Acta Neurochirurgica (2025) 167:61
https://doi.org/10.1007/s00701-025-06477-9

RESEARCH

Is Knosp enough? A novel classification for Acromegaly: a retrospective analysis of cure rates and outcome predictors in a large tertiary centre

Feras Fayez[1] · Ahmed Abougamil[1] · Francesca Vitulli[2] · James Knight[1] · Christoforos Syrris[1] · Oktay Genel[1] · Jonathan Shapey[1] · Eleni Maratos[1] · Nicholas Thomas[1] · Sinan Barazi[1]

Received: 23 June 2024 / Accepted: 24 February 2025

Abstract

Purpose This study evaluates surgical outcomes for acromegaly at King's College Hospital(2012–2022), focusing on predictive factors for surgical cure. A novel radiological metric, the MI Ratio, is introduced to enhance the prediction of clinical remission post-surgery, providing a more accurate prognosis and informing treatment planning.

Methods This single-centre cohort study involved a retrospective analysis of prospectively collected data from a UK tertiary referral centre. Included were patients with histologically proven somatotroph tumours who underwent endoscopic trans-sphenoidal surgery (TSS) between 2012 and 2022. Exclusions were made for incomplete data or lost follow-up. Patient demographics, tumour characteristics, radiological parameters, and biochemical markers were analysed. The MI Ratio was defined as the distance from the midline to the lateral maximum of a tumour, divided by the distance between the two cavernous carotid arteries on coronal MRI.

Results Out of 157 patients, 150 met the inclusion criteria. Using the 2018 consensus OGTT nadir < 0.40 ng/mL, microadenomas had a higher surgical cure rate (72%) compared to macroadenomas (48%), with an overall cure rate of 53%. Significant predictors of surgical cure included the MI Ratio ($p < 0.001$), microadenomas ($p = 0.022$), Knosp score < 2 ($p = 0.012$), immediate post-operative GH level ($p = 0.016$), and patient gender ($p = 0.005$). Pre-operative medical management did not significantly impact surgical remission ($p = 0.19$), while pre-operative GH level approached significance ($p = 0.06$). CV between operators for MI was < 5% indicating minimal Interoperator variability.

Conclusions This study is the first to describe the MI Ratio, demonstrating its utility in predicting surgical remission in acromegaly patients. A combination of radiological features, demographics, and hormone profiles can more accurately identify patients less likely to achieve surgical cure.

Keywords Trans-sphenoidal surgery · Acromegaly · Radiological classification

Introduction

Acromegaly is a neuroendocrinological disease typically caused by functioning pituitary adenomas, which are monoclonal expansions of differentiated somatotroph-secreting cells of the pituitary gland [24, 27, 35]. This results in an excess production of growth hormone (GH) which subsequently leads to increased Insulin-like Growth factor-1 (IGF-1) production and its end-organ effects [17, 21, 24, 27, 28]. The disease has been found to impact both genders and is commonly diagnosed between the 5th and 6th decade [41]. Worldwide, the incidence of acromegaly is estimated to be 3–11 new cases per million per year with a prevalence of 28–137 cases per million population [22]. Acromegaly may

Feras Fayez and Ahmed Abougamil contributed equally to this work.

Previous Presentations Presented at the SBNS Autumn 2023 meeting. Awarded skull base session prize for the meeting and the Mr. Michael Powell Pituitary Prize for the year 2023.

✉ Ahmed Abougamil
dr.ahmadbahy@gmail.com

[1] Department of Neurosurgery, King's College Hospital Foundation Trust, London, UK

[2] Department of Neurological Sciences, Division of Neurosurgery, Università Degli Studi Di Napoli Federico II, Naples, Italy

Published online: 08 March 2025

result in a variety of cardiovascular, respiratory, endocrine, metabolic, musculoskeletal, and neoplastic comorbidities as well as symptoms of local mass effect from compression of the optic apparatus [1]. Rarely these tumours can present as pituitary apoplexy or be found incidentally [2]. The clinical features of acromegaly often develop insidiously [10]. This commonly leads to a delayed diagnosis of the disease of up to 5–10 years that may result in reduced quality of life (QoL) and elevated patient morbidity and mortality [38].

The primary goal of treatment is to eliminate GH excess, and surgical resection of the adenoma has traditionally been the mainstay of treatment [7, 8, 42]. If there is a clear surgical target on MRI, Surgery, usually via an endoscopic transsphenoidal approach, is the first line treatment with the goal of suppressing GH levels [6, 7, 19, 24, 27, 28]. Depending on Octreotide test dose–response and patients' ability to tolerate side effects [17, 24, 28, 41], medical therapy with somatostatin analogues (SSA) and dopamine agonists may also be used [9, 13, 36] either as a primary treatment or as a neo-adjuvant therapy prior to surgery [33]. To assess cure, accepted guidelines [10, 14, 15, 26, 28, 41] recommend that GH levels are measured after an oral glucose load [26], typically 3–6 months post-operatively [33] and following a washout period for those on SSA. Adjusted IGF-1 levels for age and gender are used in some settings instead [28, 37].

In experienced hands, initial remission rates exceeding 85% have been reported for microadenomas, while rates ranging from 40–50% have been observed for macroadenomas [6, 14, 19, 26]. Gross total resection (GTR) is an important predictor of long-term remission [7, 19, 32] and the endocrine society guidelines [24] recommend repeat surgery in those with a residual that is a clear surgical target [7, 24]. Despite this, while there have been attempts [14], it remains difficult to predict probability of GTR.

For patients with residual tumour or for those patients who declined surgery, adjuvant medical therapy [9, 23] with SSAs [18, 24], dopamine agonists or more recently pegvisomant (a GH receptor antagonist) has been recommended [16]. Radiotherapy can also be considered in patients that are not candidates for revision surgery and those with residual or recurrent disease [20, 24, 27]. However, when feasible, surgery remains the preferred treatment modality for patients with acromegaly [24, 27, 28].

The Knosp classification [25] and its revised update [5] stratify cavernous sinus invasion (Fig. 1) and have been reported as a strong predictor of resection and therefore remission [19]. Other risk factors such as GH level [6, 15, 26, 37] age and gender [15] have also been cited as negative predictors of cure. Previous studies have had limited numbers of patients and consequently lack the statistical power to develop new clinical and diagnostic methods [1, 7].

The outcome following surgery varies considerably and can range from long-term remission following a single operation to persist disease despite multimodal treatment [24]. At present we lack the tools to predict surgical cure. The current study introduces a novel radiological classification for adenomas in patients with acromegaly which can predict surgical remission. This will help to inform treatment decisions for our patients.

Methods

Study design: Retrospective cohort study.

Inclusion criteria: This study included all patients with histologically confirmed acromegaly who underwent endoscopic TSS.

Exclusion criteria: Patients who did not have adequate follow-up, defined as a minimum of three months, were excluded from the analysis.

This manuscript was conducted ethically in accordance with the World Medical Association Declaration of Helsinki [39] and was registered at our institution.

We conducted a single centre retrospective analysis of 150 patients treated at King's College Hospital between 2012 to 2022. Within this period, we screened all patients diagnosed with acromegaly at our centre by searching our electronic patient records for the keywords "Acromegaly" "Growth Hormone", "GH" and "Somatotroph" in all pituitary MDT notes within the database. Patients who underwent TSS were screened, and their data extracted into an excel file. Patients were then split into 2 cohorts: those who received pre-operative medical treatment in the form of somatostatin receptor analogue and those who did not. All patients underwent endoscopic TSS under general anaesthetic and had a growth hormone-secreting adenoma confirmed on histopathology.

Patient data including age at time of surgery, gender, pre-operative GH level and radiological features on axial MRI were recorded. MRI data included degree of cavernous sinus invasion recorded using the revised Knosp classification Additionally, a novel radiological methodology, termed the MI Ratio, was proposed, and implemented.

