U.PORTO

MESTRADO INTEGRADO EM MEDICINA

Evaluation of the clinical applicability of the PANOMEN-3 classification in predicting the prognosis of pituitary neuroendocrine tumours

Rita Ribeiro Carvalho

M

2026

Evaluation of the clinical applicability of the PANOMEN-3 classification in predicting the prognosis of pituitary neuroendocrine tumours

Rita Ribeiro Carvalho

up202007536@icbas.up.pt

Mestrado Integrado em Medicina

Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

Orientador: Cláudia Soares do Amaral

Coorientador: Doutora Carolina Helena de Freitas Noronha e Gil

Coorientador: Diogo Brandão Neves

Serviço de Endocrinologia, Unidade Local de Saúde de Santo António, Porto, Portugal;

Serviço de Neurocirurgia, Unidade Local de Saúde de Santo António, Porto, Portugal;

Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal.

Porto, junho 2026

Declaração de Integridade

Declara-se que o conteúdo desta dissertação é original e que todo o material de terceiros foi devidamente referenciado de acordo com as normas académicas vigentes. Afirma-se, ainda, não existir qualquer conflito de interesses, atual ou potencial, de natureza financeira, pessoal ou profissional, suscetível de influenciar a realização, os resultados ou a interpretação do presente trabalho.

Autora: Rita Ribeiro Carvalho

Assinado por: **Rita Ribeiro Carvalho**
Num. de Identificação: 30930132
Data: 2026.06.04 22:42:53+01'00'
CHAVE MÓVEL

Orientadora: Cláudia Soares do Amaral

Assinado por: **CLÁUDIA SOARES DO AMARAL**
Num. de Identificação: 10300997
Data: 2026.06.05 09:03:51+01'00'
CARTÃO DE CIDADÃO

Coorientadora: Doutora Carolina Helena de Freitas Noronha e Gil

Coorientador: Diogo Brandão Neves

Assinado por: Diogo Filipe Brandão Neves
Identificação: BI15228197
Data: 2026-06-05 às 11:51:56

Porto, junho 2026

AGRADECIMENTOS

Em primeiro lugar, gostaria de expressar o meu sincero agradecimento à minha orientadora, Cláudia Soares do Amaral, por me ter proposto este tema e por ter acompanhado atentamente este trabalho desde o início. A sua disponibilidade e dedicação foram uma fonte de incentivo ao longo de todo o percurso, tornando os desafios mais fáceis de ultrapassar e contribuindo para que esta experiência se revelasse verdadeiramente enriquecedora. Sob a sua orientação, este projeto proporcionou-me um contacto mais próximo com a investigação, o que despertou em mim um interesse acrescido por esta área.

À minha coorientadora, Doutora Carolina Helena de Freitas Noronha e Gil, agradeço pelas orientações e sugestões partilhadas, bem como pela atenção e amabilidade com que sempre se disponibilizou.

Agradeço, ainda, ao meu coorientador, Diogo Brandão Neves, pelo acompanhamento próximo e pelo empenho demonstrado ao longo de todo este projeto. Esteve presente em todas as etapas do seu desenvolvimento, desde as questões iniciais até às decisões mais complexas. O seu apoio foi imprescindível para a sua concretização.

Deixo um agradecimento especial aos meus pais, verdadeiros alicerces da minha educação, pelo apoio, força e valores que me transmitiram. À minha irmã, obrigada pelo carinho, amizade e por estar sempre presente. Ao meu namorado, obrigada pelo apoio incondicional, pela tranquilidade transmitida nos momentos de maior ansiedade e, claro, por todas as explicações de estatística, sempre dadas com uma paciência admirável.

Por fim, aos meus amigos de curso, obrigada por todos os momentos partilhados, pela entreajuda e pela amizade construída ao longo destes anos. Levo comigo amizades e memórias que certamente permanecerão para além desta etapa.

RESUMO

Introdução: Os tumores neuroendócrinos da hipófise (PitNET) constituem um grupo heterogéneo de neoplasias intracranianas com comportamento clínico variável. Embora a maioria apresente um curso indolente, alguns tumores têm uma evolução atípica e agressiva, demonstrando crescimento invasivo, persistência ou recorrência tardia após cirurgia e tratamento médico. É, portanto, crucial identificar ferramentas prognósticas fiáveis. A classificação Pituitary Adenoma Nomenclature and Multidimensional Evaluation (PANOMEN-3) foi recentemente proposta como um sistema multidimensional que integra variáveis clínicas, bioquímicas, radiológicas e histopatológicas para estimar o risco de morbimortalidade associada aos PitNET. Este estudo teve como objetivo avaliar o desempenho prognóstico da classificação PANOMEN-3 numa coorte cirúrgica de doentes com PitNET, bem como analisar o valor prognóstico das suas variáveis individuais.

Métodos: Estudo observacional retrospetivo de uma coorte consecutiva de doentes adultos com PitNET confirmados histologicamente, submetidos a cirurgia hipofisária num centro terciário académico, a Unidade Local de Saúde de Santo António, no Porto, entre 2017 e 2024. Os scores PANOMEN-3 foram calculados com base em variáveis clínicas, bioquímicas, radiológicas e histopatológicas recolhidas dos registos clínicos eletrónicos. Analisámos a sobrevivência livre de progressão de acordo com os graus da classificação PANOMEN-3 e com as suas variáveis individuais.

Resultados: Foram incluídos 92 doentes, com um tempo mediano de seguimento de 69 meses (IQR 36). A maioria dos tumores correspondia a macroadenomas (72,8%). Os subtipos histológicos mais frequentes foram os PitNET corticotróficos (28,3%), gonadotróficos (27,2%) e somatotróficos (21,7%). De acordo com a classificação PANOMEN-3 pós-operatória, 5,4% dos doentes foram classificados como grau 0, 27,2% como grau 1, 57,6% como grau 2 e 9,8% como grau 3. O aumento do grau PANOMEN-3 associou-se a tempos significativamente mais curtos até recorrência ou progressão tumoral, refletindo uma redução progressiva da sobrevivência livre de progressão (log-rank $p < 0,001$). Na regressão de Cox, os tumores PANOMEN-3 grau 2 (HR 6,13; IC 95% 2,15–17,51; $p = 0,001$) e grau 3 (HR 15,58; IC 95% 4,32–56,26; $p < 0,001$) apresentaram risco significativamente superior de recorrência ou progressão tumoral.

Conclusões: A classificação PANOMEN-3 demonstrou bom desempenho prognóstico nesta coorte cirúrgica de doentes com PitNET. O risco de recidiva ou progressão tumoral foi previsto significativamente por scores mais elevados nesta escala (graus 2 e 3). Estes associaram-se a uma menor sobrevivência livre de progressão, suportando, no seu conjunto, a utilidade desta classificação para estratificação pós-operatória do prognóstico nestes doentes.

Palavras-chave: Tumores neuroendócrinos da hipófise (PitNET); classificação PANOMEN-3; progressão tumoral; prognóstico; análise de sobrevivência.

