

Supplementary Table 4: Details of included studies: Predictors of aggressive behaviour

Study (year) - Design	Population (n)	Mean duration of follow-up in months (range)	Predictors of aggressive behaviour
Lenders (2023) - Cohort	Patients with surgically treated pituitary tumours (386), of which 16.6% APT (64) - somatotroph (73) - lactotroph (24) - lactotroph/somatotroph (1) - corticotroph (34) - thyrotroph (2) - NFPA (252)	Median 90 (IQR 43.8-126.2)	Association between aggressive behaviour and: - Tumour size: macroadenomas: 95.3 vs 85.9%, p 0.038. OR 3.3 [95%CI 0.99-10.98], p 0.0512 - Cavernous sinus invasion: 65.6 vs 38.8%, p <0.001. OR 3.0 [95%CI 1.71-5.28], p 0.0001 - Sphenoid sinus invasion: 43.8 vs 22.0%, p <0.001. OR 2.7 [95%CI 1.57-4.81], p 0.0004 - Cavernous sinus and/or sphenoid sinus invasion: 79.7 vs 48.1%, p <0.001. OR 4.2 [95%CI 2.21-8.07], p <0.0001 - Ki67 ≥ 3%: 23.8 vs 13.4%, p 0.034. OR 2.0 [95%CI 1.03-3.85], p 0.0420 - Mitoses > 2/10 HPF: 17.5% vs 6.6%, p 0.005. OR 3.0 [95%CI 1.36-6.53], p 0.0065 - Grade 1a: OR 0.2 [95%CI 0.10-0.43], p <0.0001 - Grade 1b: OR 0.9 [95%CI 0.27-3.33], p 0.9243 - Grade 2a: OR 2.6 [95%CI 1.48-4.57], p 0.001 - Grade 2b: OR 3.4 [95%CI 1.35-8.57], p 0.0096 No relation between aggressive behaviour and age at surgery, gender, tumour subtype and p53 positivity
Liu (2022) - Cohort	Patients with pituitary tumours (2021), of which 44 refractory tumours, matched to 44 non-refractory tumours - somatotroph (18) - lactotroph (7) - gonadotroph (15) - corticotroph (16) - Crooke's cell (8) - null cell (33) - pit-1 positive plurihormonal (1) - acidophilic stem cell (1)	n.r.	Association between aggressive behaviour and: Whole cohort: - Male sex: 56.9 vs 40.7%, p <0.05. OR 1.9 [95%CI 1.05-3.50], p 0.0347 - Mean age at diagnosis: 46.6 vs 42.77 years, p <0.05 - Tumour growth rate*: 4.4%/month vs n.r., "significantly faster" Non-refractory tumours: - Tumour size*: 48.6 vs 22.6 mm, p <0.01 - Giant tumour at presentation: 54.5 vs 2.3%, p <0.01. OR 51.6 [95%CI 6.51-408.74], p 0.0002 - Invasiveness: cavernous sinus 84.1 vs 11.4%, suprasellar extension: 52.3 vs 0%, extension into clival region: 31.8 vs 0%, p <0.01 • Cavernous sinus OR 41.2 [95%CI 12.02-141.43], p <0.0001 • Suprasellar extension OR 97.3 [95%CI 5.64-1678.55], p 0.0016 • Extension into clival region OR 42.3 [95%CI 2.43-736.35], p 0.0102 - Atypical features at histopathological examination: 29.5 vs 0%, p <0.0001. OR 38.1 [95%CI 2.19-665.62], p 0.0126 - Mean Ki67: 8 vs. 1.2%, p <0.001 - Positive p53 immunostaining: 34.1 vs 2.3%, p <0.001. OR 22.2 [95%CI 2.78-177.74], p 0.0034 - Abundant mitosis/nuclear pleomorphism: 45.5 vs 4.5%, p <0.001. OR 17.5 [95%CI 3.76-81.44], p 0.0003
Miermeister (2015) - Cohort	10 PC, 40 typical pituitary adenomas based	n.r.	- Ki-67 ≥ 4%: OR 1.8, p 0.012, R ² 0.41 - p53 positive nuclei ≥ 4%: OR 1.8, p 0.01, R ² 0.44 Sensitivity 60%, specificity 93%, accuracy 86% for both parameters

	on histological and clinical grounds^		No association for mitoses > 2/10 HPF and invasiveness
Trouillas (2023) - Cohort	43 tumours (23 APT + 20 PC) from the ESE survey cohort on which the five-tiered classification could be applied	104 (24-216) For 6 patients, follow-up time could not be determined	Trouillas classification grade 2b at the initial surgery: sensitivity 68% for APT/PC^^
<p>APT = aggressive pituitary tumour NFPA = nonfunctioning pituitary adenoma PC = pituitary carcinoma</p> <p>IQR = interquartile range HPF = high power field EGFR = epidermal growth factor receptor MMP2 = matrix metalloproteinase 2 MM9 = matrix metalloproteinase 9</p> <p>n.r. = not reported * It was not reported if this was determined at presentation, during follow-up, or after (surgical) treatment ^Data on 98 atypical pituitary adenomas not included since they were defined based on histological features solely and may show heterogenous clinical behaviour (data retrieved after contacting the authors). ^^ The article also reported specificity of the grade 2b compared to a cohort of 365 unselected surgically treated patients with pituitary tumours. This is not included here since clinical data on behaviour was missing from this cohort.</p>			