Development and external validation of an international, multicenter machine learning algorithm for prediction of anastomotic insufficiency after colonic or colorectal anastomosis

The Prediction of Anastomotic Insufficiency risk after Colorectal surgery (PANIC) study

Anas Taha[1,2 *], Stephanie Taha-Mehlitz[3,5 *], Ahmad Hendie[1,4], Tobias Staudner[1], Michel Adamina[1,2]

1: Clinical Research and Artificial Intelligence in Surgery, Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland

2: Department of Surgery, Clinic for Visceral and Thorax Surgery, Cantonal Hospital of Winterthur, Winterthur, Switzerland

3: Department of Visceral Surgery, Clarunis, University Hospital Basel, Basel, Switzerland

4: Department of Electrical and Computer engineering, University of McGill, Montreal, Canada

**: equally contributed, shared first authorship*

Introduction

Colorectal anastomotic leaks or anastomotic insufficiency (AI) lead to consequential clinical and economic strain on patients, and significantly increases morbidity and mortality. [1] On average, hospital stay is extended by 12 days while healthcare-related expenses are increased by 30’000 US$ in patients who suffered AI. [1] In experienced centers, 3.3% is the approximated incidence of AI after colon anastomosis, with 8.6% for colorectal anastomoses. [2]

In previous publications, multiple subgroups of patients with increased hazard for AI have been found. [1, 3-7] Meticulous preoperative recognition of patients with heightened risk for AI is clinically beneficial, as it would permit improved resource preparation, enhanced patient education and informed consent, superior surgical decision-making, and possibly even adjustment of specific risk factors. However, it is often futile for clinicians to try balancing the many crucial risk factors for AI to attain a personalized risk for AI in a single patient.

Machine learning (ML) methods have been exceptionally competent at incorporating various clinical patient variables into one unified risk prediction model designed specifically for each patient. To the authors’ best knowledge, there does not yet exist a credible prediction model or a conclusive prediction score for AI after colorectal anastomosis.

The aim of the Prediction of Anastomotic Insufficiency risk after Colorectal surgery (PANIC) study is to establish and externally validate an efficient ML-based prediction tool based on multicenter data from a range of international centers.

Methods

Overview

Data is compiled from an assortment of international centers. Principally, the models will be assembled, and publication will be constructed according to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines. [8] One model will be designed for colon-to-colon anastomosis (PANIC-C) and one distinct separate model will be constructed for colorectal anastomosis (PANIC-R).

The Clinical Research and Artificial Intelligence in Surgery research group, Department of Biomedical Engineering, University of Basel and the Department of Surgery, Clinic for Visceral and Thorax Surgery, Cantonal Hospital, Winterthur are the sponsors of this study. Dr. med. A. Taha and Prof. Dr. med. Michel Adamina are the lead investigators.

Ethical Considerations

Each center will be accountable for their own ethics board / institutional review board (IRB) approval. They must acquire approval for retrospective or prospective data collection and sharing of the completely deidentified data with the sponsor. The sponsor can assist by providing said protocol, which will also be registered on ClinicalTrials.gov.

Authorship

Each center must share complete data from at least 500 patients to be included. Each cooperating center may assign two center leads and as many contributors as required to enter the data, not exceeding a total of 10 named contributors. Centre leads and center-specific contributors will be listed as full members of the PANIC study group and will be granted PubMed / Medline contributor status in a group authorship model.

Inclusion and Exclusion Criteria

Consecutive patients who underwent colocolic or colorectal anastomosis for neoplasia, diverticulitis, mesenterial ischemia, iatrogenic or traumatic perforation, or inflammatory bowel disease will be included. Patients younger than 18 years of age, who have a recurrent colorectal cancer, beared peritoneal carcinomatosis or unresectable metastatic disease at time of bowel resection and anastomosis, or who could not provide consent to participate according to individual institution's rules will be excluded. Moreover, patients who were lost to follow-up less than six weeks after surgery and who have not noted a successful ostomy reversal will be excluded - in these cases, it cannot be surely assessed whether AI occurred or not.

Data Collection

Each center will assemble their data either prospectively or retrospectively from a dedicated clinical database. Each center must provide data from at least 500 patients over not more than 10 years to be included, this in order to ensure data consistency and centre proficiency as monitored by adequate case load.

A standardized data collection form will be given by the sponsor. The data will be registered in a standardized and deidentified form which ensures data completion and anonymity. A study-specific identifiant will be produced upon completion of data entrance of each patient. The study database will not have any identifiable patient data. Contributing centers will maintain a log file in which the study-specific patient numbers can be followed to center-specific patient-numbers, in case of a necessity.

