Proposal and Validation of a Simple Grading Scale (TRANSSPHER Grade) for Predicting Gross Total Resection of Nonfunctioning Pituitary Macroadenomas After Transsphenoidal Surgery

Michael A. Mooney, MD[*]
Christina E. Sarris, MD[*]
James J. Zhou, MD[*]
Garni Barkhoudarian, MD[‡]
Michael R. Chicoine, MD[§]
Juan C. Fernandez-Miranda, MD[¶]
Paul A. Gardner, MD[¶]
Douglas A. Hardesty, MD[*]
Heidi Jahnke, RN, MSN[*]
Daniel F. Kelly, MD[‡]
Brandon D. Liebelt, MD[*]
Marc R. Mayberg, MD[‖]
Daniel M. Prevedello, MD[#]
John Sfondouris, MD[*]
John P. Sheehy, MD[*]
James P. Chandler, MD[**]
Kevin C. J. Yuen, MD[‡‡]
William L. White, MD[*]
Andrew S. Little, MD[*]
on behalf of The TRANSSPHER Study Group

[*]Department of Neurosurgery, Barrow Neurological Institute, St. Joseph's Hospital and Medical Center, Phoenix, Arizona; [‡]Pacific Neuroscience Institute and Pituitary Disorders Center, John Wayne Cancer Institute at Providence Saint John's Health Center, Santa Monica, California; [§]Department of Neurological Surgery, Washington University School of Medicine, St. Louis, Missouri; [¶]Department of Neurological Surgery, University of Pittsburgh School of Medicine, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [‖]Department of Neurological Surgery, University of Washington School of Medicine, Seattle, Washington; [#]Department of Neurological Surgery, The Ohio State University Wexner Medical Center, Columbus, Ohio; [**]Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Chicago, Illinois; [‡‡]Department of Neurology and Barrow Neuroendocrinology Clinic, Barrow Neurological Institute, St. Joseph's Hospital and Medical Center, Phoenix, Arizona

Correspondence:
Andrew S. Little, MD,
c/o Neuroscience Publications,
Barrow Neurological Institute,
St. Joseph's Hospital and Medical Center,
350 W. Thomas Rd.,
Phoenix, AZ 85013.
E-mail: Neuropub@barrowneuro.org

Received, May 17, 2018.
Accepted, December 31, 2018.

BACKGROUND: A simple, reliable grading scale to better characterize nonfunctioning pituitary adenomas (NFPAs) preoperatively has potential for research and clinical applications.

OBJECTIVE: To develop a grading scale from a prospective multicenter cohort of patients that accurately and reliably predicts the likelihood of gross total resection (GTR) after transsphenoidal NFPA surgery.

METHODS: Extent-of-resection (EOR) data from a prospective multicenter study in transsphenoidal NFPA surgery were analyzed (TRANSSPHER study; ClinicalTrials.gov NCT02357498). Sixteen preoperative radiographic magnetic resonance imaging (MRI) tumor characteristics (eg, tumor size, invasion measures, tumor signal characteristics, and parameters impacting surgical access) were evaluated to determine EOR predictors, to calculate receiver-operating characteristic curves, and to develop a grading scale. A separate validation cohort (n = 165) was examined to assess the scale's performance and inter-rater reliability.

RESULTS: Data for 222 patients from 7 centers treated by 15 surgeons were analyzed. Approximately one-fifth of patients (18.5%; 41 of 222) underwent subtotal resection (STR). Maximum tumor diameter > 40 mm; nodular tumor extension through the diaphragma into the frontal lobe, temporal lobe, posterior fossa, or ventricle; and Knosp grades 3 to 4 were identified as independent STR predictors. A grading scale (TRANSSPHER grade) based on a combination of these 3 features outperformed individual variables in predicting GTR (AUC, 0.732). In a validation cohort, the scale exhibited high sensitivity and specificity (AUC, 0.779) and strong inter-rater reliability (kappa coefficient, 0.617).

CONCLUSION: This simple, reliable grading scale based on preoperative MRI characteristics can be used to better characterize NFPAs for clinical and research purposes and to predict the likelihood of achieving GTR.

KEY WORDS: Adenoma, Extent of resection, Grading scale, Pituitary, Residual, Transsphenoidal

Operative Neurosurgery 0:1–10, 2019 DOI: 10.1093/ons/opy401

Transsphenoidal surgery is the preferred treatment for most patients with symptomatic nonfunctioning pituitary adenomas (NFPAs). Subtotal resection (STR) is reported for 15% to 30% of lesions,[1-5] although reported rates vary widely. Although numerous predictors of the extent of tumor removal, such as measures of tumor invasion (eg, Knosp grade) and tumor composition (eg, cysts, hemorrhage, and T2 signal), have been explored, there is no current system for characterizing lesions that synthesizes features into a scale predictive of gross total resection (GTR).[2,4,6-13]

The TRANSSPHER Study (Transsphenoidal Extent of Resection Study; ClinicalTrials.gov

ABBREVIATIONS: AUC, area under the curve; **EOR,** extent of resection; **GTR,** gross total resection; **MRI,** magnetic resonance imaging; **NFPAs,** nonfunctioning pituitary adenomas; **OR,** odds ratio; **ROC,** receiver-operating characteristic; **STR,** subtotal resection; **STROBE,** Strengthening the Reporting of Observational Studies in Epidemiology