ABSTRACT

Introduction: Pituitary neuroendocrine tumours (PitNETs) constitute a heterogeneous group of intracranial neoplasms with variable clinical behaviour. Although most follow an indolent course, some tumours exhibit an atypical and aggressive evolution, demonstrating invasive growth, persistence or late recurrence after surgery and medical treatment. It is therefore crucial to identify reliable prognostic tools. The Pituitary Adenoma Nomenclature and Multidimensional Evaluation (PANOMEN-3) classification has recently been proposed as a multidimensional system integrating clinical, biochemical, radiological, and histopathological variables to estimate the risk of morbidity and mortality associated with PitNETs. This study aimed to evaluate the prognostic performance of the PANOMEN-3 classification in a surgical cohort of patients with PitNETs, as well as to analyse the prognostic value of its individual variables.

Methods: Retrospective observational study of a consecutive cohort of adult patients with histologically confirmed PitNETs who underwent pituitary surgery at a tertiary academic centre, Unidade Local de Saúde de Santo António, Porto, between 2017 and 2024. PANOMEN-3 scores were calculated based on clinical, biochemical, radiological and histopathological variables collected from electronic medical records. Progression-free survival was analysed according to PANOMEN-3 classification grades and its individual variables.

Results: 92 patients were included, with a median follow-up time of 69 months (IQR 36). Most tumours were macroadenomas (72.8%). The most frequent histological subtypes were corticotroph (28.3%), gonadotroph (27.2%) and somatotroph (21.7%) PitNETs. According to postoperative PANOMEN-3 grading, 5.4% of patients were classified as grade 0, 27.2% as grade 1, 57.6% as grade 2 and 9.8% as grade 3. Increasing PANOMEN-3 grade was associated with significantly shorter times to tumour recurrence or progression, reflecting a progressive reduction in progression-free survival (log-rank $p < 0.001$). In Cox regression analysis, PANOMEN-3 grade 2 tumours (HR 6.13; 95% CI 2.15–17.51; $p = 0.001$) and grade 3 tumours (HR 15.58; 95% CI 4.32–56.26; $p < 0.001$) showed a significantly higher risk of tumour recurrence or progression.

Conclusions: The PANOMEN-3 classification demonstrated good prognostic performance in this surgical cohort of patients with PitNETs. The risk of tumour recurrence or progression was significantly predicted by higher scores on this scale (grades 2 and 3). These were associated with shorter progression-free survival, collectively supporting the utility of this classification for postoperative prognostic stratification in these patients.

Keywords: Pituitary neuroendocrine tumours (PitNETs); PANOMEN-3 classification; tumour progression; prognosis; survival analysis.

ABBREVIATIONS

CI, Confidence interval

CSF, Cerebrospinal fluid

HR, Hazard ratio

IQR, Interquartile range

MEN1, Multiple Endocrine Neoplasia Type 1

PANOMEN, Pituitary Adenoma Nomenclature and Multidimensional Evaluation

PANOMEN-3, Pituitary Adenoma Nomenclature and Multidimensional Evaluation (version 3)

PitNET, Pituitary neuroendocrine tumour

PFS, Progression-free survival

SD, Standard deviation

ULSSA, Unidade Local de Saúde de Santo António

WHO, World Health Organization

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Pituitary neuroendocrine tumours (PitNETs) represent one of the most frequent benign intracranial neoplasms and encompass a biologically and clinically heterogeneous group of diseases. Population-based imaging and autopsy studies suggest that pituitary tumours may be present in up to 15–17% of adults[1]. However, only a minority become clinically relevant, requiring medical, surgical or radiotherapeutic intervention. When clinically active, PitNETs can lead to significant morbidity through endocrine hypersecretion, local mass effects on surrounding neural structures, and varying degrees of hypopituitarism, often resulting in long-term health impairment[2,3].

Although most PitNETs follow an indolent course, their clinical behaviour varies considerably in terms of invasiveness, recurrence risk and treatment response, which is not always predictable at the time of diagnosis[4]. A substantial proportion of tumours exhibit invasive growth, surgical residue growth or persistent biochemical activity[4,5] , while others recur after an initial period of apparent remission. In addition, an even smaller but clinically relevant subset of these tumours demonstrates aggressive behaviour, characterised by rapid progression and resistance to standard multimodal therapies[2,4,5]. This wide spectrum of outcomes emphasises the importance of reliable prognostic tools capable of identifying patients at increased risk of recurrence or progression early in the disease course. This is of particular relevance for guiding risk-adapted follow-up strategies, enabling closer and more frequent clinical surveillance in higher-risk patients while avoiding unnecessary monitoring in those with a more favourable prognosis.

Historically, prognostic stratification of PitNETs relied primarily on clinical and radiological characteristics, including hormonal activity, tumour size and radiological evidence of invasion. Their utility, however, was largely confined to guiding initial management, with limited ability to predict long-term tumour behaviour[4,5]. With advances in tumour biology, pathological markers such as mitotic activity, Ki-67 proliferation index and *TP53* mutation, together with radiological evidence of cavernous sinus invasion, were increasingly recognised as indicators of aggressive tumour behaviour and poorer clinical outcomes. Nevertheless, these markers showed limited prognostic performance when considered individually[5].

Earlier World Health Organization (WHO) classifications, notably the 2004 edition, relied on the distinction between “typical” and “atypical” adenomas, largely based on proliferative markers such as Ki-67 index and *TP53* mutation[2]. Although this approach represented an attempt to incorporate biological markers into tumour classification, the prognostic value of these features proved to be inconsistent across studies, with variability in cut-off definitions and limited reproducibility.

In this context, an initial effort for integrated prognostic assessment emerged in 2013. Trouillas et al. proposed a clinicopathological classification that combined proliferative markers, including Ki-67, mitotic count and *TP53* mutation, with radiological and pathological evidence of tumour invasion, demonstrating significant prognostic value in predicting recurrence and progression, particularly for invasive tumours[6]. However, the prognostic relevance of individual proliferative markers, specifically Ki-67 and *TP53*, remained inconsistent, contributing to a progressive decline in their isolated use as reliable predictors of clinical outcomes. Additionally, this classification had several limitations, being restricted to surgically treated cases and not incorporating key clinical determinants such as endocrine activity, pituitary dysfunction or genetic predisposition[6].

The 2017 and 2022 World Health Organization (WHO) classifications further refined tumour pathology characterisation by adopting the concept of PitNETs and identifying specific high-risk histological subtypes[7,8]. In particular, the 2017 update abandoned the previous "atypical adenoma" category, reflecting the limited and inconsistent prognostic value of proliferative markers[7]. These updated classifications introduced a more detailed framework based on cell lineage, transcription factors and hormone expression, improving tumour phenotyping and contributing to a more precise histopathological definition[7,8]. Nonetheless, their clinical impact has remained limited, as they do not provide a robust grading or staging system that can adequately predict tumour biology and thus prognosis stratification. In addition, it relies on histopathological evaluation and therefore cannot be applied to a substantial proportion of clinically relevant tumours managed without surgery[7,8].

These limitations highlighted the need for a comprehensive classification framework capable of integrating clinical, radiological, biochemical and pathological information. To address these gaps, the Pituitary Society launched the Pituitary Adenoma Nomenclature and Multidimensional Evaluation (PANOMEN) initiative, which progressed through two earlier iterations before reaching its current form, PANOMEN-3[9,10]. The classification assigns each tumour a grade derived from clinical, biochemical, radiological and histological information, and can be applied both before surgery, using the clinical, biochemical and imaging items alone, and again postoperatively, to include data yielded by histology.

