Pituitary
https://doi.org/10.1007/s11102-022-01216-0

# Machine learning as a clinical decision support tool for patients with acromegaly

Cem Sulu[1] · Ayyüce Begüm Bektaş[2] · Serdar Şahin[1] · Emre Durcan[1] · Zehra Kara[1] · Ahmet Numan Demir[1] · Hande Mefkure Özkaya[1,3] · Necmettin Tanrıöver[3,4] · Nil Çomunoğlu[5] · Osman Kızılkılıç[3,6] · Nurperi Gazioğlu[3,7] · Mehmet Gönen[8,9] · Pınar Kadıoğlu[1,3]

Accepted: 10 March 2022


## Abstract

**Objective** To develop machine learning (ML) models that predict postoperative remission, remission at last visit, and resistance to somatostatin receptor ligands (SRL) in patients with acromegaly and to determine the clinical features associated with the prognosis.

**Methods** We studied outcomes using the area under the receiver operating characteristics (AUROC) values, which were reported as the performance metric. To determine the importance of each feature and easy interpretation, Shapley Additive explanations (SHAP) values, which help explain the outputs of ML models, are used.

**Results** One-hundred fifty-two patients with acromegaly were included in the final analysis. The mean AUROC values resulting from 100 independent replications were 0.728 for postoperative 3 months remission status classification, 0.879 for remission at last visit classification, and 0.753 for SRL resistance status classification. Extreme gradient boosting model demonstrated that preoperative growth hormone (GH) level, age at operation, and preoperative tumor size were the most important predictors for early remission; resistance to SRL and preoperative tumor size represented the most important predictors of remission at last visit, and postoperative 3-month insulin-like growth factor 1 (IGF1) and GH levels (random and nadir) together with the sparsely granulated somatotroph adenoma subtype served as the most important predictors of SRL resistance.

**Conclusions** ML models may serve as valuable tools in the prediction of remission and SRL resistance.

**Keywords** Machine learning · Acromegaly · Prognosis · Somatostatin receptor ligand

## Introduction

Acromegaly is a rare chronic disease characterized by excessive production of growth hormone (GH) and insulin-like growth factor 1 (IGF1), caused by pituitary adenoma in the vast majority of cases [1]. It has a worldwide prevalence of 40–130 per million [2]. Prolonged exposure to an excess of GH leads to somatic disfigurement, a wide range of

✉ Pınar Kadıoğlu
kadioglup@yahoo.com

1 Department of Internal Medicine, Division of Endocrinology, Metabolism, and Diabetes, Cerrahpasa Medical School, Istanbul University-Cerrahpaşa, Kocamustafapaşa Street No:53, 34098 Fatih Istanbul, Turkey

2 Graduate School of Sciences and Engineering, Koç University, Istanbul, Turkey

3 Pituitary Center, Istanbul University-Cerrahpasa, Istanbul, Turkey

4 Department of Neurosurgery, Cerrahpasa Medical School, Istanbul University-Cerrahpasa, Istanbul, Turkey

5 Department of Medical Pathology, Cerrahpasa Medical School, Istanbul University-Cerrahpasa, Istanbul, Turkey

6 Department of Radiology, Cerrahpasa Medical School, Istanbul University-Cerrahpasa, Istanbul, Turkey

7 Department of Neurosurgery, Istinye University, Istanbul, Turkey

8 Department of Industrial Engineering, College of Engineering, Koç University, Istanbul, Turkey

9 School of Medicine, Koç University, Istanbul, Turkey

Published online: 18 April 2022

manifestations including carcinogenesis; thus bringing the risk of increased morbidity and mortality [3, 4]. Transsphenoidal surgery (TSS) serves as the treatment of choice for patients with acromegaly, yet this approach is effective in only 60–70% of the patients at best even in centers with large experience [5–8]. A considerable percentage of the patients remain uncontrolled and need further treatments. Thus, the prediction of postoperative remission and response to other treatments is an area of active investigation.

To date, a large body of clinical studies evaluated possible predictors of surgical outcomes [9–20]. Some of these studies suggested that factors other than tumor invasiveness such as age, sex, preoperative and postoperative GH and IGF1 levels, granulation pattern, and proliferation indexes could predict postoperative remission, whereas others failed to demonstrate the predictive power of these factors [5, 10, 11, 15, 16, 18, 21–26]. In these studies, pre-determined cut-offs and timing of GH and IGF1 evaluations were also somewhat arbitrary [25, 27–30]. Besides, they have also been criticized for their narrow scope of simple analysis of prognostic factors. Prognosis and decisions about treatment strategy should not be solely based on a single factor. The decision process should be constructed on a combined analysis of multiple important clinical features. There is an inevitable need for robust prediction models that are capable of assessing multiple prognostic factors at one hand with high accuracy. Recently, machine learning (ML) models are gaining momentum [31]. ML is a subset of artificial intelligence that provides information by automatically acquiring patterns from databases instead of being conditioned with rules. The influence from users' intervention is avoided, discriminating it from traditional methods. Several studies tested its capability to foresee surgical outcomes in patients with pituitary adenoma [32, 33]. In particular, attempts have been made to predict either early or delayed postoperative remission status of patients with acromegaly using ML models [34–36]. Yet, none of them aimed to assess the clinical and biomedical predictors of resistance to long-acting somatostatin receptor ligands (SRL). We proposed that the development of a comprehensive and reproducible prediction model for postoperative remission status and drug resistance could aid in tailoring treatment decisions. In this study, our objective was to develop ML algorithms that predict remission at postoperative 3 months, remission at last visit, resistance to SRL in patients with acromegaly, and to determine the clinical features that affect these responses the most.

## Materials and methods

This single-center, retrospective study was conducted in the Pituitary Center of a tertiary care university hospital. The study was approved by the Research Ethics Committee of Istanbul University-Cerrahpaşa. The study fully adheres to the Declaration of Helsinki. Patient data were coded and stored anonymously.