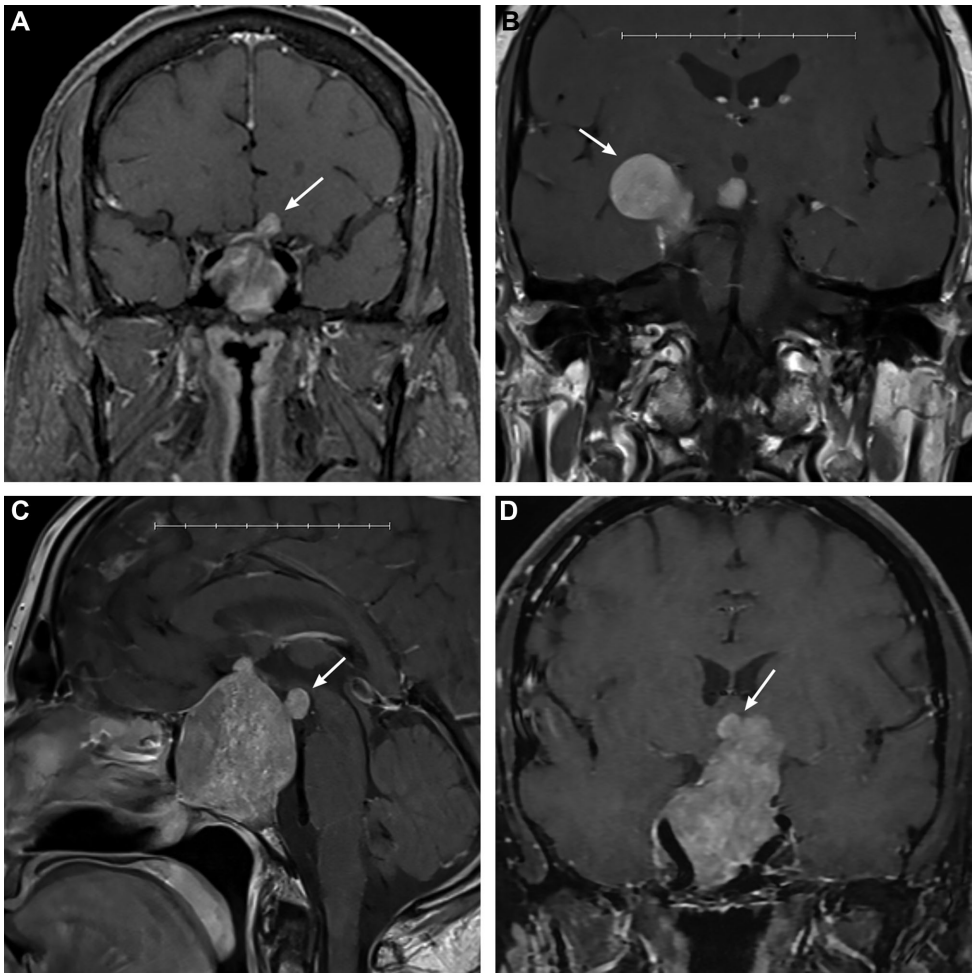

FIGURE 1. *Radiographic examples of nodular extension. Nodular extension through the diaphragma suggests subarachnoid extension or invasion of the brain parenchyma and limits tumor removal. Sequential panels demonstrate nodular extension (arrows) into the* **A**, *frontal lobe,* **B**, *temporal lobe,* **C**, *posterior fossa, and* **D**, *ventricle. Used with permission from Barrow Neurological Institute, Phoenix, Arizona.*

NCT02357498) is a prospective multicenter research collaboration investigating patients with surgically treated NFPAs. It involves 15 surgeons with various levels of experience, 7 surgical centers, and centralized adjudicated imaging review. This trial provides a unique opportunity to study factors that influence the extent of tumor removal and to develop a grading scale generalizable to proficient transsphenoidal surgeons. This scale would support research efforts by standardizing how surgical series are reported and would be clinically useful by informing preoperative decision-making, patient counseling, and prognostication.

The goal of this study was to develop a simple and reliable grading scale that predicts GTR using features easily measured on preoperative pituitary magnetic resonance imaging (MRI). We examined numerous covariates, including tumor size, invasiveness, MRI signal characteristics, and anatomical factors that affect the ease of surgical access, and we distilled them into the 3 imaging features that drive the sensitivity and specificity of the scale. We propose this model for clinical and research use, similar to scales used in other neurosurgical disciplines.[14-22]

METHODS

Study Design and Participants

Multicenter Cohort

Patients enrolled in the TRANSSPHER Study were evaluated after providing written informed consent, as approved by the institutional review boards at participating institutions. Patients were enrolled

between 2015 and 2018. The overall study adhered to the principles of the *US Code of Federal Regulations*, Title 45, Part 46, "Protection of Human Subjects" (revised January 15, 2009). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cohort studies were utilized as the reporting guidelines for this study (http://www.strobe-statement.org).

Preoperative Imaging Variables

A list of potentially relevant imaging features was compiled by reviewing the literature and obtaining input from the surgeons involved in the study.[2,9,23-28] Variables included maximum tumor diameter (any plane, including axial, coronal, sagittal, or oblique); tumor volume (measured by volumetric software); Knosp grade; presence of clival invasion; presence of a tumor "waist" (ie, dumbbell morphology); intercarotid distance; conchal; or presellar-type sphenoid sinus; presence of a "shallow sella"; T2 signal intensity within the tumor; nodular tumor extension into the frontal lobe, temporal lobe, posterior fossa, or third ventricle; presence of hemorrhage or cystic features (> 50% of tumor volume); and gadolinium enhancement pattern, such as heterogeneous enhancement. Open-source volumetric software was used for the preoperative volumetric analysis (3D Slicer, https://www.slicer.org).[29]

Clival invasion was defined as the invasion of a tumor within the clivus that would require drilling of the clivus to reach the inferior aspect of the tumor. A tumor waist was defined in the coronal plane as narrowing of the tumor caused by the diaphragma where the suprasellar portion of the tumor extended lateral to the waist (ie, the tumor had a "dumbbell" or "figure 8" shape). Tumor waist ratio was calculated as the maximum coronal diameter of a tumor above the waist compared with the maximum coronal diameter of the constriction point. Shallow sella was defined as a small tuberculum sella face (distance from the tuberculum to the inferiormost aspect of the sella turcica) less than 10 mm and potentially limiting access to a suprasellar tumor, as judged by the raters. The T2 signal was characterized as hypo-, iso-, or hyperintense to normal pituitary gland or adjacent temporal lobe gray matter, as described elsewhere.[30-32] Nodular extension was defined as a lobular invasive pattern of growth through the diaphragma into the frontal lobe, temporal lobe, posterior fossa, or ventricle (Figure 1). Intraoperatively, nodular extension often correlates with a breach in the diaphragma where the tumor enters the subarachnoid space or brain parenchyma.

Imaging Review

Patients underwent preoperative and postoperative (within 3 mo of surgery) dedicated pituitary-protocol T1-weighted contrast-enhanced MRIs, including fine-cut coronal images through the sella, on 1.5T or 3.0T magnets. Tumors and other anatomical covariates were scored by 2 independent raters (M.A.M. and C.S.). All discrepancies between raters were individually reviewed to reach a consensus after review by the senior author. Extent of resection (EOR) was dichotomized on anonymized scans as either GTR or STR by the consensus of 3 reviewers at the coordinating institution blinded to treatment group (1 neuroradiologist [J.S.] and 2 neurosurgeons [B.L. and A.S.L.]). For an MRI to be scored as STR, residual tumor had to be detected on at least 2 consecutive thin-cut MRI slices in one imaging plane and confirmed on a slice in another imaging plane.