Since its introduction, several retrospective studies conducted in surgically treated cohorts have suggested that higher PANOMEN-3 grades are associated with less favourable clinical outcomes, including increased risk of persistent disease, tumour recurrence or progression, and a greater need for additional therapeutic interventions[10–14] **Erro! A origem da referência não foi encontrada.**. These findings indicate that this classification may provide prognostic information beyond the traditional classification systems and support its potential role in postoperative risk stratification and long-term follow-up planning.

Despite these encouraging results, the available evidence remains limited to retrospective studies in which patient selection, follow-up duration and treatment strategies are highly heterogeneous[10–14]. Thus, further assessment using data from independent cohorts, including smaller, single-centre and real-world series, is warranted to determine the robustness and generalisability of the grading system across different clinical contexts.

In the present study, we evaluated the prognostic performance of the PANOMEN-3 classification in a real-world retrospective cohort of surgically treated patients with PitNETs. Most importantly, we analysed the association between postoperative PANOMEN-3 grade and clinical outcomes, tumour recurrence or progression during follow-up. Additionally, we explored the potential influence of demographic variables and individual PANOMEN-3 domains on progression risk, aiming to further clarify the relevance of this classification in everyday clinical practice.

METHODS

Study design, setting, and study period

We conducted a single-centre, retrospective observational cohort study at our tertiary hospital, Unidade Local de Saúde de Santo António (ULSSA), in Porto, Portugal, including patients with PitNETs who underwent pituitary surgery between January 2017 and December 2024. This study period was selected to ensure inclusion of a contemporary surgical cohort while allowing sufficient postoperative follow-up for meaningful assessment of clinical outcomes, specifically tumour recurrence and progression. The process of patient selection and the reasons for exclusion are summarised in the flowchart presented in **Figure 1**.

Eligible patients were adults (≥18 years) with a diagnosis of PitNET confirmed on surgical histopathology, who underwent pituitary surgery at ULSSA between January 2017 and December 2024. To ensure cohort homogeneity and allow accurate assessment of postoperative outcomes, only the first pituitary surgical intervention performed at ULSSA within the defined study period was considered eligible and a minimum postoperative follow-up of 12 months was required to permit a reliable classification of disease status and assessment of recurrence or progression. Patients who underwent surgery during this timeframe for recurrent disease were excluded.

Other exclusion criteria were insufficient clinical, biochemical, radiological and pathological data to allow calculation of the PANOMEN-3 score. Patients were not excluded based on tumour aggressiveness.

Data collection and variable definition

Data were retrospectively extracted from ULSSA electronic medical records. Sources included endocrinology and neurosurgery outpatient notes, surgical, pathology and radiological imaging reports. Histological tumour subtypes were recorded according to the terminology and diagnostic categories of the 5th edition (2022) WHO Classification of Endocrine and Neuroendocrine Tumours, based on routine diagnostic pathology reports[8]. The analysed subtypes included corticotroph, somatotroph, lactotroph, gonadotroph and plurihormonal PitNETs, as well as non-functioning tumours without specification of cellular lineage and other rare or unclassified histological variants grouped under the category “others”. Clinical and biochemical phenotype, including diagnostic criteria for functioning subtypes and the definition of biochemical control or remission, was

classified according to current consensus guidelines for prolactinomas[15], Cushing's disease[16], acromegaly[17,] clinically non-functioning adenomas[18] and thyrotropin-secreting adenomas[19].

Clinical, radiological and histopathological tumour characteristics were systematically collected from medical records and pathology reports. Radiological variables included tumour size, classified as microadenoma (<10 mm), macroadenoma (≥10 mm) or giant adenoma (≥40 mm), as well as cavernous sinus invasion according to the Knosp classification[20].

Variables were collected at two predefined timepoints corresponding to the last comprehensive preoperative assessment (T0) and the first postoperative reassessment (T1). Preoperative PANOMEN-3 variables were primarily used for baseline cohort characterisation, whereas postoperative PANOMEN-3 classification (T1) was used for prognostic analyses and evaluation of progression risk.

Time to tumour recurrence or progression was recorded individually and defined as the period from the date of surgery until the documentation of a progression event, last clinical evaluation or loss to follow-up, allowing time-to-event analysis despite heterogeneous follow-up durations.

For cohort characterisation, demographic variables including age at diagnosis and sex were recorded descriptively although not incorporated in the score.

PANOMEN-3 variables were systematically collected and scored across predefined domains: tumour phenotype, secretory status, hypopituitarism, tumour size, mass effect, cavernous sinus invasion, residual tumour, aggressive histopathological features and genetic syndromes[10]. The PANOMEN-3 domains, corresponding risk scores, and their applicability before and after surgery are summarised in **Table I**.

For each patient and timepoint, a raw PANOMEN-3 score was calculated by summing the points assigned to all applicable PANOMEN-3 domains, comprising seven domains preoperatively (T0) and nine domains postoperatively (T1). To allow standardised comparison across patients, a corrected score (ratio) was calculated by dividing the raw score by the number of PANOMEN-3 domains assessed, as previously described in the PANOMEN-3 framework[10]. The detailed scoring system and the method used to calculate the corrected score and final grade are presented in **Table II**. Based on this corrected score, patients were classified into four grades reflecting increasing disease severity: grade 0 (score = 0), grade 1 (>0 and <0.3), grade 2 (0.3–0.6) and grade 3 (>0.6).

Outcomes

The primary outcome was tumour progression during follow-up. Progression was defined by the occurrence of one or more of the following events: biochemical evidence of recurrence or progression in functioning tumours, including recurrence after remission or worsening hormonal

hypersecretion in the absence of radiological progression, radiological detection of a new tumour following complete resection and radiological regrowth after partial resection.

Progression-free survival (PFS) was defined as the time from pituitary surgery to the first documented evidence of tumour recurrence or progression.

Disease status at the last follow-up evaluation was classified as disease-free, residual disease or recurrent/progressive disease.

Statistical analysis

All statistical analyses were performed using Stata® Statistical Software, Release 19.0. Continuous variables were summarised as mean ± standard deviation (SD) or median and interquartile range (IQR), according to data distribution, whereas categorical variables were described using absolute and relative frequencies.

PFS was analysed using the Kaplan–Meier method, and differences between survival curves were assessed using the log-rank test. For survival analyses, patients were divided into two age groups (18–64 years and ≥65 years) to compare younger and elderly individuals.

To assess the independent prognostic value of the PANOMEN-3 classification, a multivariable Cox proportional hazards regression model was constructed, including PANOMEN-3 grade as the variable of primary interest and age at diagnosis and sex as standard demographic covariates for adjustment. Age was analysed as a continuous variable and PANOMEN-3 grade as a categorical variable. Due to the small number of patients classified as grade 0 (n = 5), grades 0 and 1 were combined into a single reference category for regression analysis, to improve the robustness and stability of the regression model estimates.

Hazard ratios (HRs) with corresponding 95% confidence intervals (95% CI) were calculated. All statistical tests were two-sided, and p-values < 0.05 were considered statistically significant.