The following steps were used to calculate the MI ratio (as illustrated in Fig. 2):

1. Select the pre-operative Coronal T-1 weighted Contrast-enhanced MRI.
2. Identify the slice where tumour diameter is the largest.
3. Measure the distance between the two cavernous Carotid arteries at the widest point (Dashed blue line). This is < < I > > .
4. Draw a line to indicate the midline using anatomical structures such as the superior sagittal sinus and the third ventricle as well as the Vomer inferiorly (dashed

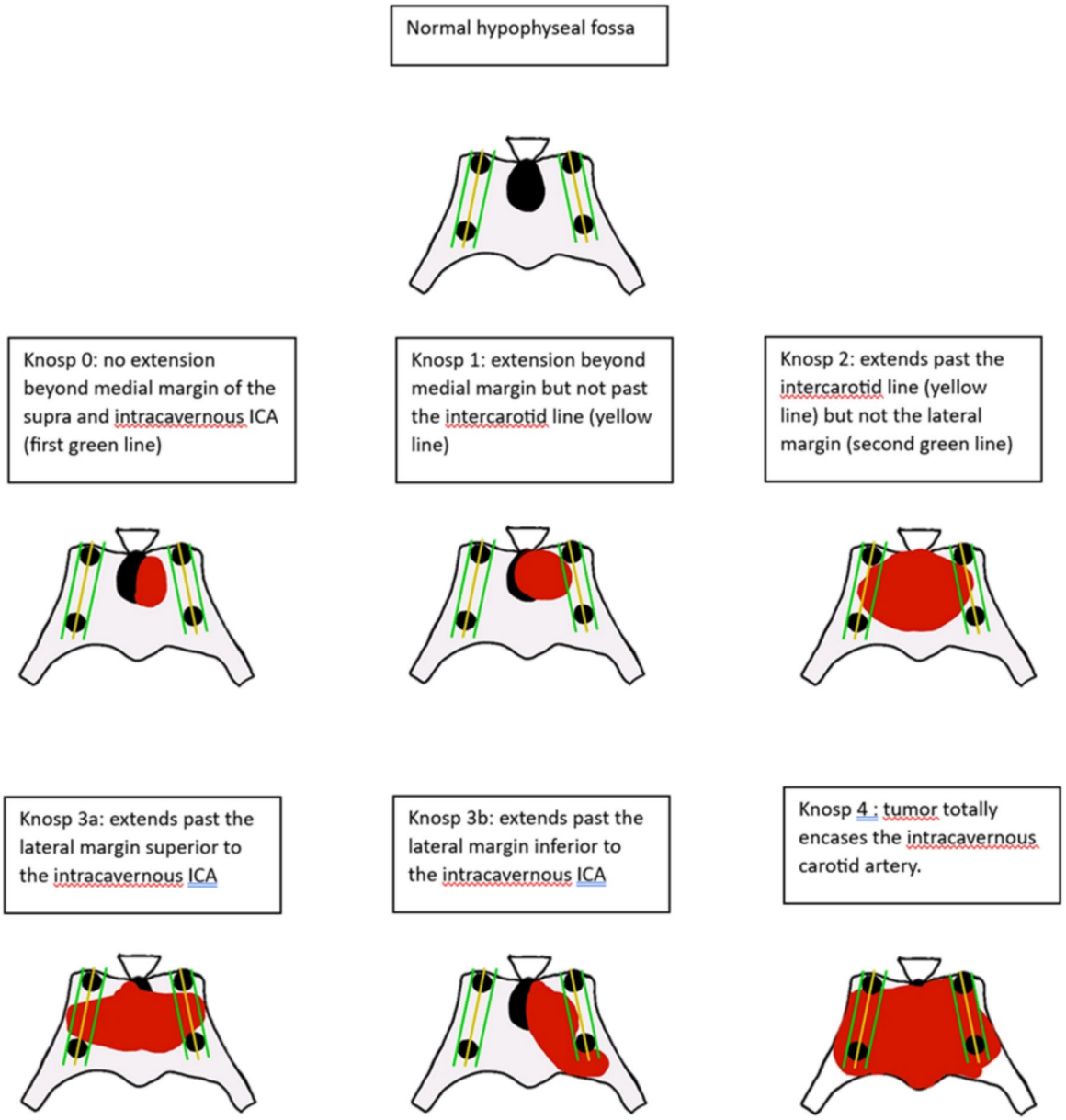

Fig. 1 The revised Knosp classification [5, 25] visualized

golden line). If there is mass effect, the middle of the two ICAs can be used as a proxy.

5. Measure the distance from the lateral maximum (light green line) of the tumour to the midline (dashed golden line). This is < < M > > .
6. Calculate the MI ratio by using these two values.

Interobserver variability

Twenty cases were randomly selected, and MI ratio was calculated by two senior operators (AA and CS) to find Interoperator variability and ensure there was no significant discrepancy. Mean, standard deviation and the coefficient of variation (CV) were calculated for each operator.

Biochemistry

Acromegaly was defined as the inability to suppress GH levels below 1.0 ng/mL after an oral glucose tolerance test (OGTT) as per the endocrine society clinical practice guidelines [24]. Overall cure rates were computed utilizing both the pituitary consensus cut-offs for postoperative nadir levels of OGTT and IGF-1 levels, as stipulated in the guidelines from 2014 [24] and 2018 [28], respectively. These criteria

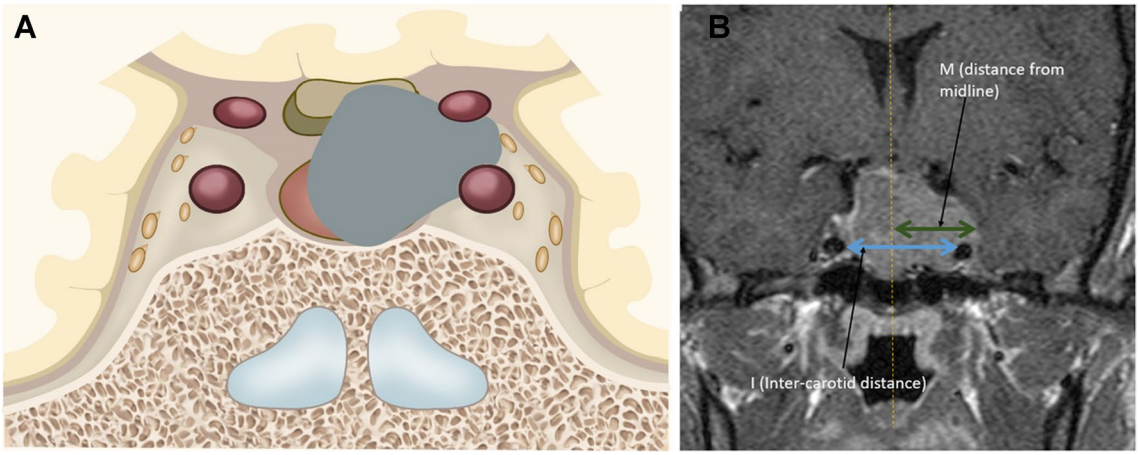

Fig. 2 **A** Coronal CT slice of and illustartion of the Hypophyseal Fossa. **B** Blue line is the intercarotid distance we call this variable I; golden line is the midline; and the dark green line is the distance between midline and lateral maxima which we call M

were applied to the initial OGTT or IGF-1 level obtained within the 3–6-month period after TSS.

A sub-group analysis (Fig. 3) was also done looking at specific populations including menopausal women [31], each gender, elderly patients and patients split by radiological parameters.

Radiological analysis

The following radiological data were recorded: Tumour size (< 10 mm microadenoma, 10 + mm macroadenoma), cavernous sinus invasion, Knosp classification [25] and our proposed classification (MI Ratio). GTR was defined as the lack of radiological residual on follow-up imaging.