Validation Cohort

Consecutive patients (n = 165) with surgically treated NFPAs from the lead institution before or after the trial enrollment period served as the validation cohort. Patients who underwent surgery between 2011 and 2017 were included. Two independent raters (D.H. and J.P.S.) not involved in the creation of the original scale and blinded to postoperative outcomes assigned TRANSSPHER grades to all scans. GTR was determined on the basis of independent radiologist assessment of postoperative MRIs performed within 1 yr of surgery; indeterminant studies were independently reviewed in a blinded fashion to classify GTR vs STR.

TABLE 1. Characteristics of Patients in the Multicenter Cohort

Variable	No. (%)* (n = 222)
Age, mean (range), yr	59.1 (20 to 83)
Sex	
Male	134 (60)
Female	88 (40)
Prior surgery	23 (10)
Presenting symptoms	
Visual deficit	103 (46)
Headache	76 (34)
Galactorrhea	6 (3)
Hypopituitarism	26 (12)
Amenorrhea	9 (4)
Incidental	62 (28)
Approach	
Endoscopic	154 (69)
Microscopic	68 (31)

*Values are number (%), unless indicated otherwise.

Statistical Methods

All statistical analyses were performed using Stata Statistical Software (StataCorp, LLC, College Station, Texas). Chi-squared tests, Fisher exact tests, and independent-samples *t*-tests were used when appropriate. First, a univariate analysis was performed to screen for covariates of interest. Covariates with $P < .10$ were subsequently included in a binary logistic regression analysis to evaluate for independent predictors of STR. Diameter and volume measurements were considered individually, given the high degree of collinearity; each was evaluated as both a continuous variable and a categorical variable in sequential models.

Covariates that retained statistical significance on multivariate logistic regression were incorporated into a grading scale designed to predict the probability of GTR. Model-predicted probabilities were used to create receiver-operating characteristic (ROC) curves and to obtain an area under the curve (AUC). The AUC values between models and individual variables were compared and assessed, with statistical significance defined as $P < .05$. Inter-rater reliability of the scale was assessed using the Cohen's kappa correlation coefficient, and coefficients were interpreted as follows: 0.00 to 0.19 = very weak; 0.20 to 0.39 = weak; 0.40 to 0.59 = moderate; 0.60 to 0.79 = strong; 0.80 to 1.00 = very strong.[33]

RESULTS

Patient Characteristics—Multicenter Exploratory Cohort

A total of 222 of 243 paired preoperative and postoperative MRIs from 7 centers and 15 surgeons were scored. For patient demographics, see Table 1. Twenty-one patients were excluded

TABLE 2. Univariate Analysis of Preoperative Patient and Tumor Characteristics

Characteristic	GTR,* No./Total (%)	STR,* No./Total (%)	*P* value
Patient age			>.99
< 70 yr	33/177 (19)	144/177 (81)	
≥70 yr	8/45 (18)	37/45 (82)	
Maximum tumor diameter			
Mean (range), mm	25 (7.3-49.3)	34 (18.0-60.9)	<.001
>30 mm	49/74 (66)	25/74 (34)	<.001
>40 mm	2/11 (18)	9/11 (82)	<.001
Tumor volume			
Mean (range), cm^3	6.3 (0.1-24.4)	13.3 (2.8-52.8)	<.01
>5 cm^3	94/126 (75)	32/126 (25)	<.01
>10 cm^3	29/49 (59)	20/49 (41)	<.001
>20 cm^3	3/9 (33)	6/9 (67)	<.01
>30 cm^3	3/3 (100)	0/3 (0)	<.01
Surgical technique			.45
Endoscope	128/154 (83)	26/154 (17)	
Microscope	53/68 (78)	15/68 (22)	
Knosp grades			
0 to 2	160/181 (88)	21/181 (12)	<.001
3 to 4	21/41 (51)	20/41 (49)	<.001
Intercarotid distance, mean (range), mm	17 (9.0-27.4)	19 (11.0-27.3)	.06
Tumor waist			
Yes	69/91 (76)	22/91 (24)	.07
Ratio ≥ 1.2	24/36 (67)	12/36 (33)	.02
Ratio ≥ 1.4	9/16 (56)	7/16 (44)	.01
Clival invasion	4/9 (44)	5/9 (56)	.01
Shallow sella	21/23 (91)	2/23 (9)	.27
Conchal sella	0/1 (0)	1/1 (100)	.19
Presellar	5/7 (71)	2/7 (29)	.62
T2 signal†			
Hypointense	8/11 (73)	3/11 (27)	.69
Isointense	84/99 (85)	15/99 (15)	.23
Hyperintense	86/109 (79)	23/109 (21)	.39
Nodular extension			
Frontal extension	6/11 (55)	5/11 (45)	.03
Temporal extension	0/3 (0)	3/3 (100)	.006
Posterior fossa extension	1/3 (33)	2/3 (67)	.09
Intraventricular extension	1/7 (14)	6/7 (86)	<.001
Any of the above	7/16 (44)	9/16 (56)	<.001
Hemorrhagic	8/9 (89)	1/9 (11)	.99
Cystic	20/21 (95)	1/21 (5)	.14

GTR, gross total resection; STR, subtotal resection.
*Values are number (%), unless indicated otherwise.
†Three MRI scans could not be analyzed because of imaging limitations or because the lesions were entirely cystic.

because of inadequate preoperative or postoperative MRIs (ie, inadequate coronal and sagittal reconstructions, lack of intravenous contrast, or loss to follow-up). Most patients (154 of 222; 69%) underwent endoscopic resection; the remainder underwent microscopic resection (68; 31%).

Independent Predictors of STR

Tumor diameter, tumor volume, Knosp grades 3 to 4, tumor waist ratio, clival invasion, and any nodular extension into the frontal lobe, temporal lobe, posterior fossa, or ventricle were predictive of STR ($P < .05$) by univariate analysis (Table 2). The presence of a tumor waist approached statistical significance ($P < .10$). There was no difference in rate of GTR by age (<70 yr, 81% [144/177]; ≥ 70 yr, 82% [37/45]) or surgical technique (endoscopic vs microscopic, $P > .45$). There was no statistical difference in rates of GTR between initial and repeat surgery, variations of sellar configuration, or T2 signal intensity within the tumor ($P > .16$). Counterintuitively, decreased intercarotid distance demonstrated an increased GTR proportion ($P = .06$) and therefore was not included in the multivariate

TABLE 3. Select Multivariate Regression Models Analyzing Predictors of Subtotal Resection