### Study Sample and Procedure

All the medical records of the 535 patients with acromegaly that were followed up at the Endocrinology, Metabolism, and Diabetes outpatient clinic of Cerrahpaşa Medical Faculty between 2000 and 2020 were reviewed. Inclusion criteria were (i) a clear-cut biochemical and pathological diagnosis of acromegaly as outlined in the current Endocrine Society guideline [37]; (ii) patients with regular follow-up for at least 12 months. Patients with acromegaly had missing information concerning any of the following clinical, biochemical, radiological, and histopathological features were excluded: (i) baseline demographic characteristics (sex, age at diagnosis and surgery); (ii) radiological features (tumor size at diagnosis based on maximum diameter in millimeters, cavernous sinus invasion); (iii) histopathological features (granulation pattern of the tumor, ki-67 index, mitosis index, immunohistochemical staining for additional hormones); (iv) biochemical results (IGF1, random/nadir GH at diagnosis and during follow-up; v) type, dosage, and duration of the postoperative treatments (when present, categorized as a dopamine agonist [DA], somatostatin analog [SRL, i.e., lanreotide ATG or octreotide LAR], and GH receptor antagonist [GHRA; i.e., pegvisomant], repeat surgery, and radiotherapy). Patients with acromegaly who received first-line medical therapy or radiotherapy due to ineligibility for TSS were also excluded. Finally, a total of 152 patients out of 535 patients were included.

The features to be tested in ML algorithms for prediction of early remission, remission at the last visit, and SRL resistance were chosen among above-mentioned parameters in accordance with previous literature and clinical cognition.

### Endocrinological Assessment

The biochemical diagnostic criteria for acromegaly were as follows: (1) adult patients with clinical symptoms of acromegaly, (2) patients with pituitary adenoma confirmed by pituitary magnetic resonance imaging (MRI), and (3) preoperative IGF1 values exceeding the upper limit of age- and the sex-adjusted reference range, (4) lack of suppression of GH to < 1.0 ng/ml during an oral glucose load [37]. The diagnosis was also confirmed with histopathological examination of the tumor. All the patients with acromegaly underwent TSS as first-line therapy performed by two senior neurosurgeons of the Pituitary Center. Off note, patients with severe pharyngeal thickness, sleep apnea syndrome, high-output heart failure or ventricular dysrhythmia received preoperative medical treatment to provide anesthetic safety as recommended by the current Endocrine Society guideline [38, 39]. We also preferred to initiate preoperative medical treatment to patients with acromegaly whose expected time to TSS was beyond one month due to workload of Pituitary Center [40]. Three months after the TSS, we measured IGF1, random GH, and nadir GH during an oral glucose tolerance test to determine remission status. Definition of early remission was based on off-medication GH levels (nadir GH < 0.4 ug/L during an oral glucose tolerance test, and random GH < 1.0 ug/L) and age- and sex-adjusted normalized IGF1 at postoperative 3 months [37]. Repeat surgery was performed for patients with persistent disease with accessible intrasellar residual mass. For the remaining patients with persistent disease, initial medical adjuvant therapy was initiated. In patients with moderate-to-severe symptoms and signs of GH excess we opted SRL as the initial adjuvant medical therapy. Patients with modest elevations of IGF1 and mild symptoms and signs of GH excess, a trial of cabergoline was also considered. Biochemical response to SRL was defined as normalization of age- and sex-adjusted IGF1 levels and random GH < 1 ug/L [37, 41]. Patients with acromegaly who did not present a biochemical response after 12 months of treatment with maximal doses of SRL (30–40 mg octreotide long-acting repeatable (LAR) or lanreotide autogel (ATG) 120 mg every 28 days) were deemed to be SRL resistant [42]. These patients either received pegvisomant or cabergoline was added [38]. In patients with inaccessible residual tumor mass after TSS and medical therapy was unsuccessful or not tolerated, radiation therapy was also considered [38].

## Response Classification using Extreme Gradient Boosting

eXtreme Gradient Boosting (XGBoost) is an ML algorithm based on decision trees [43]. It works by ensembling many decision trees for the desired classification or regression task. Since it grows the decision trees sequentially and iteratively based on the previous trees, thus correcting itself while learning. It is considered one of the most powerful tree-based ML algorithms of our day. In this study, we applied XGBoost to construct ML models to determine predictive factors related to remission and resistance to SRL.

## Determining Feature Importance for Clinical Applicability and Interpretability

One of the most challenging aspects of ML applications is their capability to interpret the results while having good predictive performance. Since, in this study, we targeted to build ML models which classify the patients with acromegaly according to their postoperative 3 months remission status, remission status at last visit, and resistance to SRL; the importance of related clinical features while making classification for each outcome is significant for this tool to become a clinical decision support tool. In our study, as training ratio, we took 80% of the whole acromegaly data set for model training and used the remaining 20% for model testing. We partitioned the data randomly to form the training and test sets, so that they will be similar in terms of features. We performed 100 independent replications of our experiment for more robust results, where for each replication we randomly selected 80% of the whole acromegaly data set. Then, we used the remaining 20% as the test dataset, which is again a collection of previously unseen data instances to each learned model in each independent replication. It should be noted here that randomness used in our independent replications are controlled by usage of different seed values and our results are reproducible We used four-fold cross-validation for hyperparameter tuning. After having performed 100 independent replications, for clinical applicability and easy interpretation, we calculated the Shapley Additive exPlanation (SHAP) values for each feature using SHAP for XGBoost R package [44] and provided SHAP summary and feature importance plots with parameters that are most commonly chosen among 100 replications.

In this study, together with classification AUROC values, we offered plots that were prepared with cross-validated XGBoost parameters, which depicted SHAP values of each feature used for a specific clinical response classification, and also visualized how well a feature served to distinguish this specific response.

## Performance Metric

For the binary classification tasks in this study, as the performance metrics, we used AUROC, whose larger values correspond to better classification performance. We reported the AUROC values resulting from 100 independent replications in detail for each of the abovementioned classification tasks.

## Results

### Patient Characteristics

After screening, 152 patients out of 535 patients with acromegaly were included in the final analysis. All of the patients had undergone TSS and had at least 12 months of follow-up. 61% of the patients were female. 40% (n = 61) achieved remission at postoperative 3 months and 84.3% of the patients (n = 128) had normal IGF1 levels at last visit. 96% (n = 59) of the patients with remission at postoperative 3 months were in remission at the last visit. Among 91 patients without remission at postoperative 3 months, 38 patients (41%) were resistant to SRL.