Procedure

All procedures were performed by at least two senior pituitary surgeons, including a senior consultant with over 15 years of experience. All procedures were done with a bi-nostril approach using a forehand technique endoscopically primarily with a 0-degree lens and 30 or even 45-degree lenses were deployed when necessary to explore lateral masses. For high grade invasive tumours, the cavernous

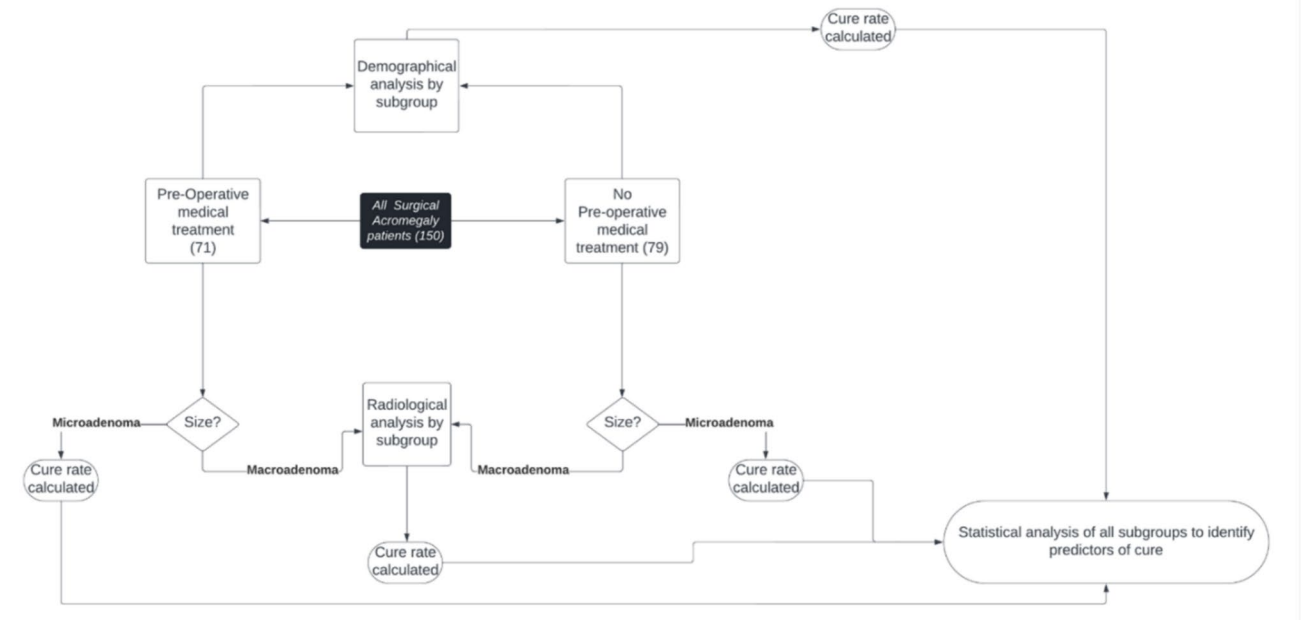

Fig. 3 Subgrouping workflow. analysis and methodology

sinus was exposed using a medial wall resection. Approaches were always midline to lateral.

Pathology

Pertinent histopathological features including Ki-67 [3] and cell type(s) were evaluated. Tumours with multiple cell types were excluded and only primary somatotroph adenomas were included in the final analysis.

Statistical analysis

The raw data was entered into Microsoft Excel. Statistical analysis was done via R (version 4.3.2; The R Foundation for Statistical Computing) and RStudio (version 2023.12.0 + 369). Standard descriptive statistics were used to describe the characteristics of cases (median with range, mean ± SD and frequencies with percentages). Univariate analysis used t-test for means and Fisher exact testing for proportions. OR are reported with 95% confidence intervals CI. p values < 0.05 were deemed statistically significant.

Subgroup analysis included both demographic splitting (menopausal status, age, gender, previous history of surgery) and radiological subgroupings (micro vs macroadenoma, Knosp classification [25] and MI score) Fig. 3 illustrates this process:

Follow-up

Initial follow up for all patients was at 3 and 6 months post-operatively and any complications were recorded. Data for patients who did not achieve remission and those who underwent repeat surgery and/or adjuvant treatment were recorded.

In patients who underwent multiple surgical procedures, the most recent procedure was used in our final analysis. If used, medical therapy (and type) and radiotherapy were noted. Long term biochemical control was recorded for all patients to assess the efficacy of the adjuvant therapies and likelihood of initial remission predicting long term cure.

Results

During the study period, 157 patients underwent endoscopic TSS for histologically confirmed acromegaly at KCH. Four patients were lost to follow up prior to OGTT/IGF-1 testing and three did not have radiological data. A total of 150 patients were thus included in the final analysis. Patient demographics are included in Table 1 below:

Cure rates

As seen in Table 1, microadenomas had a surgical cure rate of 72% and macroadenomas had a surgical cure rate of 48% using the 2018 consensus cut-off of < 0.4 ng/mL GH nadir on OGTT. This led to an overall cure rate of 53%. Using the 2014 consensus criteria our cure rates would have been 83% for microadenomas, 67% for macroadenomas and an overall 70%, i.e. there was a marked difference in cure rates observed when using the stricter 2018 consensus of < 0.40 ng/mL and the < 1.00 ng/mL cut-off recommended in the 2014 consensus.

Univariate analysis

All analysis was done using the 2018 consensus for cure of < 0.4 ng/mL on OGTT.

Factors found to be statistically significant predictors of cure (P values < 0.05) are shown in Table 1. Male gender $p = 0.005$ (OR 2.64[1.36–5.23]), microadenomas $p = 0.022$ (OR 2.80[1.18–7.27]), smaller tumour diameter $p < 0.001$(OR 0.98 per mm[0.97–0.99]), Smaller lateral extension $p < 0.001$(OR0.96 per mm [0.95–0.98]) and by proxy lower Knosp [5, 25] scores were all significant pre-operative predictors of cure.

The only significant immediate post-operative predictors of cure were radiological residual on MRI $p < 0.001$(OR 0.14[0.07–0.29]) and immediate post-operative GH level $p = 0.016$ (OR 0.96 per ng/mL [0.92–0.99]).

Univariate analysis was also performed to predict radiological residual on post-operative MRI. Higher Knosp ($p < 0.001$, OR1.97 per Knosp grade [1.46–2.66]), greater MI score ($p < 0.001$ OR 46.1 per 0.1 increase [7.30–291.78]) and cavernous sinus invasion ($p < 0.001$ OR 3.81[1.90–7.92]) were all significant pre-operative predictors of radiological residual.

Receiver operating characteristic (ROC) analysis

We performed a ROC analysis of select pre-operative variables and plotted them in Fig. 4. MI + Gender had the largest area under curve (AUC) of 0.76 followed by MI alone at 0.73 then Knosp scoring and diameter with an AUC of 0.70.

Prediction modelling and interoperator variance

MI scoring had low coefficient of variation CV (< 5%) for both operators, indicating minimal Interoperator variability. We generated two graphs (real and smoothed) for MI using both the actual results predicted by the logistic regression (Fig. 5a) and calibrating results to cure or not cure depending

Table 1 Basic descriptors, univariate statistical significance, and odds ratios (OR) with confidence intervals for patients who achieved surgical cure compared to those who did not