Primary Endpoint Definition

AI (Anastomotic insufficiency/leakage) is defined according to Gessler et al. [6] and Rahbari et al. [9] as any clinical signs of leakage, confirmed by radiological examination, endoscopy, clinical examination of the anastomosis, or upon reoperation.

Secondary Endpoint Definitions

AI will be **graded** retrospectively conferring to the system suggested by Rahbari et al. [9] Anastomosis **takedown** is defined as an interruption of the continuity of the bowel and the formation of a stoma, which will be captured in the database. In cases of AI, **death** will be recorded within 90 days from index surgery. In cases of AI, **time to diagnosis of a leakage** will be calculated as days between the index operation and diagnosis of the leakage by imaging with extraluminal contrast, endoscopy, reoperation, or when fecal containing fluid is objectified in a drainage.

Tertiary Endpoint Definitions

To investigate a potential relationship between AI and oncologic outcomes in a further dedicated study

5-year follow-up

- Disease-free survival (months)
- Overall survival (months)
- Adjuvant/additive immunochemotherapy (yes/no)
-

- Adjuvant/additive radiotherapy (yes/no)
- Additive curative surgery (no/yes: liver; lung; locoragional relapse)

Features and Their Definitions

All features are measured preoperatively within 30 days of index surgery. The following data will be collected.

For colon-to-colon anastomosis:

- Active smoking (yes/no, pack/years) [10, 11]
- Alcohol abuse (>2 alcoholic beverages per day) [11]
- Height / Weight / BMI [5]
- Preoperative leukocyte count (in x 10^3 per mm^3) [7]
- Preoperative steroid use (mg) [3, 4]
- Preoperative non steroidal inflammatory drug use (NSAID)
- Neoadjuvant chemotherapy (yes/no)
- Charlson Comorbidity Index (CCI) [12]
- American Society of Anesthesiologists (ASA) Score [13-15]
- Renal function (CKD Stages G1 (normal) to G5) [7]
- Nutritional status (NRS ≥ 3) [16]
- Hypoalbuminemia (<3.5 g/dL) [17-19]
- Hemoglobin level (in g/dL) [20, 21]
- Prior abdominal surgery (any) [22]
- Liver metastasis at time of anastomosis (any) [22]
- Age (yrs.)
- Albumin and Prealbumin
- Perioperative blood transfusion
- Psychosocial diseases
- Gender (M/F) [12, 15, 18, 22, 23]
- Anastomotic configuration (stapler anastomosis, stapled and suture closure of stapler enterotomy, hand-sewn; end-to-end; side-to-end; pouch to end)
- Emergency surgery (yes/no)
- Approach: Laparoscopic, transanal, robotic, open
- Extended resection
- Conversion to laparoscopic/open surgery (yes/no)

- Left hemicolectomy / right hemicolectomy / extended left hemicolectomy / extended right hemicolectomy / ileocaecal resection / transverse colectomy / sigmoid resection
- Perforation (yes/no; when yes within 5cm from tumor (yes/no))
- Protective ostomy (yes/no)
- Intra- and postoperative complications
- Date of discharge and discharge status
- Postoperative leucocyte count and CRP level up to 10 days postoperative
- Oncologic resection (yes/no; requires at least 12 LN and integrity of mesocolic/mesorectal planes)
- Indication (Neoplasia, IBD, perforation, diverticulitis, mesenterial ischemia)
- In case of colon cancer, preoperative carcinoembryonic antigen (CEA, ng/ml) and pathological TNM staging according to TNM 8th edition (T, N, M, L, V, Pn, grading (high, low), resection status (R0, R1, R2)
- In case of colon cancer length of follow-up, disease-free and overall survival

For colo-rectal anastomosis:

- Preoperative tumor stage (MRI): Local (cT) [24-26]
- Preoperative tumor stage (MRI): Nodal (cN) [24, 26]
- Preoperative circumferential and distal margins (MRI)
- Preoperative endovascular margin (MRI)
- Preoperative tumor stage (CT thorax abdomen): Metastasis (cM) [24, 26]
- Neoadjuvant radiotherapy (yes/no) [26-28]
- Neoadjuvant chemotherapy (yes/no)
- Active smoking (yes/no, pack/years) [11, 29]
- Alcohol abuse (>2 alcoholic beverages per day) [11, 30]
- Height / Weight / BMI [31]
- Preoperative leukocyte count (in x 10^3 per mm^3) [7]
- Preoperative steroid use (mg) [3, 4, 32]
- Preoperative non steroidal inflammatory drug use (NSAID)
- Charlson Comorbidity Index (CCI) [12]
- American Society of Anesthesiologists (ASA) Score [13, 14, 32]
- Renal function (CKD Stages G1 (normal) to G5) [7, 33]
- Nutritional status (NRS $\geq$ 3) [16, 34]
- Hypoalbuminemia (<3.5 g/dL) [17-19]
- Hemoglobin level (in g/dL) [20, 21, 35]
- Prior abdominal surgery (any) [22]

