, ML techniques should only be employed complementary to SM as exploratory tools of model's performance: they potentially provide a powerful tool to demonstrate correlations that may be missed by traditional methods, but there are some drawbacks to using them uncritically.

REFERENCES

1. Giuliani A, Romano L, Lazzarin G, et al. Relationship between hemorrhoidal grade and toilet habits. Ann Ital Chir. 2020;91:192-195
2. Thomson WH. The nature and cause of hemorrhoids. Proc R Soc Med. 1975 Sep;68(9):574-5
3. Kaidar-Person O, Person B, Wexner SD. Hemorrhoidal disease: A comprehensive review. J Am Coll Surg. 2007 Jan;204(1):102-17
4. Gallo G, Sacco R, Sammarco G (2018) Epidemiology of hemorrhoidal disease. In: Ratto C, Parello A, Litta F (eds) Hemorrhoids Coloproctology, vol 2. Springer, Cham, pp 3–7
5. Lee JH, Kim HE, Kang JH, et al. Factors associated with hemorrhoids in korean adults: korean national health and nutrition examination survey. Korean J Fam Med. 2014 Sep;35(5):227-36
6. Riss S, Weiser FA, Schwameis K, et al. The prevalence of hemorrhoids in adults. Int J Colorectal Dis. 2012 Feb;27(2):215-20
7. Tucker H, George E, Barnett D, et al. NICE Technology Appraisal on Stapled Hemorrhoidopexy for the Treatment of Hemorrhoids. Ann R Coll Surg Engl. 2008 Jan;90(1):82–4
8. Baker H. Hemorrhoids. In: Longe JL, editor Gale Encyclopedia of Medicine. 3rd ed. Detriot: Thomson Gale, 2006: 1766-1769
9. Rohde H, Christ H. [Hemorrhoids are too often assumed and treated. Survey of 548 patients with anal discomfort]. Dtsch Med Wochenschr. 2004 Sep 17;129(38):1965-9. German
10. Lohsiriwat V. Hemorrhoids: from basic pathophysiology to clinical management. World J Gastroenterol. 2012 May 7;18(17):2009-17
11. Aigner F, Gruber H, Conrad F, et al. Revised morphology and hemodynamics of the anorectal vascular plexus: impact on the course of hemorrhoidal disease. Int J Colorectal Dis. 2009 Jan;24(1):105-13
12. Morgado PJ, Suárez JA, Gómez LG, et al. Histoclinical basis for a new classification of hemorrhoidal disease. Dis Colon Rectum. 1988 Jun;31(6):474-80
13. Chung YC, Hou YC, Pan AC. Endoglin (CD105) expression in the development of hemorrhoids. Eur J Clin Invest. 2004 Feb;34(2):107-12
14. Lohsiriwat V. Treatment of hemorrhoids: A coloproctologist's view. World J Gastroenterol. 2015 Aug 21;21(31):9245-52
15. Deutsch AA, Moshkovitz M, Nudelman I, et al. Anal pressure measurements in the study of hemorrhoid etiology and their relation to treatment. Dis Colon Rectum. 1987;30:855–857
16. Gaj F, Trecca A. Hemorrhoids and rectal internal mucosal prolapse: one or two conditions? A national survey. Tech Coloproctol. 2005;9:163–165
17. Loder PB, Kamm MA, Nicholls RJ, et al. Hemorrhoids: pathology, pathophysiology and aetiology. Br J Surg. 1994;81:946–954
18. Wald A. Constipation, diarrhea, and symptomatic hemorrhoids during pregnancy. Gastroenterol Clin North Am. 2003;32:309–322