All patients	Total, *n*	Surgical cure	No surgical cure	*p*-value* (OR of cure [95%CI])
Number of patients with Acromegaly (%)	150	78 (52%)	72 (48%)	–
Age (yrs), Mean ± SD	48.39 ± 14.12	48.4 ± 14.3	48.4 ± 14.0	0.946
Gender				
Men	65 (43%)	43 (66%)	22 (34%)	0.005 (2.64[1.36–5.23])
Women	85 (57%)	36 (42%)	49 (58%)	
Pre-operative factors				
Pre-operative medical treatment	71 (47%)	33 (46%)	38(54%)	0.19 (0.63[0.33–1.19])
Microadenoma?	29 (19%)	21 (72%)	8 (28%)	0.022 (2.80[1.18–7.27])
Macroadenoma	121 (81%)	57 (47%)	64 (53%)	
Pre-operative GH level in ng/dL Mean ± SD	18.54 ± 39.07	9.72 ± 21.6	26.0 ± 48.3	0.060
Ki-67 Mean ± SD	1.9% ± 1.5%	1.9% ± 1.4%	1.9% ± 1.6%	0.963
Tumour size (largest diameter)in mm Mean ± SD	16.02 ± 8.31	13.2 ± 5.49	19.2 ± 9.70	< 0.001(0.98 per mm[0.97–0.99])
R-L Diameter in mm Mean ± SD	15.97 ± 7.39	13.4 ± 5.24	18.8 ± 8.35	< 0.001(0.98 per mm[0.97–0.99])
Distance between Intracavernous carotids in mm (I)) Mean ± SD	18.46 ± 4.03	18.5 ± 4.25	18.4 ± 3.81	0.923
Lateral maxima of tumour in mm (M)) Mean ± SD	11.24 ± 5.00	9.43 ± 3.14	13.3 ± 5.87	< 0.001(0.96 per mm [0.95–0.98])
M/I (our ratio) Mean ± SD	0.62 ± 0.25	0.52 ± 0.17	0.72 ± 0.29	< 0.001(0.45 per M/I unit [0.34–0.61])
Cavernous sinus invasion?	94 (63%)	41 (44%)	53 (56%)	0.004(0.37[0.18–0.73])
Knosp score				
0	57 (38%)	39 (68%)	18 (32%)	(control group)
1	38 (25%)	25 (66%)	13 (36%)	0.826(0.89[0.37–2.17])
2	22 (15%)	8 (36%)	14 (64%)	0.012(0.27[0.09–0.75])
3 (a and b)	24 (16%)	6 (25%)	18 (75%)	< 0.001(0.16[0.05–0.45])
4	9 (6%)	1 (11%)	8 (89%)	< 0.001(0.07[0.003–0.41])
Menopausal status (female only)	36 (24%))	14 (39%)	22 (61%)	0.659
Post-operative factors				
Radiological residual	76 (51%)	23 (30%)	53 (70%)	< 0.001(0.14[0.07–0.29])
Immediate post op GH (< 2 days) Mean ± SD ng/dL	2.55 ± 4.12	1.06 ± 1.15	3.99 ± 5.32	0.016 (0.96 per ng/dl [0.92–0.99])
Length of inpatient stay Mean ± SD (days)	8.15 ± 12.84	6.56 ± 4.91	9.53 ± 16.9	0.308

on whether they hit the 0.5 threshold and a smoothed linear model (Fig. 5b) which used the formula $Y = 0.93 + (x * -0.49)$ based on the constants and gradient generated by the binomial regression to reduce risk of overfitting.

The MI model using its optimum sensitivity and specificity (Youden's J statistic, optimal threshold of 0.55) achieved the following parameters: AUC: 73%,Accuracy 71% [95% CI: (0.63, 0.78)] on a no information rate (NIR) of 53%, Sensitivity = 68%, Specificity = 73%, PPV = 70%, NPV = 72% Leading to an F1 score of 69%.

We additionally ran multivariate analysis using a backwards stepwise logistic regression. We began with all variables that were found to be significant in a univariate analysis and ended up with Gender ($p = 0.005$) and MI ($p < 0.001$) as the only two variables still significant in a multivariate regression with an AUC of 0.76.

Modelling this regression we achieved the following parameters:

Overall accuracy 69% [95%CI (0.61, 0.77)] on a no information rate (NIR) of 53% and considering an imbalanced dataset a balanced accuracy of 70%, Sensitivity = 89%, specificity = 51%, PPV = 62% and NPV = 83% leading to an F-1 score of 73%.

Modelling the logistic regression using Knosp and gender was less accurate. This model achieved the following parameters:

AUC = 0.76, Accuracy = 67% [95% CI (0.59, 0.75)] on a no information rate (NIR) of 53%, Sensitivity = 63%, specificity = 71%, PPV = 66% and NPV= 68% leading to an F1 score of 65%.

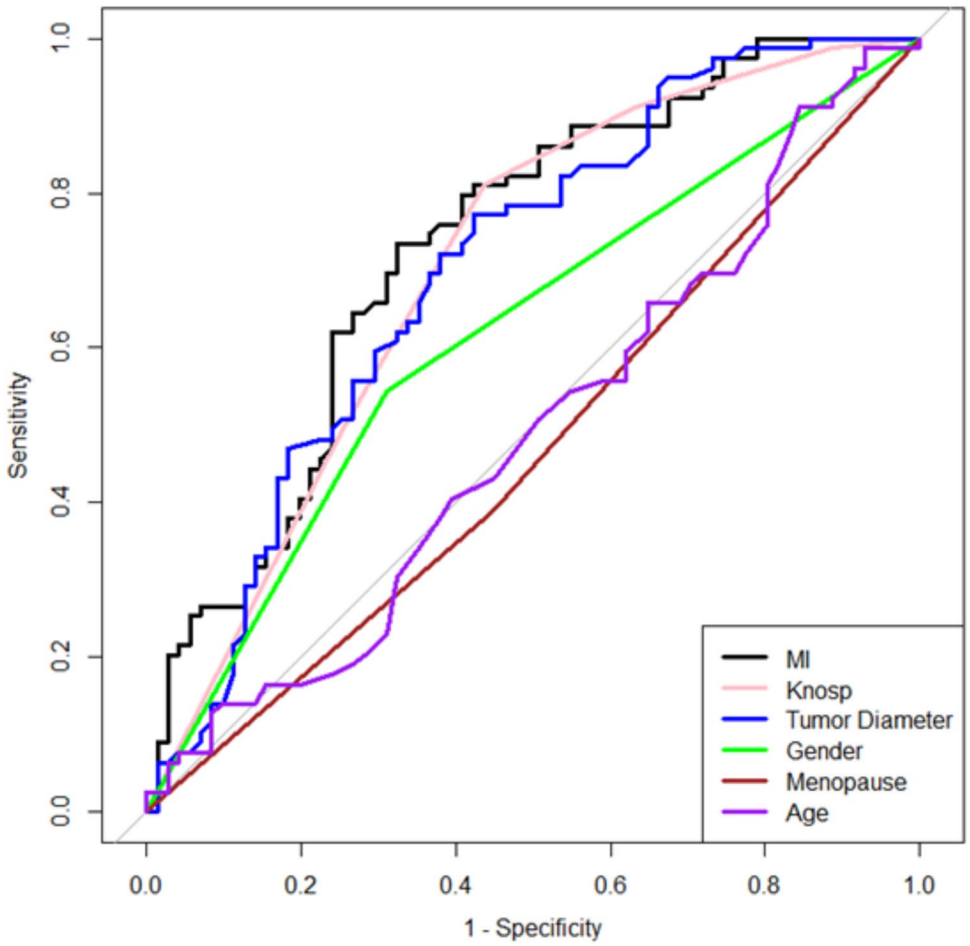

Fig. 4 Receiver operating characteristic (ROC) analysis using variables tested in our dataset. Area under Curve (AUC) denotes sensitivity and specificity of the variable in predicting cure. AUC: MI = 0.73, Knosp = 0.70, Diameter = 0.70, Gender = 0.62, Age = 0.51, Menopause = 0.47