Eighty-seven-point 5% (n = 133) of the patients had macroadenoma. Among them, 86 (67.7%) had invasive tumors. 15% of the patients with microadenoma (n = 3) had invasive tumors, which was lower than patients with macroadenoma ($p < 0.01$).

The clinical features of the patients were presented in Table 1.

**Table 1** Clinical, laboratory, and pathological features of the 152 patients with acromegaly

| Feature | | |
|---|---|---|
| Age at acromegaly diagnosis, mean ± SD | 40.9 | ± 10.3 |
| Gender, male n (%) | 59 | (38.8) |
| Preoperative IGF1/ULN, median (IQR) | 2.18 | (1.63) |
| Preoperative invasion*, n (%) | 89 | (59.6) |
| Cavernous sinus invasion, n (%) | 73 | (48) |
| Max. tumor diameter (mm), mean ± SD | 19.8 | ± 8.1 |
| Preoperative hypopituitarism, n (%) | 30 | (19.7) |
| Preoperative SRL use, n | 52 | (34.2) |
| Preoperative CBG use, n | 1 | (0.7) |
| Preoperative SRL + CBG use, n | 2 | (1.3) |
| Endoscopic TSS | 132 | (87) |
| Microscopic TSS | 20 | (13) |
| Pathology, n (%) | | |
| *GH secreting adenoma* | 113 | (72.3) |
| *GH + PRL secreting adenoma* | 39 | (27.7) |
| Ki-67 (%), median (IQR) | 1.5 | (1.5) |
| Sparsely granulated adenoma | 34 | (22.4) |
| Early remission, n (%) | 61 | (40.1) |
| Repeat TSS | 19 | (12.5) |
| Postoperative medical therapy, n (%) | | |
| *SRL* | 91 | (64.2) |
| *CBG* | 34 | (22.3) |
| *Peg.* | 20 | (13.2) |
| Radiotherapy, n (%) | | |
| Gamma Knife Radiosurgery | 12 | (7.9) |
| Cyberknife Radiosurgery | 9 | (5.9) |
| Conventional radiotherapy | 1 | (0.7) |
| SRL resistance, n (%) | 38 | (41) |
| Postoperative hypopituitarism, n (%) | 44 | (28.9) |
| Total disease duration, median (IQR) | 93 | (48) |
| Last status, n (%) | | |
| *Active* | 24 | (15.8) |
| *Remission with medication* | 78 | (51.3) |
| *Remission without medication* | 49 | (32.2) |
| Diabetes Mellitus, n (%) | 38 | (25) |
| Hypertension, n (%) | 26 | (17.1) |
| Obstructive sleep apnea, n (%) | 9 | (5.9) |
| Thyroid cancer, n (%) | 5 | (3.3) |
| Alive, n (%) | 152 | (100) |

*Invasion to surrounding structures such as cavernous sinus, basal dura, clivus and diaphragma sella

Data were analyzed using the R programming language.

SD *Standard deviation*, IQR *Interquartile range*, GH *Growth hormone*, IGF1 *Insulin-like growth factor 1*, ULN *Upper limit of normal*, TSS *Transsphenoidal surgery*, SRL *somatostatin receptor ligands*, CBG *Cabergoline*, Peg. *Pegvisomant*

### Classification Using Extreme Gradient Boosting

In this study, we performed three classification tasks on remission at postoperative 3 months, remission at the last visit, and resistance to SRL in patients with acromegaly. The mean AUROC values resulting from 100 independent replications are 0.728 for postoperative 3 months remission status classification, 0.879 for remission at the last visit classification, and 0.753 for SRL resistance status classification.

The mean values with lower and upper 95% confidence intervals (CI) of the 100 independent replications we performed for the outcomes were displayed in Table 2.

### Determining Feature Importance using SHAP Values

SHAP values are helpful to determine the importance of features that are used to predict a specific outcome. Since in this study we used the XGBoost, we calculated the SHAP values related to XGBoost model, and then we plotted these values for better visualization of each feature's importance in performing the corresponding classification task. SHAP feature importance plots corresponding to remission at postoperative 3 months, remission at the last visit and resistance to SRL were demonstrated at Figs. 1 and 2, and 3 respectively. Note that these features were ranked from top to bottom according to their importance level for this specific classification task.

A low serum GH level at diagnosis value was favorable for attaining remission at postoperative 3 months, as shown in Fig. 1. Also, when the age at operation was higher, the preoperative maximal tumor size and ki-67 index were lower; the chance of remission at postoperative 3 months was higher. Moreover, if a patient had a mammosomatotroph adenoma, the possibility of remission at postoperative 3 months increased.

**Table 2** The mean values with lower and upper 95% CI of the 100 independent replications we performed for the outcomes

| | Remission at last visit | Remission at postoperative 3 months | SRL Resistance |
|---|---|---|---|
| Mean AUROC | 0.8793 | 0.7278 | 0.7532 |
| Lower 95% CI | 0.8627 | 0.7083 | 0.7324 |
| Upper 95% CI | 0.8959 | 0.7473 | 0.7740 |

CI, *Confidence intervals;* AUROC, *Under the receiver operating characteristics;* SRL, *Somatostatin receptor ligands*

As indicated in Fig. 2, resistance to SRL was the most important feature to predict the remission at the last visit followed by preoperative maximal tumor diameter. Postoperative SRL use, gender, plurihormonal pituitary adenoma, invasive adenoma were the remaining important features, in their respective order. The figure suggests that if resistance to SRL was absent, the chance of having remission at the last visit was higher. Also, with increasing maximal tumor diameter, the chance of remission at the last visit was decreasing. Moreover, the need for postoperative SRL use and female gender were negative predictors of remission at the last visit. Note that the features with a "zero" value in the figure were not affecting this specific classification task.