Variable	*P* value	OR	95% CI
Maximum diameter > 30 mm	.10	2.0	0.86 to 4.82
Knosp grades 3 to 4	<.001	4.9	2.1 to 11.27
Tumor waist	.58	1.3	0.56 to 2.85
Clival invasion	.53	1.7	0.32 to 9.28
Nodular extension	.53	3.8	1.16 to 12.78
Maximum diameter > 40 mm	.03	6.6	1.16 to 37.63
Knosp grades 3 to 4	.001	4.5	1.92 to 10.69
Tumor waist	.35	1.5	0.66 to 3.22
Clival invasion	.5	1.9	0.30 to 11.61
Nodular extension	.07	3.3	0.90 to 12.18
Tumor volume (continuous; cm^3)	.002	1.1	1.04 to 1.21
Knosp grades 3 to 4	<.001	4.5	1.95 to 10.52
Tumor waist	.67	1.2	0.53 to 2.67
Clival invasion	.91	0.9	0.13 to 6.37
Nodular extension	.21	2.4	0.62 to 9.07
Tumor volume > 10 cm^3	.01	2.9	1.26 to 6.73
Knosp grades 3 to 4	<.001	5.1	2.24 to 11.76
Tumor waist	.44	1.4	0.62 to 3.00
Clival invasion	.68	1.5	0.26 to 8.16
Nodular extension	.04	3.7	1.08 to 12.45

Abbreviations: CI, confidence interval; OR, odds ratio.

analysis. Enhancement patterns were not studied further because of poor inter-rater agreement and a lack of consensus.

To determine the covariates that independently predicted GTR, we analyzed tumor characteristics approaching statistical significance in the univariate analysis using a series of exploratory logistic regression models. Tumor diameter (continuous variable), maximum diameter > 40 mm, tumor volume (continuous variable), tumor volume > 10 cm^3, and Knosp grades 3 to 4 were independently associated with an increased likelihood of STR ($P < .05$; Table 3). Nodular extension (frontal lobe, temporal lobe, posterior fossa, or intraventricular) demonstrated borderline significance ($P = .04$ to $P = .07$), depending on the individual model. The presence of a tumor waist, high tumor waist ratio, and clival invasion were not significant. Maximum tumor diameter and tumor volume measurements demonstrated strong collinearity ($r^2 = 0.7552$, $P < .001$).

Proposed Grading Scale (TRANSSPHER Grade)

We developed a simple grading scale to predict GTR using independent predictors from the multicenter cohort (Figure 2). One point each was assigned for tumor diameter > 40 mm (in any plane); for Knosp grades 3 to 4; and for nodular extension into the frontal lobe, temporal lobe, posterior fossa, or ventricle. The likelihood of achieving GTR was inversely related to tumor grade (Figure 3A). The ROC for the proposed grading scale was significantly better at predicting GTR than any of the 3 individual factors (Figure 3B; $P < .001$). Comparison of a grading scale utilizing the maximum diameter versus the tumor volume demonstrated no significant difference in the performance of the scale (Figure 3B; AUC, 0.732 vs 0.758; $P = .31$).

TRANSSPHER Grade Validation

First, we evaluated the performance of the proposed grading scale in a new cohort of 165 patients. The ROC analysis demonstrated comparable values in the multicenter and validation cohorts (AUC, 0.732 vs 0.779; Figure 4A). The likelihood of achieving GTR was inversely related to tumor grade, supporting the results from the exploratory cohort (Figure 4B).

Second, we assessed the inter-rater reliability of the TRANSSPHER grade. For individual variables, the maximum diameter measurement, nodular extension variable, and dichotomized Knosp grades 3 to 4 variable demonstrated strong inter-rater reliability (kappa coefficients, 0.765, 0.747, 0.615, respectively). The overall proposed TRANSSPHER grade demonstrated strong inter-rater reliability (kappa coefficient, 0.617).

DISCUSSION

The goal of this study was to develop a simple and reliable grading scale using preoperative imaging features to predict GTR in patients with NFPAs undergoing transsphenoidal surgery. This scale can be used by clinicians for preoperative surgical planning, for patient counseling and prognosticating, and for classifying patients for standardized research reporting. We explored numerous tumor characteristics, tumor MRI signal characteristics, measures of tumor invasiveness, and anatomical features that might influence surgical access, such as sella configuration and intercarotid distance. The multicenter nature of this study provides a unique opportunity to investigate data from 15 individual surgeons in 7 independent centers.

Key Findings

This study describes a simple, reliable grading scale of readily measurable tumor anatomical characteristics that predict an outcome of interest to surgeons (ie, the likelihood of achieving GTR) after transsphenoidal surgery. This study is the first, to our knowledge, to evaluate the interaction of various tumor characteristics and to synthesize them into a practical grading scale. After evaluating numerous potential predictors, we discovered that 3 characteristics emerged as strong independent predictors of GTR: tumor size ≥ 40 mm, presence of nodular tumor extension, and preoperative Knosp grades 3 to 4. The likelihood of achieving GTR is inversely related to TRANSSPHER grade. The scale was validated in a separate cohort of patients using raters not involved in the conception of the scale. The TRANSSPHER grade is a better predictor of GTR than any of its individual scale components.

Interpretation and Generalizability

Numerous studies have examined EOR for pituitary adenomas.[2,8,9,23-26,34,35] However, their patient inclusion