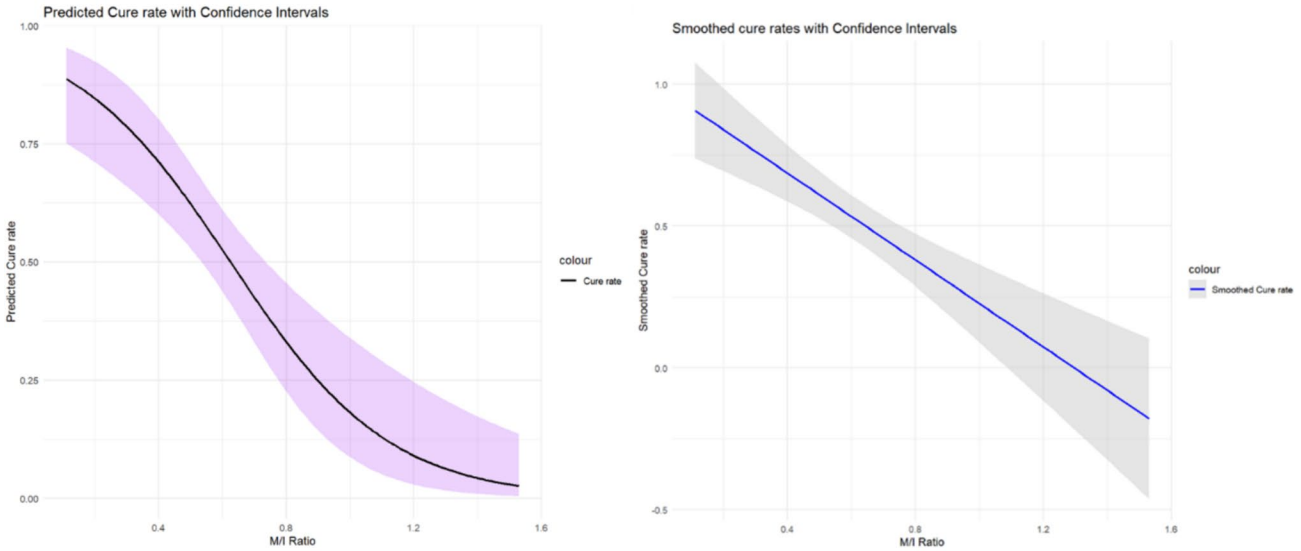

Fig. 5 **A** Actual predicted cure rates corresponding to each MI value with no smoothing and confidence intervals. **B** Smoothed predicted cure rates using the formula Y = 0.93 + (x * −0.49). This corresponds to the constant and gradient calculated using a binomial logistic regression

Discussion

This study represents one of the largest single centre studies concerning the surgical management of acromegaly. Its main contribution is the generation of a new, easy-to-use radiological metric to predict surgical resection and clinical remission (MI ratio.) It also further contributes to the literature's understanding of other factors affecting surgical cure including sex differences and the use of neo-adjuvant therapy.

Cure and outcome predictors

Previous studies have suggested a variety of cure rates as shown in Table 2. These vary greatly by institution and the literature shows a range of proposed values. The largest discrepancy is in papers that have previously used the 2014 consensus criteria (GH nadir < 1 ng/mL) against those that have used the lower 0.4 ng/mL cut-off expressed in the 2018 consensus criteria [28]. However, recent studies have suggested that achieving a cure defined by the 2018 criteria does not significantly improve the prediction of long term remission [26, 34]. The largest study evaluating differences in outcomes between the two cut-offs is Ku et al. [26], who found that in an observational study of 187 patients there was no difference in metabolic outcomes (including BMI, HbA1c, insulin and free fatty acids) which are the strongest predictors of prognosis given metabolic syndrome is the leading cause of death in patients with acromegaly [26, 28].

While some papers have found no difference between the 2 values for the purpose of long term outcomes [26, 34], we chose to use the 0.4 ng/mL cut-off as it is in line with the latest consensus criteria, and because recent studies use it almost exclusively. This will allow for the collation of more homogenous data to improve the evidence base.

Our present study demonstrated an overall cure rate of 53%: 21 of 29 microadenomas (72%) and 57 of 121 macroadenomas (47%) we cured using the 0.4 ng/mL cut-off. This study supports the view that better outcomes are achieved in higher volume centers [7, 19]. Male patients with smaller, less invasive tumours were most likely to achieve initial postsurgical remission. However, long term remission, while shown to be well correlated with initial remission [28, 34] was not the main focus of this study and can be further investigated.

In line with previous studies, our study demonstrated that pre-operative medical management with somatostatin analogues [30, 33] did not improve rates of surgical remission. Yang et al. suggested there may be some short-term benefit but when accounting for washout period they found no long-term benefit of SSA use [40]. Our results support this finding.

The Ki-67 [3] index showed no effect on initial remission rates in our study ($p = 0.963$). However, literature findings suggest a potential correlation between Ki-67 expression and increased likelihood of relapse in the future, as well as eventual non-remission in the long term [3, 19, 28].

Pre-operative GH level did not reach significance with a p-value of 0.060: the mean pre-operative GH level in the cured patients was 9.7 ng/mL and in non-cured patients 26.0 ng/mL. We hypothesize that larger studies will reach significance thresholds.

While menopause has recently been cited as a possible negative outcome predictor [31], our analysis did not identify menopause as a contributing factor in our cohort. Interestingly, our findings indicated that female sex, rather than menopausal status, was associated with worse outcomes in both bivariate and multivariate analyses. Specifically, when comparing menopausal women to non-menopausal women aged 45 and above, we observed no significant difference in initial remission rates.

Studies have conflicting evidence regarding pre-menopausal women with some finding them to be at increased risk of larger and, more invasive tumours and lower remission rates [31], However, our study, along with others [12], found minimal, if any, differences in tumour characteristics or remission rates among pre-menopausal women. Importantly, after adjusting for age, we observed no significant differences in outcomes based on menopausal status. These findings underscore the need for further research to clarify

Table 2 Previously cited cure rates in the literature

Study	Year	Cure rates			Criteria (ng/ml)	No of cases
		Micro	Macro	Overall		
Buchfelder et. Al meta analysis [8]	2017	81	59	62	Both criterions used (734 cases used 0.4 cutoff and 1579 used cutoff of 1)	2313
Bray et.al (2022) [7]	2022	78	48	54	0.4	–
Guo et al. cross sectional study [18]	2020	63	32	46	0.4	473
Guo et al. [19]	2022	51	35	37	0.4	659
KCH	2024	72	48	53	0.4	150

the relationship between menopausal status and acromegaly outcomes.

Since 2006 we have performed all transsphenoidal procedures endoscopically with improved clinical and surgical outcomes. This is supported by studies from other centres that have compared microscopic and endoscopic transsphenoidal procedures [19, 30]. In the largest recent acromegaly study Endoscopic TSS was also found to have significantly better cure rates than microscopic in both micro and macroadenomas (48% vs 17% and 83% vs 43% respectively) [19]

The rate of biochemical failure (OGTT nadir > 0.4 ng/mL) after surgery for adenoma causing acromegaly is only 0 to 19% after GTR [24], so being able to predict likelihood of GTR serves as a proxy for cure. Cavernous sinus invasion regardless of its degree, measured using both Knosp and MI classifications, has been shown to be predictors of radiological residual (indicative of lack of GTR) and non-remission aligning with findings in the current literature [19, 30]

Tumor consistency plays a critical role in determining surgical resectability and remission rates in acromegaly [9, 40]. Softer tumors are generally more amenable to complete resection, whereas firmer, fibrous adenomas are more challenging to remove and are associated with higher rates of residual disease. Unfortunately, in our study, detailed histopathological data regarding tumor consistency were not uniformly recorded due to the retrospective nature of the analysis. This limitation underscores the need for prospective studies that incorporate standardized assessments of tumor consistency, such as intraoperative grading or advanced imaging techniques like elastography. By integrating this variable into predictive models, future research could further enhance the accuracy of preoperative prognostication and inform surgical strategies for achieving optimal outcomes.

It is important to acknowledge that IGF-1 normalization, the definitive marker of biochemical remission [37], often occurs weeks to months postoperatively. The MI Ratio aims to provide an early probabilistic insight into likely outcomes, complementing rather than replacing long-term biochemical markers.

Additionally, functional remission in cases with residual tumour may result from the presence of inactive tumour regions, as supported by recent literature. Additionally, early postoperative imaging can overestimate residual volumes due to inflammatory changes or imaging artifacts. This inherent limitation necessitates cautious interpretation to avoid mislabelling surgical outcomes.

Making predictions

Recent studies such as Cohen-Cohen et al. [11] and prior to that, Fernandez-Rodriguez et al. [14] have endeavoured to construct risk scores aiming at predicting long term remission rates in acromegaly. These risk scores often incorporate various clinical and radiological parameters to make predictions. The gold standard for the radiological element of such scores has been the Knosp classification system [5] to predict the likelihood of achieving GTR and ultimately cure [32]. The Knosp classification is a qualitative measure that is a proxy for the degrees of cavernous invasion. However, it is important to recognize that qualitative measures inherently entail discrete categories, leading to substantial variations in cure rates between these different categories [4, 5].