Ethics

This study was approved by the Ethics Committee of ULSSA and by the Instituto de Ciências Biomédicas Abel Salazar, with the reference number 2025.295(253-CAC-253-CE). The research was conducted in accordance with the principles of the Declaration of Helsinki and applicable national legislation governing clinical research. Given the retrospective nature of the study and the use of routinely collected clinical data, the requirement for written informed consent was waived by the Ethics Committee. All data were anonymised prior to analysis and handled in compliance with institutional data protection policies.

RESULTS

Study population characteristics

A total of 92 patients with PitNETs who underwent pituitary surgery at ULSSA and met the study inclusion criteria were considered in the final analysis, after excluding 76 patients. The median follow-up was 69 months (IQR 36) and the median age at diagnosis was 49.5 years (IQR 24). Of the 92 patients included, 41 (45%) were male and 51 (55%) were female.

Regarding tumour size, macroadenomas were the most frequent, accounting for 72.8% of cases, followed by microadenomas (21.7%) and giant adenomas (5.4%).

With respect to histological subtype, corticotroph PitNETs represented the largest proportion (28.3%), followed by gonadotroph (27.2%) and somatotroph PitNETs (21.7%). Less frequent subtypes included lactotroph PitNETs (5.4%), plurihormonal tumours (8.7%) and tumours classified as “others” (8.7%).

Only one patient in the cohort had a documented genetic syndrome (Multiple Endocrine Neoplasia Type 1).

Some aggressive histological variants included in the PANOMEN-3 framework, namely Crooke cell and acidophilic stem-cell tumours, were not represented in this cohort, as no such cases were identified during the study period.

Preoperatively, the majority of patients were classified as PANOMEN-3 grade 2 (58.7%), followed by grade 1 (25.0%) and grade 3 (16.3%). Postoperatively, grade 2 remained the most frequent category (57.6%). Compared with the preoperative distribution, a reduction in the proportion of grade 3 patients was observed (from 16.3% to 9.8%), and 5.4% of patients were newly classified as grade 0. The baseline demographic, clinical and tumour characteristics, as well as postoperative PANOMEN-3 grade, of the study population are summarised in **Table III**.

Kaplan–Meier analysis of progression-free survival

Kaplan–Meier estimates showed a marked stratification of PFS across postoperative PANOMEN-3 grades, with an inverse association between PFS with increasing grades. The median PFS was not reached in grade 0 nor in grade 1. In grade 2 it was 43 months (95% CI 10–not reached), dropping to only 5 months (95% CI 1–not reached) in grade 3. The log-rank test confirmed a significant difference between curves ($p < 0.001$) (**Figure 2**).

Demographic variables were not associated with outcome. Median PFS was 87 months (95% CI 17–not reached) in male patients and was not reached in female patients (95% CI 39–not reached), with a log-rank p of 0.481. A similar pattern was observed by age group, with a median of 87 months (95% CI 35–not reached) in patients aged 18–64 years and a median not reached in those aged ≥65

years (95% CI 1–not reached, log-rank p = 0.214). The corresponding curves overlapped largely throughout follow-up (**Figures 3A and 3B**).

When evaluating the individual PANOMEN-3 domains, significant differences in PFS were observed across several tumour-related variables, particularly secretory status, tumour size, invasion and postoperative residual tumour.

Elevated postoperative hormone levels were associated with a substantial reduction in PFS, with a median of only 6 months (95% CI 5–7), as compared with a median not reached in patients with normal or normalized hormone levels (95% CI 87–not reached, log-rank p < 0.001) (**Figure 4B**).

Tumour size also showed a significant association with outcome (log-rank p = 0.001) **(Figure 4D).** Microadenomas had a median PFS that was not reached (95% CI 87–not reached), whereas macroadenomas had a median of 15 months (95% CI 5–not reached). Only one patient had a giant macroadenoma, precluding a reliable estimate and limiting the interpretation of this subgroup.

Radiological evidence of cavernous sinus invasion (Knosp 3–4) was accompanied by a marked decrease in PFS. The median was 8 months (95% CI 5–49) in patients with cavernous sinus invasion, against a median not reached in those without cavernous sinus invasion (Knosp 0–2) (95% CI 73–not reached, log-rank p = 0.001) (**Figure 4F**).

The presence of postoperative residual tumour emerged as an important predictor of progression risk. Patients with residual disease on postoperative imaging had a median PFS of 24 months (95% CI 7–not reached), considerably lower than that of patients with complete tumour resection, in whom the median was not reached (95% CI 87–not reached, log-rank p < 0.001) (**Figure 4G**).

In contrast, tumour phenotype as defined within the PANOMEN-3 framework did not reach statistical significance (log-rank p = 0.139), despite a numerical tendency towards less favourable outcomes in patients with GH-secreting PitNETs, whose median was 35 months (95% CI 6–not reached). The median was not reached in the reference category, comprising prolactinomas and non-functioning PitNETs (95% CI 43–not reached), and was 87 months (95% CI 7–not reached) in patients with corticotroph adenomas. Hypopituitarism (log-rank p = 0.961), mass effect (log-rank p = 0.258) and aggressive histological features (log-rank p = 0.177) were also not significantly associated with progression-free survival in this cohort (**Figures 4A, 4C, 4E and 4H**).

Taken together, the Kaplan–Meier analysis identified statistically significant differences in progression-free survival across several variables, with the most pronounced gradient observed across PANOMEN-3 grades. A Cox proportional hazards regression model was therefore constructed to further evaluate the association between PANOMEN-3 grade and progression risk after adjustment for standard demographic covariates.

Cox proportional hazards regression analysis

In the multivariable Cox proportional hazards regression model, higher PANOMEN-3 grades were independently associated with an increased risk of progression. Compared with the reference group (grades 0–1), patients classified as grade 2 showed a substantially increased risk (HR 6.13, 95% CI 2.15–17.51, $p = 0.001$), and an even stronger association was observed for grade 3 tumours (HR 15.58, 95% CI 4.32–56.26, $p < 0.001$), corresponding to over a 15-fold increase in the likelihood of tumour recurrence. After adjustment for PANOMEN-3 grade, neither sex (HR 0.88, 95% CI 0.47–1.64, $p = 0.692$) nor age at diagnosis (HR 1.00, 95% CI 0.98–1.02, $p = 0.713$) was independently linked to PFS. The results of the Cox proportional hazards regression model are presented in **Table IV**.

DISCUSSION

The present study assessed the prognostic performance of the PANOMEN-3 classification in a real-world cohort of surgically treated patients with PitNETs. Overall, our findings demonstrate that the PANOMEN-3 classification strongly predicts the risk of tumour progression following pituitary surgery. Both the Kaplan–Meier analysis and the Cox proportional hazards model revealed a clear gradient of risk across PANOMEN-3 categories, with progressively worse progression-free survival observed at higher grades. This supports the classification as a useful postoperative risk stratification tool.

The Kaplan–Meier analysis demonstrated a marked separation of survival curves according to PANOMEN-3 grade. Patients classified as grade 0 or grade 1 maintained the most favourable clinical course, whereas grade 2 tumours showed an intermediate risk of progression and grade 3 tumours experienced the earliest and most pronounced decline in PFS. This progressive separation of survival curves suggests that the PANOMEN-3 score effectively captures the cumulative impact of tumour-related risk factors on long-term clinical outcomes.