- Age (yrs.)
- Albumin/Prealbumin
- Perioperative blood transfusion
- Psychosocial diseases
- Liver metastasis at time of anastomosis (any) [22]
- Gender (M/F) [12, 18, 22, 26, 27, 34, 36, 37]
- Emergency surgery (yes/no)
- Approach: Laparoscopic, robotic, open
- Extended resection
- Conversion to open surgery (yes/no)
- Anastomotic configuration (stapler anastomosis, hand-sewn; end-to-end; side-to-end; pouch to end)
- Number of strapler cartridges used
- Intra- and postoperative complications
- Anal verge distance (cm) from the tumor lower edge [12, 22, 24-27, 32, 38]
- Anal verge distance (cm) from the anastomosis [12, 22, 24-27, 32, 38]
- Circumferential margin (mm)
- Protective ostomy (yes/no)
- Type of anasotomotic testing
- Postoperative leucocyte count and CRP level up to 10 days postoperative
- Date of discharge and discharge status
- Indication (Neoplasia, IBD, perforation, diverticulitis, mesenterial ischemia)
- In case of rectal cancer, preoperative carcinoembryonic antigen (CEA, ng/ml) and pathological TNM staging according to TNM 8th edition (T, N, M, L, V, Pn, grading (high, low), resection status (R0, R1, R2)
- In case of rectal cancer length of follow-up, disease-free and overall survival

Sample Size

While even the largest cohort with millions of patients cannot ascertain a perfect clinical prediction model if no significant input variables are included ("garbage in, garbage out" – do not expect to predict the future from age, gender, and body mass index), the relationship among predictive performance and sample size is certainly directly proportional, especially for some data-hungry ML algorithms. Moreover, to protect the generalizability of the clinical prediction model, the sample size must be both representative enough of the patient population and should consider the complexity of the algorithm. For instance, a deep neural network – as an example of a highly complex model – will need thousands of patients to assemble, while a logistic regression model may accomplish stable results with only a few hundreds of patients. In addition, the number of input variables has a role to play. Roughly, it can be said

that a bare minimum of 10 positive cases are needed per included input variable to model the relationships. Often, eccentric behavior of the models and high variance in performance among splits is noted when sample sizes are smaller than calculated with this rule of thumb. Of immense importance is also the proportion of patients who suffered the outcome. For very rare events, a much larger total sample size is required. For example, a prediction based on 10 input features for an outcome occurring in only 10% of cases would need at least 1000 patients including at least 100 who suffered the outcome, according to the above rule of thumb. In general, from personal experience, we do not propose producing ML models on cohorts with less than 100 positive cases and wisely more cases in total, regardless of the rarity of the outcome. Also, one should examine the available literature on risk factors for the outcome of interest: If epidemiological studies acquire only feeble associations with the outcome, it is probably that one will need plenty more patients to reach a model with satisfactory predictive performance, as opposed to an outcome which has several highly associated risk factors, which may be easier to predict. Bigger sample sizes also permit more generous evaluation through a larger amount of patient data dedicated to training or validation, and usually results in superior calibration measures.

For this study, based on our expertise and on the rules of thumb mentioned above, we approximated that a minimum of 250 patients with AI are needed to identify generalizable features. With an incidence of AI of 5% for colon anastomosis and 8% for colorectal anastomosis, that means that a minimum of 5000 and 3125 patients are required for training for each population (colocolic and colo-rectal anastomosis, respectively). For satisfactory evaluation of calibration at external validation, we estimate that another 2000 and 1250 patients per cohort (thus, approximately 100 with AI per cohort) will be required. Thus, in total, we estimate that **7000 (colocolic anastomosis) and 4375 (colorectal anastomosis) patients** are necessary to reach at a satisfying model. More data will likely lead to greater performance and better calibration.