In our data set we observed distinct cure rates associated with different Knosp classifications: Knosp 0 tumours exhibited a cure rate of 67%, Knosp 1 tumours had a cure rate of 66%, Knosp 2 tumours showed a cure rate of 36.4%, Knosp 3a and 3b tumours demonstrated a cure rate of 25%, and Knosp 4 tumours had a cure rate of 11%. These findings are consistent with existing literature such as the study by Araujo-Castro et al. [5] which reported similar trends, specifically that Knosp class 2 lesions had a 11% lower chance of remission (88% vs 77%) compared to class 1 lesions and class 3 lesions had a 31% lower remission rate than class 2 lesions (46% vs 77%). It is worth noting that the Araujo-Castro et al. paper describes a heterogenous mix of pituitary adenomas, functioning as well as non-functioning, and that they define biochemical cure in their 51 acromegaly patients (out of 228) according to the 2014 guidelines Fig. 6a, b, c, d:

The advantage of using a qualitative measure is that each class has more cases and therefore more statistical power, but it means that cases at the extremities of each class suffer greatly from Interoperator dependence. Mooney et al. found that in 6 reviewers reporting 50 scans of pituitary adenomas in acromegaly, only 10% of cases had unanimous agreement [29]. As illustrated in Fig. 6, distinguishing between classes can be challenging in certain tumours (as depicted in 6a and 6b), leading to significant variations in cure rate predictions. This phenomenon, characterized by large jumps in predictions across different qualitative classifications, is inherent in all qualitative measures.

Additionally, Knosp cannot differentiate between the degrees of laterality in class 3 and class 4 tumours. Looking at Fig. 6c and d, while both are considered Knosp 3a tumours, there is a noticeable discrepancy in the degree of laterality between the two lesions and therefore it is likely that there will be a difference in the chance of cure. Yet there is currently no way of statistically quantifying the increased risk this additional size confers with Knosp.

Despite the limitations of the Knosp classification, it remains a useful tool for predicting cure rates: our study reported a discrimination of AUC 0.7. Our motivation to create MI was to preserve the benefits of measuring the extent of cavernous sinus invasion, while mitigating the shortcomings inherent in qualitative classifications.

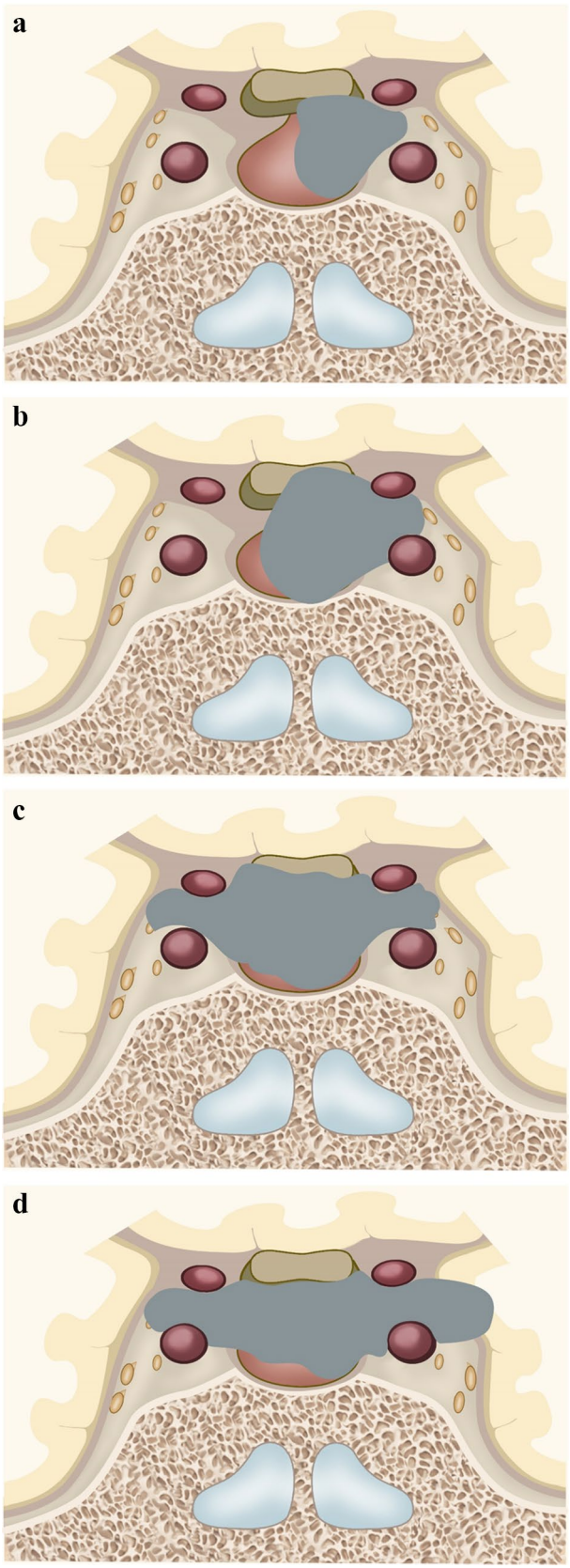

Fig. 6 The limitations of Knosp. Cases **A** and **B** are both borderline ▶ cases. Operators would call case **A** Knosp 1 or 2 and Case **B** could be considered 2 or 3a. Cases **C** and **D** on the other hand are both clearly Knosp 3a cases however they differ greatly in the extent of their invasion

MI is a novel approach aimed at addressing the limitations associated with Knosp classification. By providing a continuous variable, MI offers a more nuanced and precise assessment of cavernous sinus invasion without the large differences observed in qualitative measures. Additionally, MI accounts for differences in invasiveness between Knosp grades within the same class, thereby offering a more comprehensive evaluation. Upon reviewing the images in Fig. 6, it becomes apparent that while these examples are selected to illustrate a point, MI does not exhibit the drawbacks associated with the Knosp classification in these select cases. Overall, MI offers a promising alternative that overcomes many of the limitations of traditional qualitative measures, providing clinicians with a more accurate and robust tool for assessing cavernous sinus invasion in acromegaly patients.

One drawback however is that MI does not look at the "shape" of the laterality, meaning that classes 3a and 3b could be given the same cure rate prediction, or even class 4 in certain circumstances. The literature is divided on whether classes 3a and 3b display different cure rates and some papers group adenomas as above Knosp 3 or below rather than evaluating individual grades [5, 6, 19] and we hypothesize that laterality is more important than whether the invasion is superior or inferior to the carotid.

Our analysis demonstrates that the MI Ratio slightly outperforms the Knosp classification in predicting surgical remission, with an AUC of 0.73 compared to 0.70. Importantly, when examining incorrect predictions, the MI Ratio shows a significantly smaller average deviation from the true outcome (0.081) compared to Knosp (0.154). This reduced margin of error highlights the MI Ratio's potential for greater accuracy, particularly when refined or integrated into multivariable predictive tools. Figure 7 illustrates these findings, showing that MI predictions deviate less substantially from the actual outcome. Unlike the Knosp scale, which assigns fixed probabilities to each grade (67%, 66%, 36%, 25%, and 11% for grades 0–4 in our cohort), the MI Ratio allows for patient-specific predictions with an associated confidence level. For instance, when predictions between 40 and 60% are excluded as inconclusive, the MI Ratio's accuracy improves from 69 to 75%, an adjustment Knosp cannot accommodate.