19. Johannsson HO, Graf W, Pahlman L. Bowel habits in hemorrhoid patients and normal subjects. Am J Gastroenterol. 2005;100:401–406
20. Ray-Offor E, Amadi S. Hemorrhoidal disease: Predilection sites, pattern of presentation, and treatment. Ann Afr Med. 2019 Jan-Mar;18(1):12-16
21. Idrees JJ, Clapp M, Brady JT, et al. Evaluating the Accuracy of Hemorrhoids: Comparison Among Specialties and Symptoms. Dis Colon Rectum. 2019 Jul;62(7):867-871
22. Sengupta N, Tapper EB, Feuerstein JD. Early Versus Delayed Colonoscopy in Hospitalized Patients With Lower Gastrointestinal Bleeding: A Meta-Analysis. J Clin Gastroenterol. 2017 Apr;51(4):352-359
23. Bretthauer M, Kaminski MF, Løberg M, et al. Population-Based Colonoscopy Screening for Colorectal Cancer: A Randomized Clinical Trial. JAMA Intern Med. 2016 Jul 1;176(7):894-902
24. Mehanna D, Platell C. Investigating chronic, bright red, rectal bleeding. ANZ J Surg. 2002 Dec;72(12):923-4
25. Gralnek IM, Neeman Z, Strate LL. Acute Lower Gastrointestinal Bleeding. N Engl J Med. 2017 Mar 16;376(11):1054-1063
26. Aoki T, Hirata Y, Yamada A, et al. Initial management for acute lower gastrointestinal bleeding. World J Gastroenterol. 2019 Jan 7;25(1):69-84
27. Gallo G, Martellucci J, Sturiale A, et al. Consensus statement of the Italian society of colorectal surgery (SICCR): management and treatment of hemorrhoidal disease. Tech Coloproctol. 2020 Feb;24(2):145-164
28. Brill AI, Fleshman JW, Ramshaw BJ, et al. Minimally invasive procedures: What family physicians need to know. J Fam Pract. 2005;54[Suppl 1]:S1–S24
29. Church J, Simmang C, Standards Task Force, American Society of Colon and Rectal Surgeons; Collaborative Group of the Americas on Inherited Colorectal Cancer, and the Standards Committee of the American Society of Colon and Rectal Surgeons. Practice parameters for the treatment of patients with dominantly inherited colorectal cancer (familial adenomatous polyposis and hereditary nonpolyposis colorectal cancer). Dis Colon Rectum. 2003;46:1001–1012
30. Fripp VT, Esquivel J, Cerruto CA. Perianal melanoma disguised as hemorrhoids: case report and discussion. J Natl Med Assoc. 2005;97:726–731
31. Maqbool A, Lintner R, Bokhari A, et al. Anorectal melanoma—3 case reports and a review of the literature. Cutis. 2004;73:409–413
32. Felz MW, Winburn GB, Kallab AM, et al. Anal melanoma: an aggressive malignancy masquerading as hemorrhoids. South Med J. 2001;94:880–885
33. Scarlett Y. Medical management of fecal incontinence. Gastroenterology. 2004;126[Suppl 1]:S55–63.
34. Hemorrhoids. In: Corman ML, ed. Colon and rectal surgery. 5th ed. Philadelphia:LippincottWilliams &Wilkins; 2004:177–253