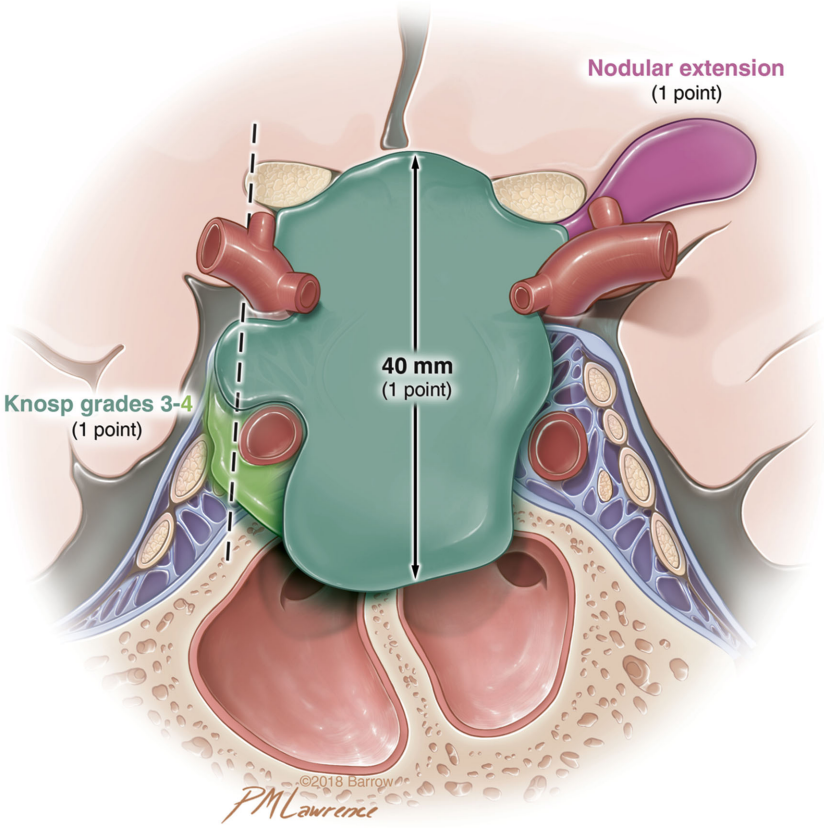

FIGURE 2. *Artist's illustration depicts the proposed grading scale (TRANSSPHER grade) for pituitary adenomas. One point each is assigned for tumor diameter > 40 mm in any plane, for Knosp grades 3 to 4, and for nodular extension. The sum of all 3 components (maximum 3 points) yields the TRANSSPHER grade. Dashed line demarcates lateral border of internal carotid artery. Illustration used with permission from Barrow Neurological Institute, Phoenix, Arizona.*

criteria, surgeon experience, and specific adenoma characteristics are highly variable. Furthermore, these series often include only a single-experienced surgeon to interpret the postoperative imaging results for the patients. To develop the TRANSSPHER grade, we evaluated a prospective multicenter study of microscopic and endoscopic transsphenoidal surgery performed by 15 different surgeons to identify variables most strongly associated with GTR. Our study design has several strengths. First, EOR in this multicenter cohort was evaluated by 3-blinded raters at the coordinating institution. Second, both microscopic and endoscopic surgeons were included, which increases the applicability of the scale. Notably, no difference in GTR outcomes was observed between endoscopic and microscopic surgeons in this study ($P = .45$). Third, we were able to validate the scale in a separate cohort of patients with a new set of raters to assess the scale performance and reliability.

Tumor Diameter

The appropriate method for estimating pituitary adenoma size and relevant cutoffs for tumor size has been debated.[7,8,36-40] For example, some authors have proposed that diameter measurements in the coronal or axial planes may have a stronger association with EOR than those in the sagittal plane, with tumor volume having the strongest association.[27] To be sure, it is more time-consuming to determine tumor volume than to determine tumor diameter, and to do so requires volumetric software. For the purposes of the TRANSSPHER grade, we chose the maximum tumor diameter in any plane because of its simplicity and reliability, which did not compromise the performance of the scale. We found that increasing tumor diameter and increasing tumor volume were both independently associated with a decreased likelihood of GTR (Table 3). We evaluated the performance of the

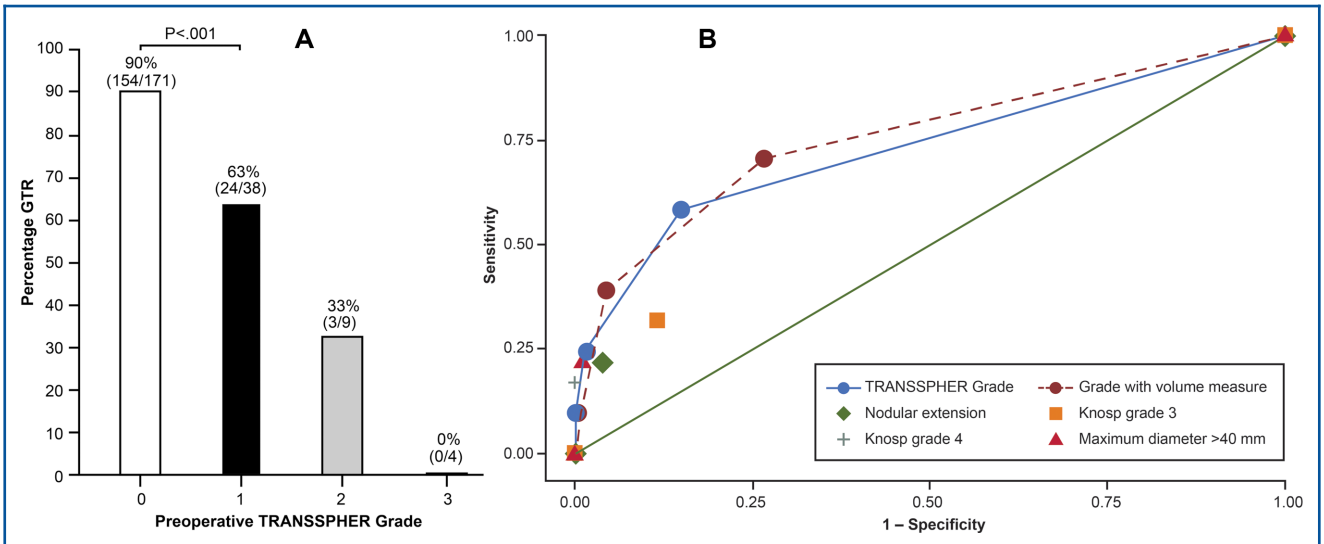

FIGURE 3. *Analysis of the TRANSSPHER grade in the multicenter cohort.* A, *Gross total resection (GTR) rates for pituitary adenomas stratified by the proposed grading scale (TRANSSPHER grade). The likelihood of achieving GTR decreased for tumors with a higher TRANSSPHER grade (P < .001).* B, *Receiver-operating characteristic curve analysis demonstrates the improved performance of the proposed grading scale (TRANSSPHER grade) compared to individual factors for predicting gross total resection (GTR; P < .001 for all variables). A version of the grading scale using the volume > 10 cm^3 cutoff value in place of diameter is also illustrated (dashed line). No significant difference in the area under the curve (AUC) between the grading scale with maximum diameter and the grading scale with the volume measurement was observed (AUC, 0.732 vs 0.758; P = .31). Used with permission from Barrow Neurological Institute, Phoenix, Arizona.*