Predictive Modeling

A KNN imputer will be co-trained to impute any missing data that may occur in future application of the model. If there is missing data in the training set, it will be imputed using said KNN imputer. [39] Features or patients with a missingness greater than 25% will be excluded. Data will be standardized and one-hot-encoded. In case of major class imbalance – which is expected for the abovementioned endpoint – random upsampling or synthetic minority oversampling (SMOTE) will be applied to the training set. [40, 41] All features will at first be given to the model for training. If needed, we will employ recursive feature elimination (RFE) to choose input features on the training data. [42]

We will trial the following algorithms for binary classification: Generalized linear model (GLM), generalized additive model (GAM), stochastic gradient boosting machine (GBM), naïve Bayes classifier, artificial neural network, support vector machine (SVM), and random forest. Each model will be fully trained and hyperparameter tuned where applicable. The final model will be selected based upon (AUC), sensitivity, and specificity, as well as calibration metrics on the resampled training performance. Training will happen in repeated 5-fold cross-validation with 10 repeats.

The one final model will then be examined on the external validation data only once. Ninety-five percent confidence intervals for external validation metrics will be derived using the bootstrap.

The threshold for binary classification will either be identified on the training data alone using the AUC-based "closest-to-(0,1)-criterion" or Youden's index to optimize both sensitivity and specificity or will be optimized on the training set based on clinical significance (rule-out model). The analyses will be carried out in R Version 3.6.2. [43] or Phyton.

Evaluation

The efficacy of classification models can be approximately evaluated along two dimensions: Model discrimination and calibration. [44] The term *discrimination* denotes the ability of a prediction model to correctly classify whether a certain patient is going to or is not going to suffer a certain outcome. Thus, discrimination describes the accuracy of a binary prediction – yes or no. *Calibration*, on the other hand, defines the degree to which a model's predicted probabilities (ranging from 0% to 100%) correspond to the observed incidence of the binary endpoint (true posterior). Many publications do not report calibration metrics, although these are of central importance, as a well-calibrated predicted probability (e.g. your predicted probability of experiencing a complication is 18%) is often much more valued to clinicians – and patients! – than a binary prediction (e.g. you are likely not going to experience a complication). [44]

Resampled training performance as well as performance on the external validation set will be examined for discrimination and calibration. In terms of discrimination, we will assess AUC, accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1 Score. In terms of calibration, we will assess the Brier score, expected-observed (E/O) ratio, calibration slope and intercept, the Hosmer-Lemeshow goodness-of-fit test, as well as visual inspection of calibration plots for both datasets, which will also be included in the publication.

Interpretability

The degree and choice of methods for interpretability will rely upon on the final chosen algorithm. Some algorithms can natively give explanations as to which factors affect the outcome in what way. Thus, in case e.g., a GLM, GAM, or naïve Bayes classifier is chosen, the parameters / partial dependence values will be provided. For simple decision trees, diagrams of the decision-making process can be given. Other models with higher degrees of complexity, such as neural networks or stochastic gradient boosting machines cannot natively provide such explanations. In that case, we will provide both AUC-based variable importance as well as model-agnostic local interpretations of variable importance using the LIME principle. [45]

Expected Results

We foresee arriving to a generalizable model based on multicenter international data that is likely to predict AI consistently with an AUC of at least 0.70 (more realistically 0.8 or greater) and that is well-calibrated. [44] A web-based prediction tool will also be designed for each of the two models applying the *shiny* [46] environment, much likened to e.g., https://neurosurgery.shinyapps.io/impairment. This web-based app will be accessible for free on any internet-capable device (mobile or desktop), and should be stable on most devices due to the server-based computing. The costs for maintaining the server will be carried by the main investigators.

The assembled data will be saved for 10 years by the sponsor. The large dataset will be available to further analysis and will be given to any of the participating centers at request and after approval by all other centers. The goal is to facilitate other analyses using the collected dataset. If any other analyses reach publication, all contributors will be included as co-authors and all co-authors will have the opportunity to review said manuscript beforehand. Any contributing study center has the right to veto publication of any subsequent analyses containing their own data.

Acknowledgement: The study idea and the study protocol were developed together with Victor E. Staartjes (Machine Intelligence in Clinical Neuroscience (MICN) Laboratory, Clinical Neuroscience Center, University Hospital Zurich).