Knosp 4 cases are clearly less responsive to treatment (11% cure rate compared to 25% for class 3) and while MI

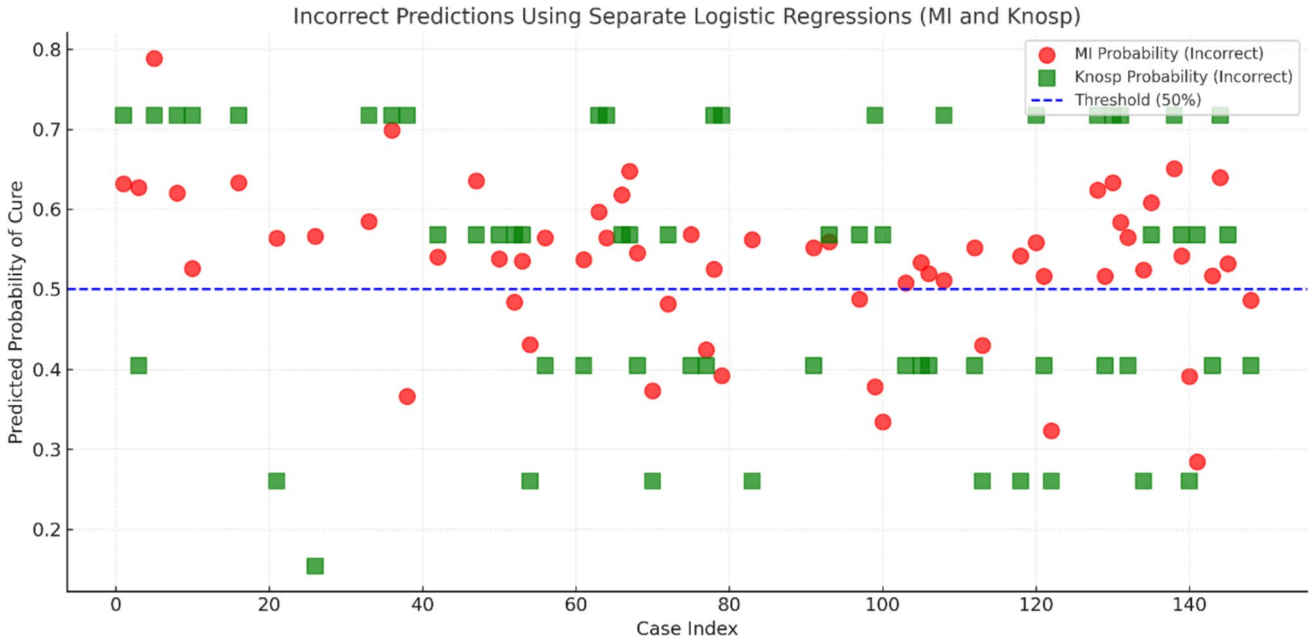

Fig. 7 A scatter graph of "Incorrect" predictions, showing that incorrect MI predictions are almost exclusively low confidence predictions that aggregate near the threshold for prediction and are easily influenced by other factors. In comparison, as Knosp predictions are categorical the distance to accurate result is much larger and more difficult to predict and compensate for in more complex models. Average deviation from the true outcome for MI: 0.081 compared to Knosp: 0.154

ratio will give them a very low chance of cure, invasion by certain shapes of tumours may be underestimated by MI if they encase the carotid without further extension. In our dataset the average extension from midline for Knosp 3 lesions was 15.1 mm compared to 22.5 mm for Knosp 4 cases. Consequently, we believe the effect of shape overall will have a very minor effect on the accuracy of our models.

The MI Ratio also addresses a key limitation of Knosp's categorical nature, which restricts its ability to capture nuance. For example, the smallest grade 2 adenoma in our cohort was considered just over half as likely to be cured as the largest grade 1 adenoma, reflecting the large and inflexible jumps in Knosp predictions. A continuous variable like the MI Ratio enables greater precision and adaptability, making it a valuable addition for advanced modelling, including AI-based approaches, where thresholds can be optimized for specific datasets. While Knosp performs effectively on an aggregated basis and may occasionally outperform MI in certain datasets, MI predictions are consistently closer to the true probability of cure on a case-by-case basis. This utility is further emphasized by the significantly larger deviation observed when Knosp is incorrect, as detailed in Fig. 7. Focusing on incorrect cases, we can clearly visualise the rigidity of incorrect Knosp predictions compared to the MI predictions which were almost exclusively filtered near the midline (50% probability). This is ultimately a proof of concept, and the degree of "incorrectness" is an important metric. This is because as we identify more factors that predict failure or success, the closer the invasiveness metric is to the threshold of 50%, the more likely it is to switch to the other result if other factors suggest this. Importantly, when examining incorrect predictions, the MI Ratio shows a significantly smaller average deviation from the true outcome (0.081) compared to Knosp (0.154). This reduced margin of error highlights the MI Ratio's potential for greater accuracy, particularly when refined or integrated into multivariable predictive tools.

The main strengths of this study lie in its large cohort size for such a rare pathology and period of coverage. Being a retrospective analysis, it allows us to include only the cases with adequate follow-up. Of the 157 patients identified to have surgically treated acromegaly only 7 were excluded from the analysis (less than 5% of the cases at our institution), reducing the possibility of selection bias.

Additionally, all cases being from our large tertiary centre allows for consistency in imaging, pathology, diagnostic criteria, equipment, and surgical experience. Having consistency in these parameters allows us to compare cases with greater confidence as there are fewer confounding variables to be accounted for, and also increases the power of our predictions. This does, however, increase the risk of over-fitting

and in ideal sample size analysis, we would need 384 participants to confidently rule out over-fitting for our MI predictive model. Larger prospective studies are therefore needed to validate MI for general use.

This paper focuses exclusively on initial surgical cure (up to 6 months). Larger, prospective studies are needed to ascertain whether the results shown here can be extrapolated to long term remission and ability to predict relapse. Additionally, in sensitivity analysis, at least 235 cases were recommended as an ideal sample size, suggesting a larger cohort is needed to generalise these findings to the wider population.

The MI Ratio, as presented, is limited by its retrospective validation in a single-center study. Its immediate practical utility in early management remains constrained by the necessity of long-term endocrinological confirmation, such as IGF-1 normalization. Future multicentre prospective studies are needed to establish its role in standard clinical practice.

Conclusions

Acromegaly can be a life-changing disease so a better understanding of prognosis can be a valuable tool in informing patients as well as practitioners.

Non-modifiable factors such as age, gender and anatomy of the cavernous sinuses have been identified as predictors of cure. The degree of laterality in an adenoma and thus cavernous sinus invasion is an important metric that should be considered. Knosp is a highly accurate parameter in predicting success in TSS, however its qualitative nature makes it less flexible. A novel radiological categorization (MI) has been introduced and given further work this may potentially be an accurate metric by which to predict cure for individual patients.

While the MI Ratio is not a standalone tool for immediate clinical decision-making, its value lies in enabling early predictive modelling. By offering individualized probabilistic outcomes, it provides a framework for future AI-driven risk stratification and patient-specific treatment optimization. This aligns with the goal of advancing personalized medicine, particularly in rare and challenging conditions like acromegaly.

Author contributions Conceptualization AA,FF; Data curation FF, AA, CS; Formal analysis FF, FV, JK; Funding acquisition NA; Investigation NA; Methodology FF, AA, FV; Project administration NA; Resources NA; Software FF; Supervision SB, NT, EM, JS; Validation FF,JK; Visualization NA; Roles/Writing—original draft FF,AA,JS; and Writing—review & editing FF,AA,JK,EM,SB.

Funding None of the authors have any conflicts of interest, financial or otherwise. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability Pseudoanonymised data is available upon request as per our institutions policies.

Declarations

Ethical approval Not applicable.

Human Ethics and Consent to Participate Not applicable.

Competing interests The authors declare no competing interests.