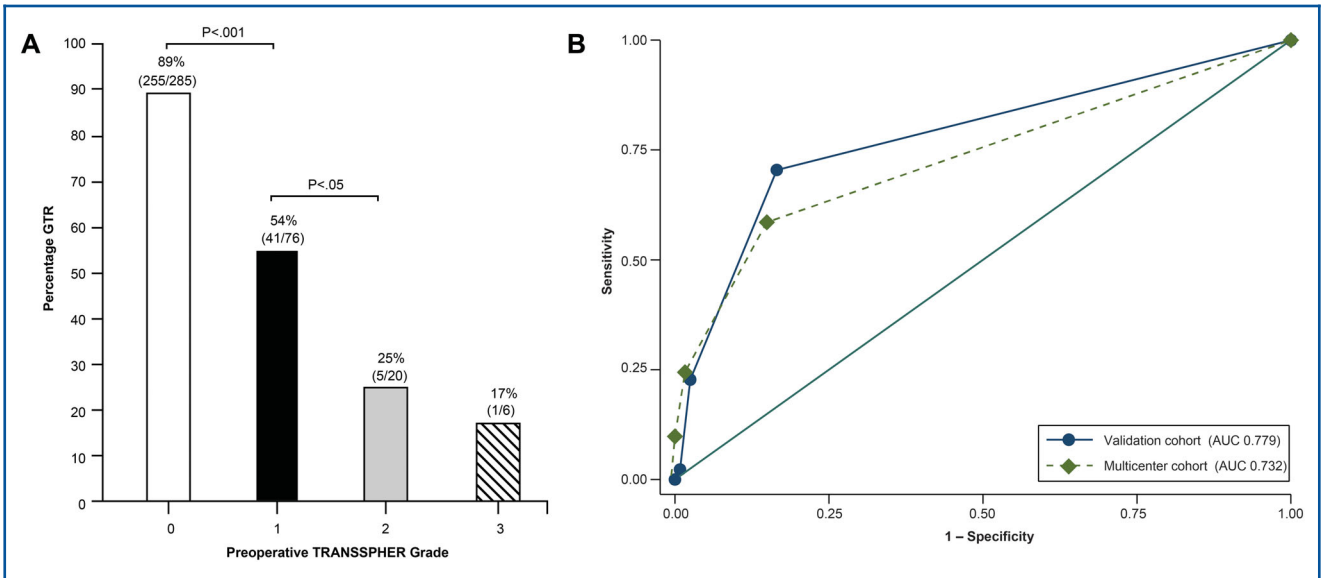

FIGURE 4. *Analysis of the TRANSSPHER grade in the combined multicenter and validation cohort.* A, *The likelihood of achieving gross total resection (GTR) in the combined multicenter and validation cohort decreased for tumors with a higher TRANSSPHER grade.* B, *Receiver-operating characteristic curves for the TRANSSPHER grade in the multicenter cohort (area under the curve [AUC], 0.732) and in the validation cohort (AUC, 0.779), demonstrating reliable performance of the grading scale.* Used with permission from Barrow Neurological Institute, Phoenix, Arizona.

TRANSSPHER grade using both diameter and volume measurements and found no significant difference in its predictive ability. The ROC analysis demonstrated the strength of this variable and supports its use in the grading scale.

Interestingly, we examined various maximum diameter cutoff values for the TRANSSPHER grade and found that maximum tumor diameter > 30 mm was not associated with STR, whereas maximum diameter > 40 mm was associated with STR (Table 2). This finding is consistent with other reports,[23,26,36,41] and a diameter of 40 mm has been suggested in the literature as a cutoff point for giant adenomas.

Nodular Extension

Adenoma extension through the diaphragma into the frontal lobe, temporal lobe, posterior fossa, or ventricle poses a substantial challenge during transsphenoidal surgery. Nodular tumor extension as we describe it (Figure 1) implies a rent in the diaphragma that the tumor has transgressed to reach the subarachnoid space or brain parenchyma. Intraoperative observations suggest that these invasive tumors have a poorly defined tumor capsule and may place adjacent cortex, cerebral vessels, and cranial nerves at risk, which limits the likelihood of safe resection.[25]

Other studies have used different measures to capture the complexity of lesions with clinically significant suprasellar extension. Our definition for the TRANSSPHER grade was informed by this work. For example, Nishioka et al[25] determined that an "intracranial extension index" (a calculated ratio of intracranial tumor to total tumor) was an independent predictor of EOR in giant adenomas. In their examination of multilobular configuration, Koutourousiou et al[26] found that this measure also correlated with EOR in patients with giant adenomas. Goel et al[7] developed a classification system of 4 groups for giant adenomas defined by levels of aberrant extension and proposed surgical management strategies for each group.

Although numerous methods of capturing the features of adenomas likely exist, we isolated nodular extension because it can easily be recognized by raters and is less likely to overlap with other relevant features, such as overall tumor size or Knosp grade. Our analysis of inter-rater reliability confirmed this hypothesis, as the nodular extension variable was associated with strong reliability between raters in our exploratory and validation cohorts.

Knosp Grade

Lesions involving the cavernous sinus are a challenge to surgeons because of their relationship to the cavernous carotid artery and adjacent cranial nerves. Preoperative Knosp grade is well established as a predictor of intraoperative invasion of the cavernous sinus, with incremental increases in the likelihood of invasion among the higher grades.[10] It is also an independent predictor of EOR.[2,4,6-8,23,42]

We chose to dichotomize the Knosp grades into lesions less likely to have cavernous sinus invasion (grades 0-2) and more likely to have cavernous sinus invasion (grades 3-4) since this is the most common method for reporting these results and because it is associated with improved inter-rater and intrarater reliabilities.[42] We did not examine intraoperative findings of cavernous sinus invasion, because our aim was to describe lesions on the basis of preoperative imaging characteristics.

Other Preoperative Imaging Characteristics Not Included in the Scale

We evaluated numerous other preoperative imaging characteristics thought to potentially influence NFPA EOR, including sellar configurations, intercarotid distance, presence or absence of a tumor waist, measurements of tumor-to-waist ratio, and tumor T2 signal intensity. Interestingly, none of these factors retained significance on multivariate analysis as independent predictors of STR. Although some of these anatomical factors may indeed make surgery more challenging, they were not associated with decreased likelihood of GTR in our multicenter cohort and thus were not included in the TRANSSPHER grade. T2 signal intensity has been proposed as a potential indicator of tumor consistency and has been linked to tumor functional status and pathologic characteristics.[30-32] Our study found no association between T2 signal intensity and NFPA EOR, which to our knowledge represents the most extensive examination of this topic in the literature to date.