Figure 3 denotes that when IGF1 to the upper limit of normal (age- and sex-adjusted) ratio and GH at postoperative 3 months were higher, the patient was more likely to have resistance to SRL. Also, if the adenoma was sparsely granulated somatotroph adenoma, the possibility of SRL resistance was higher.

Fig. 2

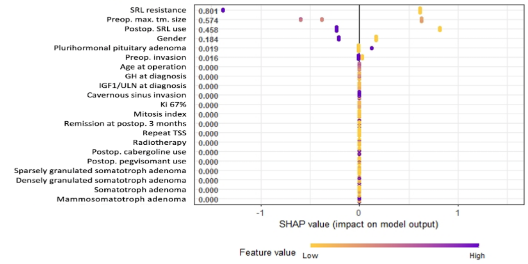


**Fig. 2** SHapley Additive exPlanations (SHAP) values for predictors of remission at postoperative 3 months (2), and resistance to somatostatin receptor ligands (SRL) GH growth hormone, IGF1 insulin-like growth factor 1, ULN upper limit of normal, TSS transsphenoidal surgery, Preop. preoperative, Postop. postoperative, Max. maximum, OGTT oral glucose tolerance test

For easy usage of our machine learning implementation and for better and simple visualization of our work, we prepared a user interface that can be found at https://midas.ku.edu.tr/Acromegaly/. Through this website, clinicians can easily test their patients' remission and SRL response.

## Discussion

The present study was prompted by a need for a more effective and widely applicable preoperative prediction method for the outcome of the patients with acromegaly. Accurate prediction of early and long-term remission can identify patients who might benefit from preoperative adjuvant medical therapy, primary medical therapy or require

Fig. 1

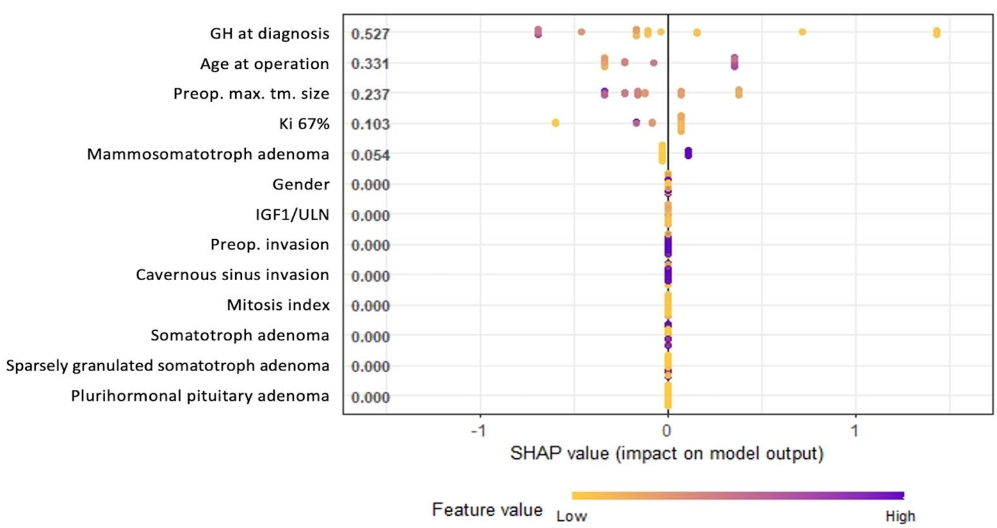


**Fig. 1** SHapley Additive exPlanations (SHAP) values for predictors of remission at postoperative 3 months (1), remission at last visit GH growth hormone, IGF1 insulin-like growth factor 1, ULN upper limit of normal, TSS transsphenoidal surgery, Preop. preoperative, Postop. postoperative, Max. maximum, OGTT oral glucose tolerance test

Fig. 3

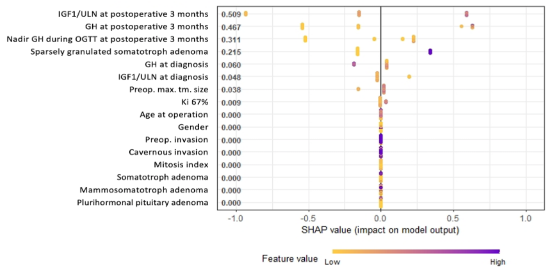


**Fig. 3** SHapley Additive exPlanations (SHAP) values for predictors of remission at postoperative 3 months (3). The features are ranked based on the permutation importance method in the XGBoost model, according to the sum of the SHAP values for all patients. The SHAP values are used to show the distribution of the effect of each feature on the XGBoost model outputs. Every dot in the figure denotes a patient. Negative SHAP value means a negative contribution on predicting the composite outcome studied, positive SHAP value means a positive contribution, and zero means no contribution. Purple designates that the value of a feature is high, and yellow designates that the value of a feature is low. Among dichotomous variables female gender, SRL resistance, postoperative SRL use, and sparsely granulated adenoma, preoperative invasion, plurihormonal pituitary adenoma and mammosomatotroph adenoma were coded as 1; and their counterparts were coded as 0 GH growth hormone, IGF1 insulin-like growth factor 1, ULN upper limit of normal, TSS transsphenoidal surgery, Preop. preoperative, Postop. postoperative, Max. maximum, OGTT oral glucose tolerance test

multimodality treatment. In this regard, we applied an ML algorithm (i.e., XGBoost) the features that are clinically related to postoperative 3 months remission, remission at last visit, and SRL resistance. We demonstrated the high performance of this novel approach in all of the studied outcomes in both training and test data sets.

The variables for postoperative 3 months remission and remission at the last visit chosen in this study included all the available information. The predictive factors for early remission determined by the ML algorithm were preoperative GH levels, age at operation, maximal tumor diameter, Ki-67 expression, and the presence of prolactin co-staining in decreasing order of importance. The top three features, namely low preoperative GH levels, older age, and small tumor maximal diameter as predictors of early remission were consistent with previous reports based on conventional statistical methods [22, 24, 45, 46]. The importance of these three features as prognostic factors also concurred with recent ML algorithm surveys conducted on patients with acromegaly [34–36]. The previous conventional studies and the above-mentioned ML surveys proposed that cavernous sinus invasion represents one of the important factors affecting surgical remission [34–36, 46]. Our model has failed to show the cavernous sinus invasion as a predictive factor for early remission. This might be a reflection of a methodological drawback of our study, as we determined the presence of cavernous sinus invasion based on preoperative MRI findings as a binary category instead of Knosp grading. This might have hampered the precise assessment.