Multivariable Cox regression confirmed this pattern. Grade 2 tumours were independently associated with a more than six-fold increase in progression risk compared with the reference group, while grade 3 tumours showed an approximately fifteen-fold increase in risk, after adjustment for demographic covariates. Collectively, these results suggest that the classification captures a marked prognostic gradient and may be useful for identifying patients at especially higher risk of postoperative disease progression.

Beyond the composite score itself, several individual PANOMEN-3 domains demonstrated significant associations with PFS. Persistent hormonal hypersecretion after surgery emerged as an important determinant of clinical outcome, as these patients demonstrated significantly poorer PFS. This observation corroborates the well-established relationship between persistent hormonal hypersecretion and tumour persistence or regrowth in functioning PitNETs[4,5,21]. Likewise, tumour size was significantly associated with progression risk, with macroadenomas demonstrating worse outcomes compared with microadenomas. Larger tumours are associated with a lower likelihood of complete surgical resection and, together with invasive behaviour, have been linked to higher rates of tumour persistence and recurrence[21].

A similar pattern was observed for cavernous sinus invasion. Tumours invading surrounding structures, particularly the cavernous sinus, exhibited significantly lower PFS compared with non-invasive lesions. This finding aligns with previous studies identifying invasion as one of the most relevant predictors of surgical outcome and tumour progression in pituitary tumours. A likely

explanation is that cavernous sinus invasion restricts surgical access and often precludes complete resection, thereby increasing the likelihood of postoperative residual tumour[6].

In line with this observation, the presence of postoperative residual tumour was significantly associated with an increased risk of progression. Patients with residual tumour tissue identified on postoperative imaging demonstrated markedly lower PFS compared with those achieving complete tumour resection. Residual tumour has consistently been associated with an increased risk of tumour regrowth following pituitary surgery and remains a notable determinant of long-term disease control[22].

Conversely, several PANOMEN-3 domains did not demonstrate statistically significant associations with progression-free survival in our cohort. Tumour phenotype, hypopituitarism, mass effect and aggressive histological features, including recognised high-risk histological variants such as immature PIT1-lineage, null cell and sparsely granulated somatotroph tumours, were not significantly associated with progression risk in this analysis. These findings may partly reflect the relatively limited sample size and the heterogeneity of tumour subtypes represented in the cohort but also reinforce the difficulty in finding adequate prognostic factors. Aggressive histological variants are relatively uncommon, and their prognostic impact may be difficult to detect in smaller retrospective series.

Evaluation of the genetic syndrome domain was limited by the presence of only one patient with a documented hereditary syndrome, namely Multiple Endocrine Neoplasia Type 1 (MEN1). Although this syndrome is known to present with larger, more invasive and treatment-resistant pituitary tumours[23], the proportion of syndromic cases in our cohort was considerably lower than the 20–50% prevalence of pituitary involvement reported in dedicated MEN1 series, and far below the rates of up to 65% described in autopsy studies[23]. This discrepancy is unlikely to reflect a true epidemiological difference and probably stems from the retrospective nature of our series, in which neither systematic biochemical screening for concurrent parathyroid or enteropancreatic involvement nor germline testing of the *MEN1* gene was routinely performed, leaving occult syndromic cases unrecognised. As a result, the contribution of the genetic syndrome domain to the prognostic models was inevitably constrained.

Demographic factors appeared to have limited prognostic relevance in the present study. In the multivariable Cox model, neither age at diagnosis nor sex was independently associated with PFS, an observation consistent with the absence of meaningful separation observed in the corresponding Kaplan–Meier curves. Concordant observations have been reported previously, where tumour-specific biological characteristics, notably invasion, proliferation and immunohistological subtype, have repeatedly been identified as the strongest predictors of long-term tumour behaviour in

PitNETs, with demographic variables showing a more modest and inconsistent contribution across studies[4,5,22,24].

The findings observed in our cohort mirror the emerging evidence supporting the prognostic utility of the PANOMEN-3 classification. Recent retrospective analyses conducted in surgically treated cohorts have similarly reported that higher PANOMEN-3 grades are associated with increased risks of tumour persistence, recurrence and progression[10–14]. By integrating clinical, biochemical, radiological and histopathological variables into a single multidimensional framework, PANOMEN-3 attempts to overcome the limitations of earlier classification systems that relied primarily on isolated markers of tumour aggressiveness. Our results add to this evidence by providing additional data on the applicability of this framework in routine clinical practice.

From a clinical perspective, the subgroup of patients classified as PANOMEN-3 grade 3 warrants special attention. It accounts for 16.3% of the cohort preoperatively and 9.8% postoperatively, proportions broadly consistent with those reported in previous validation series, in which grade 3 frequencies ranged from approximately 10% to 16% across surgically treated cohorts[11,13]. Despite representing a minority of cases, these patients showed the most adverse postoperative course in our cohort, with a median PFS of only 5 months and a more than fifteen-fold increase in progression risk relative to the reference group. Moreover, a comparable pronounced prognostic gradient has been described in independent cohorts, with higher PANOMEN-3 grades frequently associated with increased risk of progression, earlier need for additional therapeutic intervention, and poorer long-term outcomes, even after multimodal treatment[11,12]. These data, taken together with our own results, raise the question of whether timely postoperative PANOMEN-3 grading could assist in identifying patients for whom watchful waiting is insufficient. Although our study was not designed to assess the impact of specific postoperative interventions on outcomes, the very short PFS observed in our grade 3 patients is difficult to reconcile with this standard follow-up approach and supports the notion that this subgroup merits early multidisciplinary discussion regarding adjuvant treatment strategies.

Strengths and limitations

The multidimensional structure of the PANOMEN-3 framework addresses a limitation that has long constrained prognostic assessment in pituitary tumours. Earlier models relied predominantly on histopathological features such as proliferative indices or invasion patterns. While these parameters remain clinically relevant, they do not fully capture the complex biological and clinical heterogeneity of PitNETs, a limitation explicitly acknowledged by the Pituitary Society in the rationale for the development of the PANOMEN-3 classification[9,10]. By integrating tumour burden, endocrine activity,

radiological features and histopathological characteristics, PANOMEN-3 provides a more comprehensive representation of overall disease severity.

Several limitations should nevertheless be considered when interpreting the results of this study. First, the retrospective design introduces the potential for selection bias and limits the ability to control for all possible confounding factors. Second, the study was conducted in a single centre and included a relatively modest number of patients, which may have reduced the statistical power for the analysis of certain individual PANOMEN-3 domains and limited the ability to detect potentially small but clinically relevant effects of demographic variables on outcome. Third, some aggressive histological variants incorporated in the PANOMEN-3 framework were not represented in this cohort, preventing assessment of their specific prognostic impact. In addition, syndromic disease was certainly underrepresented in our cohort, precluding a meaningful analysis of the genetic syndrome domain.

Whilst late tumour progressions are known to occur in a subset of PitNETs beyond five years, we consider the follow-up duration appropriate for the primary aim of identifying patients at higher risk of early postoperative recurrence or progression, consistent with the intended use of the PANOMEN-3 classification as a postoperative risk stratification tool rather than as a measure of long-term natural history.

Notwithstanding these limitations, the present study provides additional evidence supporting the prognostic relevance of the PANOMEN-3 classification in surgically treated PitNETs.