Limitations

Several limitations of this study should be considered. First, our analysis of radiographic predictors is limited by the variables we examined. We surveyed experienced pituitary surgeons and reviewed the literature to compile a list of 16 imaging variables; however, there may be other anatomical characteristics that we did not examine. Second, although we identified a statistically significant decrement in the rate of GTR between TRANSSPHER grades 0 to 1 and 1 to 2, we did not identify a similar decrement for grades 2 to 3, although the percentage of GTR decreased for grade 3 compared with grade 2 in both the exploratory cohort and the combined cohort (0% vs 33% and 17% vs 25%). Statistical analysis of these findings was limited by the relative rarity of grade 3 lesions. Future studies with larger numbers of high-grade lesions are warranted. Third, the TRANSSPHER grade does not incorporate surgeon intent. For example, it may be acceptable in a patient with visual disturbance to debulk a tumor for near-total resection (rather than strive for GTR) to decompress the optic chiasm. This decision may be partly influenced by the radiographic variables we isolated in this scale, prohibiting further analysis with this study design.

CONCLUSION

Our examination of numerous imaging variables (eg, tumor size, invasiveness, MRI signal characteristics, and anatomical features that impact surgical access) identified

tumor diameter > 40 mm, nodular tumor extension, and Knosp grades 3 to 4 as strong, independent predictors of GTR in a prospective multicenter trial of transsphenoidal pituitary surgery for NFPAs. We developed and validated a simple, reliable grading scale (TRANSSPHER grade) based on these characteristics in a separate cohort of NFPA patients. The scale predicts GTR and has strong inter-rater reliability. This scale better characterizes NFPAs and may be used for clinical prognostication and to standardize reporting for research purposes.

Disclosures

Dr Little is a consultant for Kogent Surgical, LLC. The other authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article.

REFERENCES

1. Zhang X, Fei Z, Zhang J, et al. Management of nonfunctioning pituitary adenomas with suprasellar extensions by transsphenoidal microsurgery. *Surg Neurol*. 1999;52(4):380-385.
2. Mortini P, Losa M, Barzaghi R, Boari N, Giovanelli M. Results of transsphenoidal surgery in a large series of patients with pituitary adenoma. *Neurosurgery*. 2005;56(6):1222-1233; discussion 1233.
3. Dallapiazza R, Bond AE, Grober Y, et al. Retrospective analysis of a concurrent series of microscopic versus endoscopic transsphenoidal surgeries for Knosp grades 0–2 nonfunctioning pituitary macroadenomas at a single institution. *J Neurosurg*. 2014;121(3):511-517.
4. Zaidi HA, Awad AW, Bohl MA, et al. Comparison of outcomes between a less experienced surgeon using a fully endoscopic technique and a very experienced surgeon using a microscopic transsphenoidal technique for pituitary adenoma. *J Neurosurg*. 2016;124(3):596-604.
5. Bodhinayake I, Ottenhausen M, Mooney MA, et al. Results and risk factors for recurrence following endoscopic endonasal transsphenoidal surgery for pituitary adenoma. *Clin Neurol Neurosurg*. 2014;119:75-79.
6. Fatemi N, Dusick JR, de Paiva Neto MA, Kelly DF. The endonasal microscopic approach for pituitary adenomas and other parasellar tumors: a 10-year experience. *Neurosurgery*. 2008;63(4 Suppl 2):244-256; discussion 256.
7. Goel A, Nadkarni T, Muzumdar D, Desai K, Phalke U, Sharma P. Giant pituitary tumors: a study based on surgical treatment of 118 cases. *Surg Neurol*. 2004;61(5):436-445; discussion 445-436.
8. de Paiva Neto MA, Vandergrift A, Fatemi N, et al. Endonasal transsphenoidal surgery and multimodality treatment for giant pituitary adenomas. *Clin Endocrinol (Oxf)*. 2010;72(4):512-519.
9. Hofstetter CP, Nanaszko MJ, Mubita LL, Tsiouris J, Anand VK, Schwartz TH. Volumetric classification of pituitary macroadenomas predicts outcome and morbidity following endoscopic endonasal transsphenoidal surgery. *Pituitary*. 2012;15(3):450-463.
10. Micko AS, Wohrer A, Wolfsberger S, Knosp E. Invasion of the cavernous sinus space in pituitary adenomas: endoscopic verification and its correlation with an MRI-based classification. *J Neurosurg*. 2015;122(4):803-811.
11. Mooney MA, Hardesty DA, Sheehy JP, et al. Rater reliability of the Hardy classification for pituitary adenomas in the magnetic resonance imaging era. *J Neurol Surg B*. 2017;78(05):413-418.
12. Hardy J, Vezina JL. Transsphenoidal neurosurgery of intracranial neoplasm. *Adv Neurol*. 1976;15:261-273.
13. Knosp E, Steiner E, Kitz K, Matula C. Pituitary adenomas with invasion of the cavernous sinus space: a magnetic resonance imaging classification compared with surgical findings. *Neurosurgery*. 1993;33(4):610-617; discussion 617–618.
14. Spetzler RF, A proposed grading system for arteriovenous malformations. *J Neurosurg*. 1986;65(4):476-483.
15. Lawton MT, Kim H, McCulloch CE, Mikhak B, Young WL. A supplementary grading scale for selecting patients with brain arteriovenous malformations for surgery. *Neurosurgery*. 2010;66(4):702-713; discussion 713.
16. Garcia RM, Ivan ME, Lawton MT. Brainstem cavernous malformations. *Neurosurgery*. 2015;76(3):265-278; discussion 277-268.
17. Chang EF, Smith JS, Chang SM, et al. Preoperative prognostic classification system for hemispheric low-grade gliomas in adults. *J Neurosurg*. 2008;109(5):817-824.
18. Pollock BE, A proposed radiosurgery-based grading system for arteriovenous malformations. *J Neurosurg*. 2002;96(1):79-85.
19. Magill ST, Morshed RA, Lucas CG, et al. Tuberculum sellae meningiomas: grading scale to assess surgical outcomes using the transcranial versus transsphenoidal approach. *Neurosurg Focus*. 2018;44(4):E9.
20. Neidert MC, Maldaner N, Stienen MN, et al. The Barrow Neurological Institute Grading Scale as a predictor for delayed cerebral ischemia and outcome after aneurysmal subarachnoid hemorrhage: data from a nationwide patient registry (Swiss SOS). *Neurosurgery*. 2018;83(6):1286-1293.
21. Dengler NF, Diesing D, Sarrafzadeh A, Wolf S, Vajkoczy P. The Barrow Neurological Institute Scale revisited: predictive capabilities for cerebral infarction and clinical outcome in patients with aneurysmal subarachnoid hemorrhage. *Neurosurgery*. 2017;81(2):341-349.
22. Nonaka Y, Grossi PM, Bulsara KR, Taniguchi RM, Friedman AH, Fukushima T. Microsurgical management of hypoglossal schwannomas over 3 decades: a modified grading scale to guide surgical approach. *Neurosurgery*. 2011;69(2 Suppl Operative):ons121-140; discussion ons140.
23. Gondim JA, Almeida JP, Albuquerque LA, Gomes EF, Schops M. Giant pituitary adenomas: surgical outcomes of 50 cases operated on by the endonasal endoscopic approach. *World Neurosurg*. 2014;82(1-2):e281-e290.
24. Musluman AM, Cansever T, Yilmaz A, et al. Surgical results of large and giant pituitary adenomas with special consideration of ophthalmologic outcomes. *World Neurosurg*. 2011;76(1-2):141-148; discussion 163-146.
25. Nishioka H, Hara T, Nagata Y, Fukuhara N, Yamaguchi-Okada M, Yamada S. Inherent tumor characteristics that limit effective and safe resection of giant nonfunctioning pituitary adenomas. *World Neurosurg*. 2017;106:645-652.
26. Koutourousiou M, Gardner PA, Fernandez-Miranda JC, Paluzzi A, Wang EW, Snyderman CH. Endoscopic endonasal surgery for giant pituitary adenomas: advantages and limitations. *J Neurosurg*. 2013;118(3):621-631.
27. Chohan MO, Levin AM, Singh R, et al. Three-dimensional volumetric measurements in defining endoscope-guided giant adenoma surgery outcomes. *Pituitary*. 2016;19(3):311-321.
28. Smith KA, Leever JD, Chamoun RB. Prediction of consistency of pituitary adenomas by magnetic resonance imaging. *J Neurol Surg B*. 2015;76(05):340-343.
29. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. 2012;30(9):1323-1341.
30. Heck A, Emblem KE, Casar-Borota O, Bollerslev J, Ringstad G. Quantitative analyses of T2-weighted MRI as a potential marker for response to somatostatin analogs in newly diagnosed acromegaly. *Endocrine*. 2016;52(2):333-343.
31. Potorac I, Petrossians P, Daly AF, et al. Pituitary MRI characteristics in 297 acromegaly patients based on T2-weighted sequences. *Endocr Relat Cancer*. 2015;22(2):169-177.
32. Dogansen SC, Yalin GY, Tanrikulu S, et al. Clinicopathological significance of baseline T2-weighted signal intensity in functional pituitary adenomas. *Pituitary*. 2018;21(4):347-354.
33. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46.
34. Juraschka K, Khan OH, Godoy BL, et al. Endoscopic endonasal transsphenoidal approach to large and giant pituitary adenomas: institutional experience and predictors of extent of resection. *J Neurosurg*. 2014;121(1):75-83.
35. Karki M, Sun J, Yadav CP, Zhao B. Large and giant pituitary adenoma resection by microscopic trans-sphenoidal surgery: surgical outcomes and complications in 123 consecutive patients. *J Clin Neurosci*. 2017;44:310-314.
36. Mortini P, Barzaghi R, Losa M, Boari N, Giovanelli M. Surgical treatment of giant pituitary adenomas: strategies and results in a series of 95 consecutive patients. *Neurosurgery*. 2007;60(6):993-1004; discussion 1003-1004.
37. Jefferson G. Extrasellar extensions of pituitary adenomas: (section of neurology). *Proc R Soc Med*. 1940;33(7):433-458.
38. Symon L, Jakubowski J, Kendall B. Surgical treatment of giant pituitary adenomas. *J Neurol Neurosurg Psychiatry*. 1979;42(11):973-982.
39. Lee SH, Park JS, Lee S, Kim SW, Hong YK. Parasellar extension grades and surgical extent in endoscopic endonasal transsphenoidal surgery for pituitary adenomas : a single surgeon's consecutive series with the aspects of reliability and clinical validity. *J Korean Neurosurg Soc*. 2016;59(6):577-583.