Our ML model has revealed that a low ki-67 index is a predictor of remission at postoperative 3 months. In some studies where remission and non-remission groups were compared, the mean ki-67 index was lower in patients with acromegaly in remission [47, 48]. On the other hand, one retrospective multivariate analysis from our tertiary center did not demonstrate the relation of ki-67 with postoperative remission [8]. In another ML study where the ki-67 index was evaluated along with other prognostic factors for prediction of remission within 6 months of surgery, ki-67 index only showed a significant relationship with delayed remission in the training dataset, but there was no statistical difference in the test dataset [35]. The conflicting results related to the effect of ki-67 index on acromegaly remission might be due to varying cut-offs among studies (3%, 5%, etc.) [35, 47]. We assessed the effect of ki-67 as a continuous parameter rather than arbitrary cut-offs and the XGboost algorithm suggested that as the ki-67 index increased, the possibility of remission at postoperative 3 months decreased. Further large-scale analyses are warranted to investigate the effect of the ki-67 index on early remission.

In our ML model analysis of early remission, another significant feature was prolactin staining of GH secreting adenoma. The effect of prolactin staining of GH secreting adenoma on outcome remains elusive. Some studies reported lower, whereas others reported similar or higher remission rates in comparison to pure GH secreting adenoma [21, 49, 50]. In studies where dual staining was proposed to portend aggressiveness, the pathogenic mechanism was not entirely clear. This also holds true for our model, as we observed that mammosomatotroph adenomas were more likely to be in remission at postoperative 3 months, yet our analysis did not allow us to draw definite conclusions regarding the exact mechanism. To our knowledge, our study is the first ML-based report assessing the effect of prolactin co-staining in patients with acromegaly. Studies relying upon genetic analysis can provide insights into the biological behavior of these unique type of adenoma.

In the current study, we also sought to identify the predictors of remission at last visit, irrespective of whether postoperative treatments were administered or not. Feature importance explanations showed that resistance to SRL, increased tumor size (i.e., maximal tumor diameter), postoperative SRL use, and female gender were the top-tier variables for predicting non-remission. A large body of evidence concerning the effect of tumor size and gender on IGF1 levels has emerged [51, 52], but the effect of the need for SRL in the postoperative period and SRL resistance on the outcome is an evolving concept. Despite its clinical significance, the data assessing resistance to SRL as a discrete measure of long-term outcome is scarce and has never been evaluated employing ML-based algorithms. Our ML-based predictive model indicated that resistance to SRL is the most important factor affecting remission status in the long term. In our study population, all of the patients with acromegaly have undergone TSS by two senior neurosurgeons of Pituitary Center. In the postoperative period, despite the availability and wisely application of other treatment options as indicated, SRL resistance hold its significance on remission status at the last visit. Previous studies on determinants of SRL responsiveness demonstrated that besides radiological, clinical, and biochemical predictors, various histopathological and molecular features of GH secreting adenoma may have a role in SRL resistance [53–55]. Taken together, we propose that resistance to SRL has implications beyond simple therapeutic non-responsive to one class of drug. At this point, it becomes imperative to identify the predictors of SRL resistance with the support of ML tools.

To date, several predictors of SRL resistance have been proposed. IGF1 and GH levels at diagnosis, densely granulated adenoma, tumor size, ki-67 expression, gender, and age have been purported to predict SRL response via conventional methods. Except for the latter two, these parameters were confirmed by our ML model as predictors of SRL resistance. Recently, Gadelha et al. that targeted to develop a ML-based model to identify predictors of therapeutic response to SRL in patients with acromegaly [56]. They determined granulation pattern, pretreatment IGF1 and GH levels, age, sex and somatostatin receptor subtype 2 and 5 expression as predictors of therapeutic response to SRL. The granulation pattern, pretreatment IGF1 and GH levels also coincided with our registry. Moreover, we also demonstrated that IGF1 and GH levels (random and nadir) at postoperative 3 months were also important determinants of therapeutic response to SRL. It has been shown that percent reduction of growth hormone levels correlates closely with percent resected tumor volume in acromegaly [57]. This is suggestive that low IGF1 and GH levels (random and nadir) at postoperative 3 months are indeed a measure of successful surgical resection, which in turn has been shown to improve the response to SRL [58, 59]. Thus, IGF1 and GH levels (random and nadir) at postoperative 3 months as a predictor of SRL response was indirectly supported by previous studies. Our observation further expanded the spectrum of the predictors of SRL resistance but still requires verification by larger-scale ML models.

ML is a novel approach with enhanced predictive power compared to conventional methods and thus becoming increasingly used in the field of neuroendocrinology [32–36, 60]. Recently, three separate groups used a similar ML approach to predict surgical remission of patients with acromegaly in particular, but none of them assessed predictors of remission status in long-term and/or resistance to SRL [34–36]. We expanded our ML-based analysis with remission at the last visit, together with the predictive factors of SRL resistance. In our study, the XGBoost algorithms yielded a good performance overall for all the studied outcomes, as indicated by mean AUROC values. This is supported by other studies conducted on patients with acromegaly, as one of them ranked XGBoost as the best, and another indicated that it was one of the best predictive models for surgical remission among others [34, 35]. In our analyses, the XGBoost model demonstrated that preoperative GH level, age at operation, and preoperative tumor size were the most important predictors for early remission; resistance to SRL and preoperative tumor size represented the most important predictors of remission at last visit; postoperative 3-month IGF1 and GH levels (random and nadir) together with the sparsely granulated somatotroph adenoma served as the most important predictors of SRL resistance. These results were in accordance with clinical cognition and practice, further verifying the reliability of the XGboost model. Despite its favorably good performance, ML models have caveats that need to be taken into account. First ML is regarded as a "black box". This means that it is devoid of a transparent interpretation of the learning process or the outputs, and the function between the predictive factors and the response is invisible to the researcher [61]. We need to set out the reasons for the ML models to make such predictions in clinical settings. Therefore, we utilized SHAP, which is a conceptual agnostic interpretation method, to explain our prediction models. In the past, researchers used partial dependence plots or feature importance to explain the ML models. These methods show the contribution made by their features to the predictive ability of the model but do not allow us to delineate whether the influence of these features on the output is positive or negative. In 2017, Lundberg et al. presented SHAP, which is a game-theoretic approach that assigns each feature an importance value for a particular prediction to explain a complex black box model [62]. It has two vital advantages over conventional methods. The first one is that in addition to the influence of a single feature, it also considers the synergy between features and thus handles the problem of multicollinearity [35]. Second, it determines whether the influence of a single feature is positive or negative. Since it helps to interpret the output of the ML models, it is promising for ML applications on medical data sets.