References

1. Lee, S.W., D. Gregory, and C.L. Cool, Clinical and economic burden of colorectal and bariatric anastomotic leaks. Surg Endosc, 2019.
2. Krell, R.W., et al., Hospital readmissions after colectomy: a population-based study. J Am Coll Surg, 2013. 217(6): p. 1070-9.
3. Konishi, T., et al., Risk factors for anastomotic leakage after surgery for colorectal cancer: results of prospective surveillance. J Am Coll Surg, 2006. 202(3): p. 439-44.
4. Slieker, J.C., et al., Long-term and perioperative corticosteroids in anastomotic leakage: a prospective study of 259 left-sided colorectal anastomoses. Arch Surg, 2012. 147(5): p. 447-52.
5. Silva-Velazco, J., et al., Is there anything we can modify among factors associated with morbidity following elective laparoscopic sigmoidectomy for diverticulitis? Surg Endosc, 2016. 30(8): p. 3541-51.
6. Gessler, B., O. Eriksson, and E. Angenete, Diagnosis, treatment, and consequences of anastomotic leakage in colorectal surgery. Int J Colorectal Dis, 2017. 32(4): p. 549-556.
7. Alves, A., et al., Factors associated with clinically significant anastomotic leakage after large bowel resection: multivariate analysis of 707 patients. World J Surg, 2002. 26(4): p. 499-502.
8. Collins, G.S., et al., Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Bmj, 2015. 350: p. g7594.
9. Rahbari, N.N., et al., Definition and grading of anastomotic leakage following anterior resection of the rectum: a proposal by the International Study Group of Rectal Cancer. Surgery, 2010. 147(3): p. 339-51.
10. Kwak, H.D., et al., Risk Factors and Oncologic Outcomes of Anastomosis Leakage After Laparoscopic Right Colectomy. Surg Laparosc Endosc Percutan Tech, 2017. 27(6): p. 440-444.
11. Sorensen, L.T., et al., Smoking and alcohol abuse are major risk factors for anastomotic leakage in colorectal surgery. Br J Surg, 1999. 86(7): p. 927-31.
12. Trencheva, K., et al., Identifying important predictors for anastomotic leak after colon and rectal resection: prospective study on 616 patients. Ann Surg, 2013. 257(1): p. 108-13.
13. Buchs, N.C., et al., Incidence, consequences, and risk factors for anastomotic dehiscence after colorectal surgery: a prospective monocentric study. Int J Colorectal Dis, 2008. 23(3): p. 265-70.
14. Choi, H.K., W.L. Law, and J.W. Ho, Leakage after resection and intraperitoneal anastomosis for colorectal malignancy: analysis of risk factors. Dis Colon Rectum, 2006. 49(11): p. 1719-25.
15. Bakker, I.S., et al., Risk factors for anastomotic leakage and leak-related mortality after colonic cancer surgery in a nationwide audit. Br J Surg, 2014. 101(4): p. 424-32; discussion 432.
16. Kwag, S.J., et al., The nutritional risk is a independent factor for postoperative morbidity in surgery for colorectal cancer. Ann Surg Treat Res, 2014. 86(4): p. 206-11.
17. Ionescu, D., C. Tibrea, and C. Puia, Pre-operative hypoalbuminemia in colorectal cancer patients undergoing elective surgery - a major risk factor for postoperative outcome. Chirurgia (Bucur), 2013. 108(6): p. 822-8.
18. Suding, P., et al., Definitive risk factors for anastomotic leaks in elective open colorectal resection. Arch Surg, 2008. 143(9): p. 907-11.
19. Lai, C.C., et al., Low preoperative serum albumin in colon cancer: a risk factor for poor outcome. Int J Colorectal Dis, 2011. 26(4): p. 473-81.
20. Iancu, C., et al., Host-related predictive factors for anastomotic leakage following large bowel resections for colorectal cancer. J Gastrointestin Liver Dis, 2008. 17(3): p. 299-303.
21. Zaharie, F., et al., Risk factors for anastomotic leakage following colorectal resection for cancer. Chirurgia (Bucur), 2012. 107(1): p. 27-32.
22. Lipska, M.A., et al., Anastomotic leakage after lower gastrointestinal anastomosis: men are at a higher risk. ANZ J Surg, 2006. 76(7): p. 579-85.
23. Krarup, P.M., et al., A nationwide study on anastomotic leakage after colonic cancer surgery. Colorectal Dis, 2012. 14(10): p. e661-7.
24. Zhu, Q.L., et al., Laparoscopic low anterior resection for rectal carcinoma: complications and management in 132 consecutive patients. World J Gastroenterol, 2010. 16(36): p. 4605-10.
25. Eberl, T., et al., Risk factors for anastomotic leakage after resection for rectal cancer. Am J Surg, 2008. 196(4): p. 592-8.