35. Goligher JC, Duthie HL, Nixon HH, editors. Surgery of the anus, rectum and colon. 5. London: Baillière Tindall; 1984. pp. 98–149.
36. Nyström PO, Qvist N, Raahave D, et al. Stapled or Open Pile Procedure (STOPP) trial study group (2010) Randomized clinical trial of symptom control after stapled anopexy or diathermy excision for hemorrhoid prolapse. Br J Surg. 97:167–176
37. Lee MJ, Morgan J, Watson AJM, et al. A validated severity score for hemorrhoids as an essential prerequisite for future hemorrhoid trials. Tech Coloproctol. 2019 Jan;23(1):33-41
38. Pucher PH, Qurashi M, Howell AM, et al. Development and validation of a symptom-based severity score for hemorrhoidal disease: the Sodergren score. Colorectal Dis. 2015 Jul;17(7):612-8
39. Rørvik HD, Styr K, Ilum L, et al. Hemorrhoidal Disease Symptom Score and Short Health ScaleHD: New Tools to Evaluate Symptoms and Health-Related Quality of Life in Hemorrhoidal Disease. Dis Colon Rectum. 2019 Mar;62(3):333-342
40. Giordano P, Nastro P, Davies A, et al. Prospective evaluation of stapled hemorrhoidopexy versus transanal hemorrhoidal dearterialisation for stage II and III hemorrhoids: three-year outcomes. Tech Coloproctol. 2011 Mar;15(1):67-73
41. Giordano P, Tomasi I, Pascariello A, et al. Transanal dearterialization with targeted mucopexy is effective for advanced hemorrhoids. Colorectal Dis. 2014 May;16(5):373-6
42. Ratto C, Campenni P, Papeo F, et al. Transanal hemorrhoidal dearterialization (THD) for hemorrhoidal disease: a single-center study on 1000 consecutive cases and a review of the literature. Tech Coloproctol. 2017 Dec;21(12):953-962
43. Misra MC. Drug treatment of hemorrhoids. Drugs. 2005;65:1481-1491
44. Ho YH, Tan M, Seow-Choen F. Micronized purified flavonidic fraction compared favorably with rubber band ligation and fiber alone in the management of bleeding hemorrhoids: randomized controlled trial. Dis Colon Rectum. 2000;43:66–69
45. Thanapongsathorn W, Vajrabukka T. Clinical trial of oral diosmin (Daflon) in the treatment of hemorrhoids. Dis Colon Rectum. 1992;35:1085–1088
46. Perera N, Liolitsa D, Iype S, et al. Phlebotonics for hemorrhoids. Cochrane Database Syst Rev. 2012 Aug 15;(8):CD004322
47. Aziz Z, Huin WK, Badrul Hisham MD, et al. Efficacy and tolerability of micronized purified flavonoid fractions (MPFF) for hemorrhoids: A systematic review and meta-analysis. Complement Ther Med. 2018 Aug;39:49-55
48. van Tol RR, Kleijnen J, Watson AJM, et al. European Society of ColoProctology: guideline for hemorrhoidal disease. Colorectal Dis. 2020 Jun;22(6):650-662
49. Armstrong DN, Ambroze WL, Schertzer ME, et al. Harmonic Scalpel vs. electrocautery hemorrhoidectomy: a prospective evaluation. Dis Colon Rectum. 2001;44:558–564
50. Ho YH, Buettner PG. Open compared with closed hemorrhoidectomy: meta-analysis of randomized controlled trials. Tech Coloproctol. 2007;11:135-143

51. Arbman G, Krook H, Haapaniemi S. Closed vs. open hemorrhoidectomy--is there any difference? Dis Colon Rectum. 2000;43:31-34
52. Yang J, Cui PJ, Han HZ, et al. Meta-analysis of stapled hemorrhoidopexy vs LigaSure hemorrhoidectomy. World J Gastroenterol. 2013;19:4799-4807
53. Faucheron JL, Voirin D, Abba J. Rectal perforation with lifethreatening peritonitis following stapled hemorrhoidopexy. Br J Surg. 2012;99:746-753
54. Giordano P, Overton J, Madeddu F, et al. Transanal hemorrhoidal dearterialization: a systematic review. Dis Colon Rectum. 2009 Sep;52(9):1665-71
55. Consalvo V, D'Auria F, Salsano V. Transanal Hemorrhoidal Dearterialization With Doppler Arterial Identification Versus Classic Hemorrhoidectomy: A Retrospective Analysis of 270 Patients. Ann Coloproctol. 2019 Jun 30;35(3):118-122
56. Medina-Gallardoa A, Curbelo-Penaa Y, De Castroa X, et al. Is the severe pain after Milligan-Morgan hemorrhoidectomy still currently remaining a major postoperative problem despite being one of the oldest surgical techniques described? A case series of 117 consecutive patients. Int J Surg Case Rep. 2017;30:73–5
57. Morinaga K, Hasuda K, Ikeda T. A novel therapy for internal hemorrhoids: ligation of the hemorrhoidal artery with a newly devised instrument (Moricorn) in conjunction with a Doppler flowmeter. Am J Gastroenterol. 1995 Apr;90(4):610-3
58. Dal Monte PP, Tagariello C, Sarago M, et al. Transanal hemorrhoidal dearterialisation: nonexcisional surgery for the treatment of hemorrhoidal disease. Tech Coloproctol. 2007 Dec;11(4):333-8; discussion 338-9
59. Hung AJ, Chen J, Gill IS. Automated performance metrics and machine learning algorithms to measure surgeon performance and anticipate clinical outcomes in robotic surgery. JAMA Surg. 2018;153(8):770–1
60. Loftus TJ, Tighe PJ, Filiberto AC, et al. Artificial intelligence and surgical decision making. JAMA Surg. 2020;155(2):148–58
61. Komorowski M, Celi LA, Badawi O, et al. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med. 2018;24(11):1716–20
62. Horie Y, Yoshio T, Aoyama K, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc. 2019;89(1):25–32
63. Guo L, Xiao X, Wu C, et al. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). Gastrointest Endosc. 2020;91(1):41–51
64. Hashimoto R, Requa J, Dao T, et al. Artificial intelligence using convolutional neural networks for realtime detection of early esophageal neoplasia in Barrett's esophagus (with video). Gastrointest Endosc. 2020;91(6):1264-1271 e1