The present study has limitations. This was a single-center study involving limited number of patients. The model may behave differently with more data and patients from multiple sources. Although we use four-fold cross-validation to improve our analyses, there is an inevitable need for large-scale designs for the validation of robustness and reproducibility of the model.

On the other hand, the ML model constructed in this study has advantages. Introduction of artificial intelligence to predict outcome and to tailor treatment decisions of patients with acromegaly is a recently elaborating topic [56]. This study represents one of the initial efforts to use ML for predicting therapeutic response to SRL in addition to early surgical remission and remission at last visit. In this context, it provides a wider range of knowledge that can guide clinical decisions.

## Conclusions

It is feasible to develop ML-based models to predict early and long-term remission in patients with acromegaly. This subset of artificial intelligence can serve as an effective non-invasive method to foresee outcomes and aid in clinical decisions. It may provide a new scope of information compatible with clinical cognition and widen our perspective concerning prognostic features.

**Author contributions** All authors made substantial contributions to conception and design, and/or acquisition of data, and/or analysis and interpretation of data; participated in drafting the article or revising it critically for important intellectual content; gave final approval of the version to be submitted.

### Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Disclosure Statement** The authors have nothing to disclose. Cem Sulu and Ayyüce Begüm Bektaş have equally contributed to this article.

**Ethical approval** All procedures performed in studies involving human participants were in accordance with ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent** Informed consent was obtained from all living individual participants included in the study.

## References


1. Melmed S (2009) Acromegaly pathogenesis and treatment. J Clin Invest 119(11):3189–3202 Epub 2009/11/04. doi: https://doi.org/10.1172/jci39375. PubMed PMID: 19884662; PubMed Central PMCID: PMCPMC2769196.
2. Colao A, Ferone D, Marzullo P, Lombardi G (2004) Systemic complications of acromegaly: epidemiology, pathogenesis, and management. Endocr Rev 25(1):102–152
3. Sherlock M, Ayuk J, Tomlinson JW, Toogood AA, Aragon-Alonso A, Sheppard MC et al (2010) Mortality in patients with pituitary disease. Endocr Rev 31(3):301–342
4. Renehan AG, O'Connell J, O'Halloran D, Shanahan F, Potten CS, O'Dwyer ST et al (2003) Acromegaly and colorectal cancer: a comprehensive review of epidemiology, biological mechanisms, and clinical implications. Horm Metab Res 35(11/12):712–725
5. Babu H, Ortega A, Nuno M, Dehghan A, Schweitzer A, Bonert HV et al (2017) Long-term endocrine outcomes following endoscopic endonasal transsphenoidal surgery for acromegaly and associated prognostic factors. Neurosurgery 81(2):357–366
6. Shimon I, Cohen ZR, Ram Z, Hadani M (2001) Transsphenoidal surgery for acromegaly: endocrinological follow-up of 98 patients. Neurosurgery 48(6):1239–1245
7. Aydin S, Ozoner B, Sahin S, Alizada O, Comunoglu N, Oz B et al (2020) A follow-up study on outcomes of endoscopic transsphenoidal approach for acromegaly. Clin Neurol Neurosurg 198:106201 Epub 2020/09/15. doi: https://doi.org/10.1016/j.clineuro.2020.106201. PubMed PMID: 32927330.
8. Haliloglu O, Kuruoglu E, Ozkaya HM, Keskin FE, Gunaldi O, Oz B et al (2016) Multidisciplinary Approach for Acromegaly: A Single Tertiary Center's Experience. World Neurosurg 88:270–276 Epub 2016/01/26. doi: https://doi.org/10.1016/j.wneu.2015.12.092. PubMed PMID: 26806060.
9. Briceno V, Zaidi HA, Doucette JA, Onomichi KB, Alreshidi A, Mekary RA et al (2017) Efficacy of transsphenoidal surgery in achieving biochemical cure of growth hormone-secreting pituitary adenomas among patients with cavernous sinus invasion: a systematic review and meta-analysis. Neurol Res 39(5):387–398
10. Jane JA Jr, Starke RM, Elzoghby MA, Reames DL, Payne SC, Thorner MO et al (2011) Endoscopic transsphenoidal surgery for acromegaly: remission using modern criteria, complications, and predictors of outcome. J Clin Endocrinol Metabolism 96(9):2732–2740
11. Sun H, Brzana J, Yedinak CG, Gultekin SH, Delashaw JB, Fleseriu M (2014) Factors associated with biochemical remission after microscopic transsphenoidal surgery for acromegaly. J Neurol Surg Part B: Skull Base 75(01):047–52
12. Petrossians P, Borges-Martins L, Espinoza C, Daly A, Betea D, Valdes-Socin H et al (2005) Gross total resection or debulking of pituitary adenomas improves hormonal control of acromegaly by somatostatin analogs. Eur J Endocrinol 152(1):61–66
13. Kiseljak-Vassiliades K, Carlson NE, Borges MT, Kleinschmidt-DeMasters B, Lillehei KO, Kerr JM et al (2015) Growth hormone tumor histological subtypes predict response to surgical and medical therapy. Endocrine 49(1):231–241
14. Heck A, Ringstad G, Fougner SL, Casar-Borota O, Nome T, Ramm-Pettersen J et al (2012) Intensity of pituitary adenoma on T2-weighted magnetic resonance imaging predicts the response to octreotide treatment in newly diagnosed acromegaly. Clin Endocrinol 77(1):72–78
15. Nomikos P, Buchfelder M, Fahlbusch R (2005) The outcome of surgery in 668 patients with acromegaly using current criteria of biochemical 'cure'. Eur J Endocrinol 152(3):379–387
16. Ahmed S, Elsheikh M, Stratton I, Page R, Adams C, Wass J (1999) Outcome of transphenoidal surgery for acromegaly and its relationship to surgical experience. Clin Endocrinol 50(5):561–567
17. Shen M, Zhang Q, Liu W, Wang M, Zhu J, Ma Z et al (2016) Predictive value of T2 relative signal intensity for response to somatostatin analogs in newly diagnosed acromegaly. Neuroradiology 58(11):1057–1065
18. van Bunderen CC, van Varsseveld NC, Baayen JC, van Furth WR, Aliaga ES, Hazewinkel MJ et al (2013) Predictors of endoscopic transsphenoidal surgery outcome in acromegaly: patient and tumor characteristics evaluated by magnetic resonance imaging. Pituitary 16(2):158–167