40. Agrawal A, Cincu R, Goel A. Current concepts and controversies in the management of non-functioning giant pituitary macroadenomas. *Clin Neurol Neurosurg*. 2007;109(8):645-650.
41. Sinha S, Sharma BS. Giant pituitary adenomas–an enigma revisited. Giant pituitary adenomas—An enigma revisited. Microsurgical treatment strategies and outcome in a series of 250 patients. *Br J Neurosurg*. 2010;24(1):31-39.
42. Mooney MA, Hardesty DA, Sheehy JP, et al. Interrater and intrarater reliability of the Knosp scale for pituitary adenoma grading. *J Neurosurg*. 2017;126(5):1714-1719.

Acknowledgments

The authors thank the staff of Neuroscience Publications at Barrow Neurological Institute for assistance with manuscript preparation. The TRANSSPHER Study Group: Barrow Neurological Institute, Phoenix, Arizona: William L. White, MD; Andrew S. Little, MD; John Sfondouris, MD; Brandon D. Liebelt, MD; and Heidi Jahnke, RN, MSN. University of Pittsburgh, Pittsburgh, Pennsylvania: Paul A. Gardner, MD; Juan C. Fernandez-Miranda, MD; and Benita Valappil, MPH. Washington University, St. Louis, Missouri: Michael R. Chicoine, MD; Ralph Dacey, MD; Gregory Zipfel, MD; Albert Kim, MD; and John Evans, RN. Pacific Neuroscience Institute and Pituitary Disorders Center, John Wayne Cancer Institute at Providence Saint John's Health Center, Santa Monica, California: Daniel F. Kelly, MD; Garni Barkhoudarian, MD; and Annie Heng, BSN. Northwestern University, Chicago, Illinois: James P. Chandler, MD; Matt S. Lesniak, MD; Orin Bloch, MD; and Christina Amidei, PhD. Swedish Neuroscience Institute, Seattle, Washington: Kevin C. J. Yuen, MD; and Caryl Tongco, CCRP. University of Washington, Seattle, Washington: Marc Mayberg, MD. The Ohio State University, Columbus, Ohio: Daniel M. Prevedello, MD; Bridget Hoskins, BS. Harvard Medical School, Boston, Massachusetts: Timothy R. Smith, MD.

COMMENT

The authors present a simple grading scale for predicting gross-total resection of non-functioning pituitary macroadenomas after transsphenoidal surgery. The study is well-designed with sufficiently small interobserver variability of metrics assessed. As presented, this scale may not guide the choice of surgery vs no surgery since nonfunctional adenomas lack a medical treatment and would go to surgery even if GTR could not be achieved. This scale may however prove useful in future prospective multicenter trials for nonfunctional adenomas or could even help guide a similar future study of functional adenomas such as prolactinomas to help preoperatively predict the chances of biochemical remission, which if low could shift decision making towards other therapeutic modalities. The authors are to be commended for their work.

Manish Aghi
San Francisco, California