26. Park, J.S., et al., Multicenter analysis of risk factors for anastomotic leakage after laparoscopic rectal cancer excision: the Korean laparoscopic colorectal surgery study group. Ann Surg, 2013. 257(4): p. 665-71.
27. Hamabe, A., et al., Preventive effect of diverting stoma on anastomotic leakage after laparoscopic low anterior resection with double stapling technique reconstruction applied based on risk stratification. Asian J Endosc Surg, 2018. 11(3): p. 220-226.
28. Kobayashi, M., et al., Risk factors for anastomotic leakage and favorable antimicrobial treatment as empirical therapy for intra-abdominal infection in patients undergoing colorectal surgery. Surg Today, 2014. 44(3): p. 487-93.
29. Kim, M.J., et al., The impact of heavy smoking on anastomotic leakage and stricture after low anterior resection in rectal cancer patients. World J Surg, 2011. 35(12): p. 2806-10.
30. Makela, J.T., H. Kiviniemi, and S. Laitinen, Risk factors for anastomotic leakage after left-sided colorectal resection with rectal anastomosis. Dis Colon Rectum, 2003. 46(5): p. 653-60.
31. Yamamoto, S., et al., Risk factors for anastomotic leakage after laparoscopic surgery for rectal cancer using a stapling technique. Surg Laparosc Endosc Percutan Tech, 2012. 22(3): p. 239-43.
32. Lai, R., et al., Risk factors for anastomotic leakage following anterior resection for colorectal cancer: the effect of epidural analgesia on occurrence. Int J Colorectal Dis, 2013. 28(4): p. 485-92.
33. Piecuch, J., et al., Mortality and anastomotic leakage after anterior resection for rectal cancer. Hepatogastroenterology, 2012. 59(115): p. 721-3.
34. Kang, C.Y., et al., Risk factors for anastomotic leakage after anterior resection for rectal cancer. JAMA Surg, 2013. 148(1): p. 65-71.
35. Hayden, D.M., et al., Patient factors may predict anastomotic complications after rectal cancer surgery: Anastomotic complications in rectal cancer. Ann Med Surg (Lond), 2015. 4(1): p. 11-6.
36. Tanaka, K., et al., Risk factors for anastomotic leakage after laparoscopic surgery with the double stapling technique for stage 0/I rectal carcinoma: a subgroup analysis of a multicenter, single-arm phase II trial. Surg Today, 2017. 47(10): p. 1215-1222.
37. Law, W.I., et al., Risk factors for anastomotic leakage after low anterior resection with total mesorectal excision. Am J Surg, 2000. 179(2): p. 92-6.
38. Akiyoshi, T., et al., Incidence of and risk factors for anastomotic leakage after laparoscopic anterior resection with intracorporeal rectal transection and double-stapling technique anastomosis for rectal cancer. Am J Surg, 2011. 202(3): p. 259-64.
39. Templ M, K.A., Alfons A, Prantner B, VIM: Visualization and Imputation of Missing Values [Internet] Available from: https://CRAN.R-project.org/package=VIM, 2019 [cited 2020 Jan 5].
40. Chawla, N., et al., SMOTE: Synthetic Minority Over-sampling Technique. J. Artif. Intell. Res. (JAIR), 2002. 16: p. 321-357.
41. Staartjes, V.E. and M.L. Schröder, Letter to the Editor. Class imbalance in machine learning for neurosurgical outcome prediction: are our models valid? J Neurosurg Spine, 2018. 29(5): p. 611-612.
42. Granitto, P., et al., Recursive Feature Elimination with Random Forest for PTR-MS analysis of agroindustrial products. Chemometrics and Intelligent Laboratory Systems. 2006 Sept: 83 (2) 83-90.
43. R Core Team, R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austrau: R Foundation for Statistical Computing; 2020. Available from: https://www.R-project.org/
44. Staartjes, V.E. and J.M. Kernbach, Letter to the Editor. Importance of calibration assessment in machine learning-based predictive analytics. J Neurosurg Spine, 2020: p. 1-2.
45. Ribeiro, M.T., S. Singh, and C. Guestrin, "Why Should I Trust You?": Explaining the Predictions of Any Classifier, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, Association for Computing Machinery: San Francisco, California, USA. p. 1135–1144.
46. Chang, W., et al., Shiny: web application framework for *R* [Internet]. 2020 [cited 2020 May 28]. Available from: https://CRAN.R-project.org/package=shiny