Future studies involving larger multicentre cohorts and prospective validation will be important to further clarify the role of PANOMEN-3 in clinical practice. Such studies should aim to further define the potential role of this classification system in guiding personalised postoperative surveillance strategies and informing therapeutic decision-making in patients with PitNETs.

Conclusion

In summary, the present study suggests that PANOMEN-3 grade predicts postoperative outcomes in surgically treated PitNETs. Increasing grade was consistently linked to shorter progression-free survival and a substantially higher probability of tumour progression, with grade 3 patients showing an especially adverse course in the early postoperative period. These results support the role of this classification in postoperative risk stratification and its potential to guide more personalised decisions regarding the intensity and timing of follow-up and adjuvant treatment in higher-risk patients.

REFERENCES

1. Ezzat S, Asa SL, Couldwell WT, et al. The prevalence of pituitary adenomas. Cancer. 2004 Aug 28;101(3):613–9. doi:10.1002/cncr.20412

2. Scheithauer BW, Gaffey TA, Lloyd R V., et al. Pathobiology of Pituitary Adenomas and Carcinomas. Neurosurgery. 2006 Aug;59(2):341–53. doi:10.1227/01.NEU.0000223437.51435.6E

3. Molitch ME. Diagnosis and Treatment of Pituitary Adenomas. JAMA. 2017 Feb 7;317(5):516. doi:10.1001/jama.2016.19699

4. Melmed S, Kaiser UB, Lopes MB, et al. Clinical Biology of the Pituitary Adenoma. Endocr Rev. 2022 Nov 25;43(6):1003–37. doi:10.1210/endrev/bnac010

5. Raverot G, Burman P, McCormack A, et al. European Society of Endocrinology Clinical Practice Guidelines for the management of aggressive pituitary tumours and carcinomas. Eur J Endocrinol. 2018 Jan;178(1):G1–24. doi:10.1530/EJE-17-0796

6. Trouillas J, Roy P, Sturm N, et al. A new prognostic clinicopathological classification of pituitary adenomas: a multicentric case–control study of 410 patients with 8 years post-operative follow-up. Acta Neuropathol. 2013 Jul 12;126(1):123–35. doi:10.1007/s00401-013-1084-y

7. Mete O, Lopes MB. Overview of the 2017 WHO Classification of Pituitary Tumors. Endocr Pathol. 2017 Sep 1;28(3):228–43. doi:10.1007/s12022-017-9498-z

8. Asa SL, Mete O, Perry A, Osamura RY. Overview of the 2022 WHO Classification of Pituitary Tumors. Endocr Pathol. 2022 Mar 15;33(1):6–26. doi:10.1007/s12022-022-09703-7

9. Ho KKY, Kaiser UB, Chanson P, et al. Pituitary adenoma or neuroendocrine tumour: the need for an integrated prognostic classification. Nat Rev Endocrinol. 2023 Nov 17;19(11):671–8. doi:10.1038/s41574-023-00883-8

10. Ho KKY, Fleseriu M, Wass J, et al. A proposed clinical classification for pituitary neoplasms to guide therapy and prognosis. Lancet Diabetes Endocrinol. 2024 Mar;12(3):209–14. doi:10.1016/S2213-8587(23)00382-0

11. Chiloiro S, Giampietro A, Migliore R,et al. The clinicopathological PANOMEN-3 classification predicts pituitary adenoma prognosis: a real-world retrospective single center study of a surgically treated cohort. Pituitary. 2025 Oct 1;28(5). doi:10.1007/s11102-025-01562-9

12. Fookeerah P, Varikatt W, Shingde M, et al. Assessing the clinical application of the PANOMEN 3 classification in a large cohort of pituitary tumors. Eur J Endocrinol. 2025 Jul 1;193(1):56–64. doi:10.1093/ejendo/lvaf122

13. Mondin A, Bovo G, Barbot M,et al. PANOMEN-3 grading score is reliable in predicting pituitary adenoma behavior and prognosis: a single center cohort study. Endocrine. 2025 Sep 1;89(3):834–45. doi:10.1007/s12020-025-04292-x

14. Araujo-Castro M, Menéndez Torre E, Lozano-Aida C, et al. Validation and Limitations of the PANOMEN-3 Predictive Model for Tumor Recurrence and Progression in Pituitary Tumors. J Clin Endocrinol Metab. 2025 Nov 18;110(12):3391–9. doi:10.1210/clinem/dgaf252

15. Petersenn S, Fleseriu M, Casanueva FF, et al. Diagnosis and management of prolactin-secreting pituitary adenomas: a Pituitary Society international Consensus Statement. Nat Rev Endocrinol. 2023 Dec 5;19(12):722–40. doi:10.1038/s41574-023-00886-5

16. Fleseriu M, Auchus R, Bancos I, et al. Consensus on diagnosis and management of Cushing's disease: a guideline update. Lancet Diabetes Endocrinol. 2021 Dec;9(12):847–75. doi:10.1016/S2213-8587(21)00235-7

17. Giustina A, Biermasz N, Casanueva FF, et al. Consensus on criteria for acromegaly diagnosis and remission. Pituitary. 2024 Feb 3;27(1):7–22. doi:10.1007/s11102-023-01360-1

18. Mercado M, Melgar V, Salame L, Cuenca D. Clinically non-functioning pituitary adenomas: Pathogenic, diagnostic and therapeutic aspects. Endocrinol Diabetes Nutr. 2017 Aug;64(7):384–95. doi:10.1016/j.endinu.2017.05.009

19. Beck-Peccoz P, Lania A, Beckers A, Chatterjee K, Wemeau JL. 2013 European Thyroid Association Guidelines for the Diagnosis and Treatment of Thyrotropin-Secreting Pituitary Tumors. Eur Thyroid J. 2013;2(2):76–82. doi:10.1159/000351007

20. Knosp E, Steiner E, Kitz K, Matula C. Pituitary Adenomas with Invasion of the Cavernous Sinus Space. Neurosurgery. 1993 Oct;33(4):610–8. doi:10.1227/00006123-199310000-00008

21. Lu L, Wan X, Xu Y, Chen J, Shu K, Lei T. Prognostic Factors for Recurrence in Pituitary Adenomas: Recent Progress and Future Directions. Diagnostics. 2022 Apr 13;12(4):977. doi:10.3390/diagnostics12040977

22. Raverot G, Dantony E, Beauvy J, et al. Risk of Recurrence in Pituitary Neuroendocrine Tumors: A Prospective Study Using a Five-Tiered Classification. J Clin Endocrinol Metab. 2017 Sep 1;102(9):3368–74. doi:10.1210/jc.2017-00773

23. Vergès B, Boureille F, Goudet P, et al. Pituitary Disease in MEN Type 1 (MEN1): Data from the France-Belgium MEN1 Multicenter Study. J Clin Endocrinol Metab. 2002 Feb;87(2):457–65. doi:10.1210/jcem.87.2.8145

24. Rosinha P, Fonseca L, Amaral C, Ribeiro I, Cardoso MH. Pituitary adenomas in the elderly: Retrospective comparative analysis of clinical/tumor features and surgical data by age group. Medicine. 2022 Sep 30;101(39):e30825. doi:10.1097/MD.0000000000030825

Table I – PANOMEN-3 scoring criteria and corresponding risk scores, with indication of applicability in the preoperative and postoperative settings.