References

1. Adigun OO, Nguyen M, Fox TJ, Anastasopoulou C (2023) Acromegaly. [Updated 2023 Feb 2]. In: StatPearls [Internet]. StatPearls Publishing, Treasure Island.Available from: https://www.ncbi.nlm.nih.gov/books/NBK431086/
2. Alam S, Kubihal S, Goyal A, Jyotsna VP (2021) Spontaneous remission of acromegaly after pituitary apoplexy in a middle-aged male. Ochsner J 21(2):194–199
3. Alimohamadi M, Ownagh V, Mahouzi L, Ostovar A, Abbassioun K, Amirjmshidi A (2014) The impact of immunohistochemical markers of Ki-67 and p53 on the long-term outcome of growth hormone-secreting pituitary adenomas: a cohort study. Asian J Neurosurg 9(3):130–136
4. Altman DG, Royston P (2006) The cost of dichotomising continuous variables. BMJ 332(7549):1080
5. Araujo-Castro M, Acitores Cancela A, Vior C, Pascual-Corrales E, Rodríguez Berrocal V (2021) Radiological knosp, revised-knosp, and hardy-wilson classifications for the prediction of surgical outcomes in the endoscopic endonasal surgery of pituitary adenomas: study of 228 cases. Front Oncol 11((Journal Article)):807040
6. Araujo-Castro M, Pascual-Corrales E, Martínez-Vaello V, Baonza Saiz G, Quiñones de Silva J, Acitores Cancela A, García Cano AM, Rodríguez Berrocal V (2021) Predictive model of surgical remission in acromegaly: age, presurgical GH levels and Knosp grade as the best predictors of surgical remission. J Endocrinol Invest 44(1):183–193
7. Bray DP, Mannam S, Rindler RS, Quillin JW, Oyesiku NM (2022) Surgery for acromegaly: indications and goals. Front Endocrinol 13((Journal Article)):924589
8. Buchfelder M, Schlaffer S-M (2017) The surgical treatment of acromegaly. Pituitary 20(1):76–83

9. Chanson P (2015) Medical treatment of acromegaly with dopamine agonists or somatostatin analogs. Neuroendocrinology 103(1):50–58
10. Chanson P, Salenave S (2008) Acromegaly. Orphanet J Rare Dis 3:17
11. Cohen-Cohen S, Rindler R, Botello Hernandez E, Donegan D, Erickson D, Meyer FB, Atkinson JL, Van Gompel JJ (2023) A novel preoperative score to predict long-term biochemical remission in patients with growth-hormone secreting pituitary adenomas. World Neurosurg S1878–8750(23):01809
12. Dal J, Skov BG, Andersen M, Feldt-Rasmussen U, Feltoft CL, Karmisholt J, Nielsen EH, Dekkers OM, Jørgensen JOL (2021) Sex differences in acromegaly at diagnosis: a nationwide cohort study and meta-analysis of the literature. Clin Endocrinol (Oxf) 94(4):625–635
13. Díez JJ, Iglesias P (2000) Current management of acromegaly. Expert Opin Pharmacother 1(5):991–1006
14. Fernandez-Rodriguez E, Casanueva FF, Bernabeu I (2014) Update on prognostic factors in acromegaly: is a risk score possible? Pituitary 18(3):431–440
15. Freda PU, Landman RE, Sundeen RE, Post KD (2001) Gender and age in the biochemical assessment of cure of acromegaly. Pituitary 4(3):163–171
16. Giustina A, Arnaldi G, Bogazzi F et al (2017) Pegvisomant in acromegaly: an update. J Endocrinol Invest 40(6):577–589
17. Giustina A, Chanson P, Bronstein MD, Klibanski A, Lamberts S, Casanueva FF, Trainer P, Ghigo E, Ho K, Melmed S (2010) A consensus on criteria for cure of acromegaly. J Clin Endocrinol Metab 95(7):3141–3148
18. Guo X, Wang K, Yu S, Gao L, Wang Z, Zhu H, Xing B, Zhang S, Dong D (2020) Patient characteristics, diagnostic delays, treatment patterns, treatment outcomes, comorbidities, and treatment costs of acromegaly in china: a nationwide study. Front Endocrinol 11((Journal Article)):610519
19. Guo X, Zhang R, Zhang D, et al (2014) Determinants of immediate and long-term remission after initial transsphenoidal surgery for acromegaly and outcome patterns during follow-up: a longitudinal study on 659 patients. J Neurosurg (Journal Article):1–11
20. Hannon MJ, Barkan AL, Drake WM (2016) The role of radiotherapy in acromegaly. Neuroendocrinology 103(1):42–9. https://doi.org/10.1159/000435776
21. Holdaway IM, Rajasoorya C (1999) Epidemiology of acromegaly. Pituitary 2(1):29–41
22. Kamusheva M, Vandeva S, Mitov K, et al (2020) New Epidemiological, clinical and economic data for patients with acromegaly in Bulgaria. Front Public Health 8
23. Katznelson L (2006) Drug insight: primary medical therapy of acromegaly. Nat Clin Pract Endocrinol Metab 2(2):109–117. https://doi.org/10.1038/ncpendmet0096
24. Katznelson L, Laws ER Jr, Melmed S, Molitch ME, Murad MH, Utz A, Wass JAH (2014) Acromegaly: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 99(11):3933–3951
25. Knosp E, Steiner E, Kitz K, Matula C (1993) Pituitary adenomas with invasion of the cavernous sinus space: a magnetic resonance imaging classification compared with surgical findings. Neurosurgery 33(4):610–618
26. Ku CR, Choe EY, Hong JW, Kim EH, Park SH, Kim SH, Lee EJ (2016) No differences in metabolic outcomes between nadir GH 0.4 and 1.0 ng/mL during OGTT in surgically cured acromegalic patients (observational study). Medicine (Baltimore) 95(24):e3808
27. Melmed S (2009) Acromegaly pathogenesis and treatment. J Clin Invest 119(11):3189–3202
28. Melmed S, Bronstein MD, Chanson P, Klibanski A, Casanueva FF, Wass JAH, Strasburger CJ, Luger A, Clemmons DR, Giustina A (2018) A consensus statement on acromegaly therapeutic outcomes. Nat Rev Endocrinol 14(9):552–561
29. Mooney MA, Hardesty DA, Sheehy JP, Bird R, Chapple K, White WL, Little AS (2017) Interrater and intrarater reliability of the Knosp scale for pituitary adenoma grading. J Neurosurg 126(5):1714–1719
30. Papaioannou C, Druce M (2023) Preoperative medical treatments and surgical approaches for acromegaly: a systematic review. Clin Endocrinol (Oxf) 98(1):14–31
31. 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 Metab 103(3):909–916
32. Petrossians P, Borges-Martins L, Espinoza C, Daly A, Betea D, Valdes-Socin H, Stevenaert A, Chanson P, Beckers A (2005) Gross total resection or debulking of pituitary adenomas improves hormonal control of acromegaly by somatostatin analogs. Eur J Endocrinol 152(1):61–66
33. Pita-Gutierrez F, Pertega-Diaz S, Pita-Fernandez S, Pena L, Lugo G, Sangiao-Alvarellos S, Cordido F (2013) Place of preoperative treatment of acromegaly with somatostatin analog on surgical outcome: a systematic review and meta-analysis. PLoS ONE 8(4):e61523
34. Ronchi CL, Arosio M, Rizzo E, Lania AG, Beck-Peccoz P, Spada A (2007) Adequacy of current postglucose GH nadir limit (< 1 microg/l) to define long-lasting remission of acromegalic disease. Clin Endocrinol (Oxf) 66(4):538–542
35. RV L, RY O, G K, J R WHO Classification of tumours of endocrine organs
36. Sandret L, Maison P, Chanson P (2011) Place of cabergoline in acromegaly: a meta-analysis. J Clin Endocrinol Metab 96(5):1327–1335. https://doi.org/10.1210/jc.2010-2443
37. Subbarayan SK, Fleseriu M, Gordon MB et al (2012) Serum IGF-1 in the diagnosis of acromegaly and the profile of patients with elevated IGF-1 but normal glucose-suppressed growth hormone. Endocr Pract Off J Am Coll Endocrinol Am Assoc Clin Endocrinol 18(6):817–825
38. Webb SM, Badia X (2015) Quality of life in acromegaly. Neuroendocrinology 103(1):106–111
39. World Medical Association (2013) World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 310(20):2191–2194
40. Yang C, Li G, Jiang S, Bao X, Wang R (2019) Preoperative somatostatin analogues in patients with newly-diagnosed acromegaly: a systematic review and meta-analysis of comparative studies. Sci Rep 9(1):14070
41. Zahr R, Fleseriu M (2018) Updates in diagnosis and treatment of acromegaly. Eur Endocrinol 14(2):57–61
42. Zheng Y, Chen D-M, Wang Y, Mai R-K, Zhu Z-F (2020) Surgical management of growth hormone-secreting pituitary adenomas: a retrospective analysis of 33 patients. Medicine (Baltimore) 99(19):e19855