19. Mazal P, Czech T, Sedivy R, Aichholzer M, Wanschitz J, Klupp N et al (2001) Prognostic relevance of intracytoplasmic cytokeratin pattern, hormone expression profile, and cell proliferation in pituitary adenomas of akromegalic patients. Clin Neuropathol 20(4):163–171
20. Bakhtiar Y, Hirano H, Arita K, Yunoue S, Fujio S, Tominaga A et al (2010) Relationship between cytokeratin staining patterns and clinico-pathological features in somatotropinomae. Eur J Endocrinol 163(4):531
21. Bourdelot A, Coste J, Hazebroucq V, Gaillard S, Cazabat L, Bertagna X et al (2004) Clinical, hormonal and magnetic resonance imaging (MRI) predictors of transsphenoidal surgery outcome in acromegaly. Eur J Endocrinol 150(6):763–771
22. Starke RM, Raper DM, Payne SC, Vance ML, Oldfield EH, Jane JA Jr (2013) Endoscopic vs microsurgical transsphenoidal surgery for acromegaly: outcomes in a concurrent series of patients using modern criteria for remission. J Clin Endocrinol Metabolism 98(8):3190–3198
23. Antunes X, Ventura N, Camilo GB, Wildemberg LE, Guasti A, Pereira PJM et al (2018) Predictors of surgical outcome and early criteria of remission in acromegaly. Endocrine 60(3):415–422
24. Abosch A, Tyrrell JB, Lamborn KR, Hannegan LT, Applebury CB, Wilson CB (1998) Transsphenoidal microsurgery for growth hormone-secreting pituitary adenomas: initial outcome and long-term results. J Clin Endocrinol Metabolism 83(10):3411–3418
25. Shirvani M, Motiei-Langroudi R (2014) Transsphenoidal surgery for growth hormone–secreting pituitary adenomas in 130 patients. World Neurosurg 81(1):125–130
26. Kovacs K, Horvath E (1986) Pathology of growth hormone-producing tumors of the human pituitary. Semin Diagn Pathol 3(1):18–33 Epub 1986/02/01. PubMed PMID: 3303228
27. Krieger MD, Couldwell WT, Weiss MH (2003) Assessment of long-term remission of acromegaly following surgery. J Neurosurg 98(4):719–724
28. Dutta P, Korbonits M, Sachdeva N, Gupta P, Srinivasan A, Devgun JS et al (2016) Can immediate postoperative random growth hormone levels predict long-term cure in patients with acromegaly? Neurology India 64(2):252–8
29. Feelders R, Bidlingmaier M, Strasburger C, Janssen J, Uitterlinden P, Hofland L et al (2005) Postoperative evaluation of patients with acromegaly: clinical significance and timing of oral glucose tolerance testing and measurement of (free) insulin-like growth factor I, acid-labile subunit, and growth hormone-binding protein levels. J Clin Endocrinol Metabolism 90(12):6480–6489
30. Hazer DB, Işık S, Berker D, Güler S, Gürlek A, Yücel T et al (2013) Treatment of acromegaly by endoscopic transsphenoidal surgery: surgical experience in 214 cases and cure rates according to current consensus criteria. J Neurosurg 119(6):1467–1477
31. Cleophas TJ (2016) Machine learning in therapeutic research: the hard work of outlier detection in large data. Am J Ther 23(3):e837–e43
32. Hollon TC, Parikh A, Pandian B, Tarpeh J, Orringer DA, Barkan AL et al (2018) A machine learning approach to predict early outcomes after pituitary adenoma surgery. NeuroSurg Focus 45(5):E8
33. Staartjes VE, Serra C, Muscas G, Maldaner N, Akeret K, van Niftrik CH et al (2018) Utility of deep neural networks in predicting gross-total resection after transsphenoidal surgery for pituitary adenoma: a pilot study. NeuroSurg Focus 45(5):E12
34. Fan Y, Li Y, Li Y, Feng S, Bao X, Feng M et al (2020) Development and assessment of machine learning algorithms for predicting remission after transsphenoidal surgery among patients with acromegaly. Endocrine 67(2):412–422
35. Dai C, Fan Y, Li Y, Bao X, Li Y, Su M et al (2020) Development and interpretation of multiple machine learning models for predicting postoperative delayed remission of acromegaly patients during long-term follow-up.Frontiers in endocrinology. 2020;11:643.
36. Qiao N, Shen M, He W, He M, Zhang Z, Ye H et al (2021) Machine learning in predicting early remission in patients after surgical treatment of acromegaly: a multicenter study. Pituitary 24(1):53–61
37. Katznelson L, Laws ER Jr, Melmed S, Molitch ME, Murad MH, Utz A et al (2014) Acromegaly: an endocrine society clinical practice guideline. J Clin Endocrinol Metabolism 99(11):3933–3951
38. Katznelson L, Laws ER Jr, Melmed S, Molitch ME, Murad MH, Utz A et al (2014) Acromegaly: An Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metabolism 99(11):3933–3951. doi: https://doi.org/10.1210/jc.2014-2700
39. Lombardi G, Colao A, Marzullo P, Biondi B, Palmieri E, Fazio S (2002) Improvement of left ventricular hypertrophy and arrhythmias after lanreotide-induced GH and IGF-I decrease in acromegaly. A prospective multi-center study. J Endocrinol Invest 25(11):971–976 Epub 2003/01/30. doi: https://doi.org/10.1007/bf03344070. PubMed PMID: 12553557.
40. Polat Korkmaz O, Gurcan M, Nuhoglu Kantarci FE, Haliloglu O, Ozkaya HM, Sahin S et al (2019) The effects of pre-operative somatostatin analogue therapy on treatment cost and remission in acromegaly. Pituitary 22(4):387–396 Epub 2019/05/18. doi: https://doi.org/10.1007/s11102-019-00968-6. PubMed PMID: 31098838.
41. Giustina A, Chanson P, Kleinberg D, Bronstein MD, Clemmons DR, Klibanski A et al (2014) Expert consensus document: a consensus on the medical treatment of acromegaly. Nat Reviews Endocrinol 10(4):243–248
42. Colao A, Auriemma RS, Lombardi G, Pivonello R (2011) Resistance to somatostatin analogs in acromegaly. Endocr Rev 32(2):247–271
43. Chen T, editor Guestrin C, Mining D (2016) (KDD'16);
44. Du M, Liu N, Hu X (2019) Techniques for interpretable machine learning. Commun ACM 63(1):68–77
45. Minniti G, Jaffrain-Rea M, Esposito V, Santoro A, Tamburrano G, Cantore G (2003) Evolving criteria for post-operative biochemical remission of acromegaly: can we achieve a definitive cure? An audit of surgical results on a large series and a review of the literature. Endocrine-related Cancer 10(4):611–619
46. Araujo-Castro M, Pascual-Corrales E, Martínez-Vaello V, Baonza Saiz G, Quiñones de Silva J, Acitores Cancela A et al (2021) Predictive model of surgical remission in acromegaly: age, presurgical GH levels and Knosp grade as the best predictors of surgical remission. J Endocrinol Investig 44:183–193
47. Asha MJ, Takami H, Velasquez C, Oswari S, Almeida JP, Zadeh G et al (2019) Long-term outcomes of transsphenoidal surgery for management of growth hormone–secreting adenomas: single-center results. J Neurosurg 133(5):1360–1370
48. Fusco A, Zatelli MC, Bianchi A, Cimino V, Tilaro L, Veltri F et al (2008) Prognostic significance of the Ki-67 labeling index in growth hormone-secreting pituitary adenomas. J Clin Endocrinol Metabolism 93(7):2746–2750
49. Wang M, Mou C, Jiang M, Han L, Fan S, Huan C et al (2012) The characteristics of acromegalic patients with hyperprolactinemia and the differences in patients with merely GH-secreting adenomas: clinical analysis of 279 cases. Eur J Endocrinol 166(5):797–802
50. Rick J, Jahangiri A, Flanigan PM, Chandra A, Kunwar S, Blevins L et al (2018) Growth hormone and prolactin-staining tumors causing acromegaly: a retrospective review of clinical presentations and surgical outcomes. J Neurosurg 131(1):147–153
51. Park SH, Ku CR, Moon JH, Kim EH, Kim SH, Lee EJ (2018) Age-and sex-specific differences as predictors of surgical remission among patients with acromegaly. J Clin Endocrinol Metabolism 103(3):909–916