	Risk scores	No surgery	After surgery
Phenotype	0=non-functioning, prolactinoma; 1=acromegaly, thyrotropin-secreting pituitary adenoma; 2=Cushing's disease	X	X
Secretory status	0=normal or biochemically controlled; 1=elevated	X	X
Hypopituitarism	0=absent; 1=partial with no adrenal or arginine vasopressin deficiency; 2=hypopituitarism with adrenal or arginine vasopressin deficiency	X	X
Size	0=micro; 1=macro; 2=giant	X	X
Mass effect	0=absent; 1=visual field defect, cranial nerve palsy, CSF leak	X	X
Invasion	0=absent; 1=Knosp 3 or 4	X	X
Residual tumour	0=absent; 1=present		X
Histopathology	1=high-risk subtype; 1=increased proliferation		X
Genetic syndrome	0=absent; 1=present	X	X

Table II - PANOMEN-3 scoring system used to calculate the corrected score and assign disease severity grades in PitNETs.

Item	Entries	Points	Notes
Phenotype	CD	2	
	GH or TSH secreting PitNETs	1	
	Prolactinoma or NF PitNETs	0	
Secretory status	Elevated	1	
	Normal/normalized	0	
Hypopituitarism	None	0	
	Partial (no hypocortisolism or AVP-D)	1	
	Presence of hypocortisolism or AVP-D	2	
Size	Microadenoma	0	
	Non-giant macroadenoma	1	
	Giant macroadenoma	2	
Mass effect	Absent	0	
	Present	1	Visual defect, cranial nerve palsy
Cavernous sinus invasion	Absent (Knosp 0-1-2)	0	
	Present (Knosp 3-4)	1	
Genetic syndrome	Absent	0	
	Present	1	*AIP, GPR101, MEN1, CDKN1B, PRKAR1A, GNAS*
Residual tumour	Absent	0	
	Present	1	
Aggressive histology	Absent	0	
	Present (at least one feature)	1	High proliferation and/or high risk histotype*

$$\frac{\boldsymbol{Total\ Points}}{\boldsymbol{Number\ of\ items\ used}} = \boldsymbol{Ratio}$$

Ratio	**0**	**0 – 0.3**	**0.3 – 0.6**	**> 0.6**
Grade	**0**	**1**	**2**	**3**

* Increased proliferative activity defined as Ki-67 >10% and/or aggressive histology (either immature PIT1-lineage, null cell or sparsely granulated somatotroph).

Abbreviations: CD, Cushing's disease; GH, growth hormone; TSH, thyroid-stimulating hormone; NF PitNET, non-functioning pituitary neuroendocrine tumour; AVP-D, arginine-vasopressin deficiency; AIP, aryl hydrocarbon receptor-interacting protein; GPR101, G-protein-coupled receptor 101; MEN1, Multiple Endocrine Neoplasia Type 1; CDKN1B, cyclin-dependent kinase inhibitor 1B; PRKAR1A, cAMP-dependent protein kinase A type 1α regulatory subunit; GNAS, guanine nucleotide-binding protein alpha subunit.

Table III - Baseline demographic, clinical and tumour characteristics, including postoperative PANOMEN-3 grade, of the study population of surgically treated patients with PitNETs.

Characteristic	Patients (n=92)
Follow-up time (months), median (IQR)	69 (36)
Age (y), median (IQR)	49.5 (24)
Sex, n (%)	
Male	41 (45)
Female	51 (55)
Tumour size, n (%)	
Microadenoma	20 (21.7)
Macroadenoma	67 (72.8)
Giant adenoma	5 (5.4)
Histological type, n (%)	
Lactotroph	5 (5.4)
Somatotroph	20 (21.7)
Corticotroph	26 (28.3)
Gonadotroph	25 (27.2)
Plurihormonal	8 (8.7)
Others	8 (8.7)
PANOMEN-3 grade preoperatively, n (%)	
1	23 (25.0)
2	54 (58.7)
3	15 (16.3)
PANOMEN-3 grade postoperatively, n (%)	
0	5 (5.4)
1	25 (27.2)
2	53 (57.6)
3	9 (9.8)

Table IV - Cox proportional hazards regression analysis of progression-free survival according to PANOMEN-3 grade and demographic variables.

Variable	HR	95% CI	p-value
Sex	0.88	0.47-1.64	0.692
Age at diagnosis, years	1.00	0.98-1.02	0.713
PANOMEN-3 grade 2	6.13	2.15-17.51	0.001
PANOMEN-3 grade 3	15.58	4.32-56.26	<0.001

Note: PANOMEN-3 grades 0 and 1 were combined and used as the reference category.

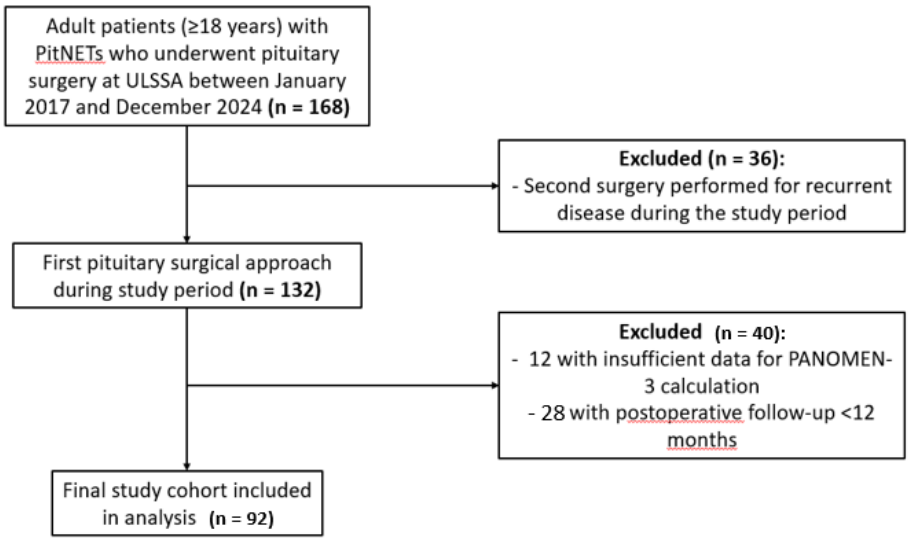

Figure 1 - Flowchart of study population selection. Adult patients with PitNETs undergoing pituitary surgery at ULSSA (2017–2024) were screened for eligibility. Exclusion criteria were recurrent surgery, insufficient data for PANOMEN-3 calculation, and inadequate follow-up.