65. Sakamoto T, Goto T, Fujiogi M, et al. Machine learning in gastrointestinal surgery. Surg Today. 2022 Jul;52(7):995-1007
66. Deo RC. Machine learning in medicine. Circulation. 2015;132(20):1920–30
67. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015 Jan 6;162(1):W1-73
68. Joshi RD, Dhakal CK. Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches. Int J Environ Res Public Health. 2021 Jul 9;18(14):7346
69. Ingwersen EW, Stam WT, Meijs BJV, et al. Dutch Pancreatic Cancer Group. Machine learning versus logistic regression for the prediction of complications after pancreatoduodenectomy. Surgery. 2023 Sep;174(3):435-440
70. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019 Jun;110:12-22
71. Berg U, W-Dahl A, Nilsdotter A, et al. Fast-Track Programs in Total Hip and Knee Replacement at Swedish Hospitals-Influence on 2-Year Risk of Revision and Mortality. J Clin Med. 2021 Apr 14;10(8):1680
72. Song X, Liu X, Liu F, et al. Comparison of machine learning and logistic regression models in predicting acute kidney injury: A systematic review and meta-analysis. Int J Med Inform. 2021 Jul;151:104484
73. Milligan E, Morgan C, Jones L, et al. Surgical anatomy of the anal canal and the operative treatment of hemorrhoids. Lancet. 1937; 2:1119–1124
74. Ratto C, Giordano P, Donisi L, et al. Transanal hemorrhoidal dearterialization (THD) for selected fourth-degree hemorrhoids. Tech Coloproctol. 2011;15:191-7
75. Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. Ann Surg. 2009 Aug;250(2):187-96
76. Garrow CR, Kowalewski KF, Li L, et al. Machine Learning for Surgical Phase Recognition: A Systematic Review. Ann Surg. 2021 Apr 1;273(4):684-693
77. Bi Q, Goodman KE, Kaminsky J, et al. What is Machine Learning? A Primer for the Epidemiologist. Am J Epidemiol. 2019 Dec 31;188(12):2222-2239
78. Riahi V, Hassanzadeh H, Khanna S, et al. Improving preoperative prediction of surgery duration. BMC Health Serv Res. 2023 Dec 2;23(1):1343
79. Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In: Proceed- ings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016. p. 785–94
80. Giuliani A, Romano L, Necozione S, et al; EMODART3 Study Group. Excisional Hemorrhoidectomy Versus Dearterialization with Mucopexy for the Treatment of Grade III Hemorrhoidal Disease: The EMODART3 Multicenter Study. Dis Colon Rectum. 2023 Dec 1;66(12):e1254-e1263

81. Kantidakis G, Putter H, Litière S, et al. Statistical models versus machine learning for competing risks: development and validation of prognostic models. BMC Med Res Methodol. 2023 Feb 24;23(1):51
82. Hashimoto DA, Rosman G, Rus D, et al. Artificial Intelligence in Surgery: Promises and Perils. Ann Surg. 2018 Jul;268(1):70-76
83. Alwosheel A, van Cranenburgh S, Chorus CG, Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis, Journal of Choice Modelling (2018), doi: 10.1016/j.jocm.2018.07.002
84. Price WN. Big data and black-box medical algorithms. Sci Transl Med. 2018 Dec 12;10(471):eaao5333