52. Albarel F, Castinetti F, Morange I, Conte-Devolx B, Gaudart J, Dufour H et al (2013) Outcome of multimodal therapy in operated acromegalic patients, a study in 115 patients. Clin Endocrinol 78(2):263–270
53. Cuevas-Ramos D, Fleseriu M (2014) Somatostatin receptor ligands and resistance to treatment in pituitary adenomas. J Mol Endocrinol 52(3):R223–R40
54. Brzana J, Yedinak CG, Gultekin SH, Delashaw JB, Fleseriu M (2013) Growth hormone granulation pattern and somatostatin receptor subtype 2A correlate with postoperative somatostatin receptor ligand response in acromegaly: a large single center experience. Pituitary 16(4):490–498
55. Gadelha MR, Kasuki L, Korbonits M (2013) Novel pathway for somatostatin analogs in patients with acromegaly. Trends in Endocrinology & Metabolism 24(5):238–246
56. Wildemberg LE, da Silva Camacho AH, Miranda RL, Elias PC, de Castro Musolino NR, Nazato D et al (2021) Machine learning-based prediction model for treatment of acromegaly with first-generation somatostatin receptor ligands. J Clin Endocrinol Metabolism 106(7):2047–2056
57. Schwyzer L, Starke RM, Jane JA Jr, Oldfield EH (2015) Percent reduction of growth hormone levels correlates closely with percent resected tumor volume in acromegaly. J Neurosurg 122(4):798–802 Epub 2014/11/26. doi: https://doi.org/10.3171/2014.10.Jns14496. PubMed PMID: 25423276.
58. Melmed S (2006) Acromegaly. New England. J Med 355(24):2558–2573
59. Ben-Shlomo A, Schmid H, Wawrowsky K, Pichurin O, Hubina E, Chesnokova V et al (2009) Differential ligand-mediated pituitary somatostatin receptor subtype signaling: implications for corticotroph tumor therapy. J Clin Endocrinol Metabolism 94(11):4342–4350
60. Liu Y, Liu X, Hong X, Liu P, Bao X, Yao Y et al (2019) Prediction of recurrence after transsphenoidal surgery for Cushing's disease: the use of machine learning algorithms. Neuroendocrinology 108(3):201–210
61. Wang X, Wang D, Yao Z, Xin B, Wang B, Lan C et al (2019) Machine learning models for multiparametric glioma grading with quantitative result interpretations. Front NeuroSci 12:1046
62. Lundberg SM, Lee S-I (eds) (2017) Proceedings of the 31st international conference on neural information processing systems;