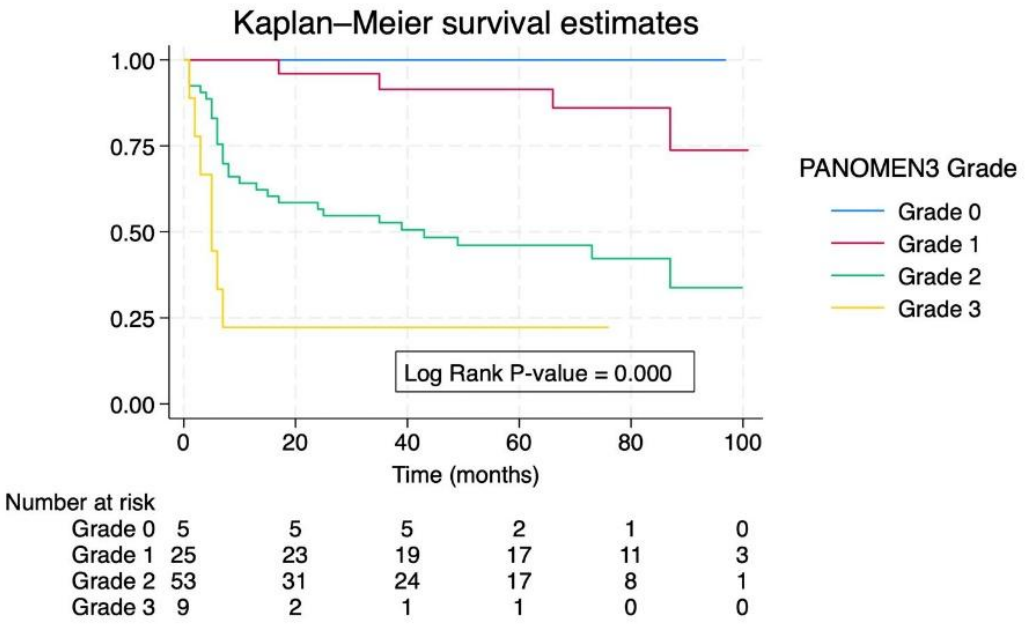

Figure 2 - Kaplan–Meier estimates of Progression-free survival according to postoperative PANOMEN-3 grade.

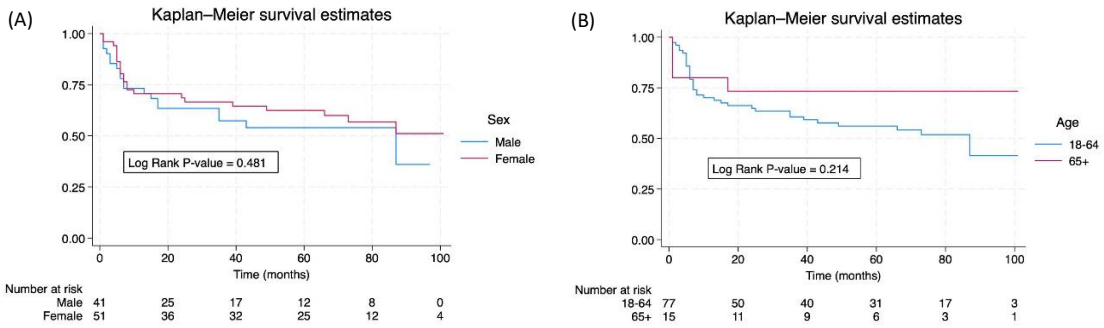

Figure 3 - Kaplan–Meier estimates of progression-free survival according to demographic variables. (A) Sex. (B) Age group.

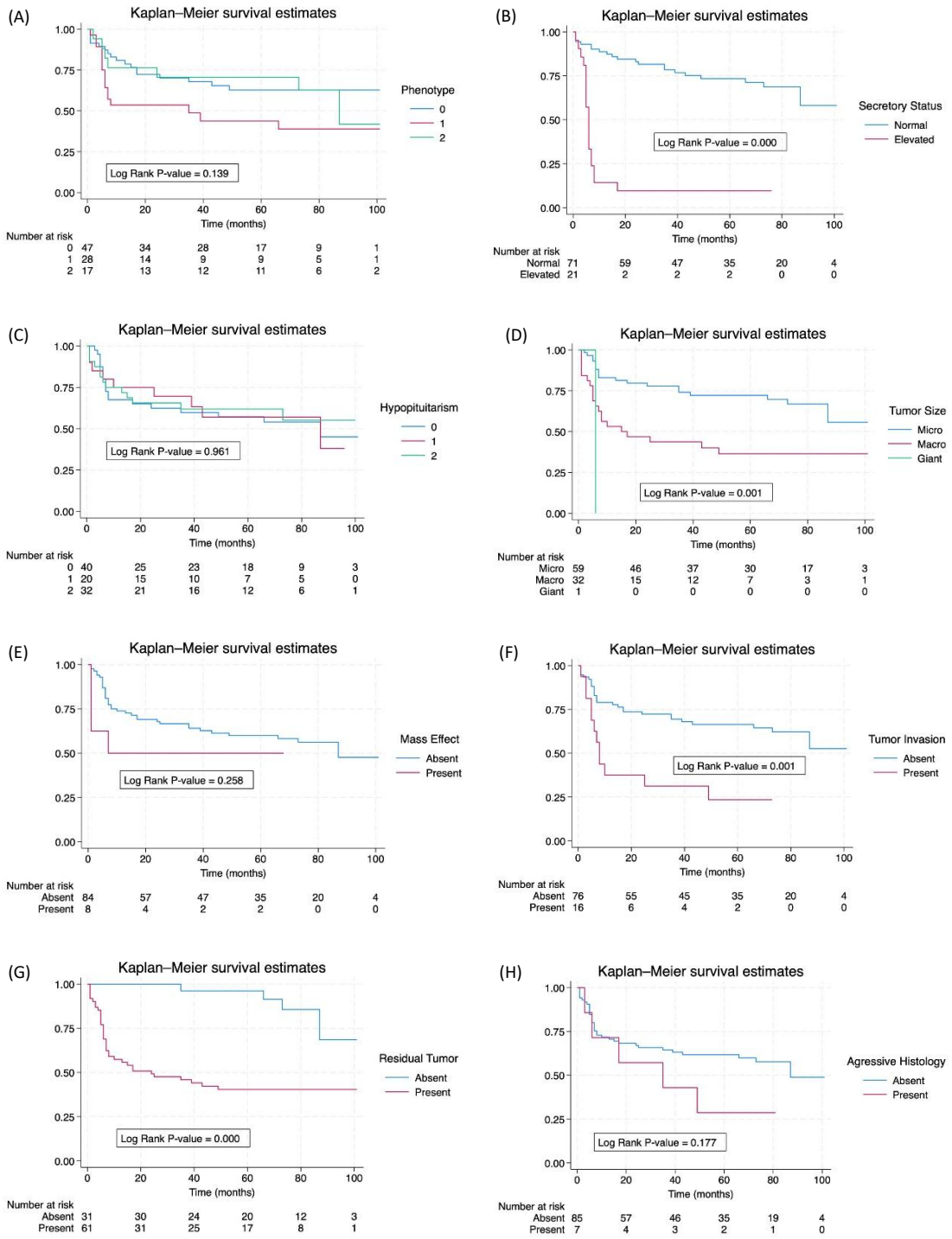

Figure 4 - Kaplan–Meier estimates of progression-free survival according to individual PANOMEN-3 score domains. (A) Tumour phenotype. (B) Hormonal secretory status. (C) Hypopituitarism. (D) Tumour size. (E) Mass effect. (F) Tumour invasion. (G) Postoperative residual tumour. (H) Aggressive histology.

INSTITUTO DE **CIÊNCIAS BIOMÉDICAS ABEL SALAZAR**