diagnostics

MDPI

Article

Association Between Peripheral-Blood Tumor Abnormal Protein and Gastrointestinal Pathological Status: A Retrospective Single-Center Study

Zhi Duan [1,2], Ning Wang [3], Meiyan Wei [1,2], Jiao He [1,2] and Zhen Li [1,2,*]

[1] Department of Pathology, The First Hospital of Changsha/The Affiliated Changsha Hospital of Xiangya School of Medicine, Central South University, Changsha 410005, China; duanzhi910@163.com (Z.D.); 15973181048@163.com (M.W.); freeman2he@gmail.com (J.H.)
[2] Changsha High-Throughput Sequencing Technology Innovation Center for Gastrointestinal Tumor Precision Medicine, The First Hospital of Changsha, Changsha 410005, China
[3] Department of Infectious and Immunology, The First Hospital of Changsha/The Affiliated Changsha Hospital of Xiangya School of Medicine, Central South University, Changsha 410005, China; wang_ning_729@163.com
* Correspondence: linxinchen52@gmail.com

Abstract

Background/Objectives: We examined the association between tumor abnormal protein (TAP) and gastrointestinal pathological status in a symptomatic surgical population. **Methods:** This retrospective single-center study assessed 485 patients tested in 2024 without requiring matched general laboratory measurements. The principal analysis included 259 patients: 63 with malignant and 196 with nonmalignant pathology. Pathology ascertainment extended into 2025 using a 90-day post-TAP window adopted during revision. Analyses used the fixed institutional threshold of 121 μm^2 and included age-and-sex adjustment, bootstrap optimism correction, exploratory chronological assessment, and selection weighting. **Results:** The area under the receiver operating characteristic curve (AUC) was 0.775 (95% CI 0.713–0.834). Sensitivity was 82.5% and specificity 64.3%, with 11 false negatives and 70 false positives. Against 70 benign neoplasms, polyps, or adenomas, AUC was 0.716. Adding TAP increased apparent model AUC from 0.715 to 0.790; the paired fixed-score difference was 0.075 (95% CI 0.024–0.125), and optimism-corrected AUC was 0.780. Later-period model AUCs ranged from 0.677 to 0.728; selection-weighted AUC was 0.775. **Conclusions:** TAP showed modest discrimination with substantial overlap between malignant and nonmalignant disease. Incremental discrimination did not establish clinical utility. Incomplete diagnostic and treatment histories and selective reference availability restrict interpretation to association with pathological status, rather than accuracy in an untreated, previously undiagnosed population. Independent prospective evaluation is required.

Keywords: tumor abnormal protein; gastrointestinal malignancy; pathological status; glycoproteins; internal validation; verification bias

Check for updates

Academic Editor: Ivana Lapić

Received: 22 September 2026
Revised: 10 November 2026
Accepted: 16 November 2026
Published: 25 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Gastrointestinal malignancies remain a major global health burden. The latest GLOBOCAN estimates identify colorectal and stomach cancers among the most frequently diagnosed cancers and major causes of cancer mortality worldwide [1]. Endoscopy, cross-sectional imaging, and histopathology underpin diagnosis, but the practical clinical problem is often not the separation of patients with cancer from healthy volunteers; rather, it is the

distinction between malignancy and symptomatic benign neoplasia, inflammatory disease, or other surgical conditions. Although blood-based assays are increasingly investigated as accessible adjuncts [2], their clinical value depends on how well they perform within the disease spectrum in which they are actually used.

Cancer-associated glycosylation provides a biologically plausible source of circulating biomarkers. Experimental work in colorectal cancer has linked abnormal O-glycosylation to altered Cosmc/T-synthase activity, Tn-antigen expression, and oncogenic cellular behavior [3]. Serum lectin profiling and targeted O-glycoproteomics have also identified colorectal cancer-associated glycan signatures and candidate glycoproteins [4,5]. However, these mechanistic and platform-specific findings should not be interpreted as proof that a routine aggregate-based TAP assay measures a single molecular entity or provides disease-specific discrimination.

Tumor abnormal protein (TAP) is a peripheral-blood assay that quantifies the area of agglomerates associated with abnormal glycoprotein reactivity. Prior gastric-cancer work reported high diagnostic performance when patients with gastric cancer were compared with control participants [6], and a recent colorectal study evaluated TAP alongside fecal syndecan-2 methylation and fecal occult blood testing [7]. Additional investigations have addressed prognostic associations or broader mixed-cancer populations [8,9]. Collectively, these reports support further investigation but leave several clinically important questions unresolved: how TAP performs against symptomatic nonmalignant gastrointestinal disease; whether its association is independent of age and sex; whether apparent performance is optimistic; and how selective availability of histopathology affects estimated accuracy.

These issues are particularly relevant in retrospective surgical datasets because histopathology is not available for every patient, and TAP testing may occur before or after a previous diagnosis or treatment. Restricting analysis to patients with suitable pathological references can introduce selection and partial verification bias, while comparisons against healthy controls may amplify spectrum effects [10,11]. We therefore investigated the association between reported TAP values and episode-relevant gastrointestinal pathological status in a retrospective gastrointestinal-surgery cohort. Our objectives were to characterize discrimination within this selected disease spectrum at a fixed institutional threshold, assess the association after adjustment for age and sex, evaluate incremental discrimination and internal optimism, and examine sensitivity to reference ascertainment, comparator composition, and selection into the analytical cohort. Because prediagnostic and pretreatment sampling could not be established consistently, the study was not intended to estimate diagnostic accuracy in a uniformly untreated, clinically undiagnosed population.

2. Materials and Methods

2.1. Study Design, Setting and Reporting Framework

This retrospective, single-center observational study examined the association between peripheral-blood TAP values and episode-relevant gastrointestinal pathological status among patients evaluated in the Department of Gastrointestinal Surgery at the First Hospital of Changsha, Changsha, Hunan, China. The source population comprised patients undergoing TAP testing during the 2024 calendar year. TAP was the index measurement, and eligible biopsy or resection histopathology provided the reference classification. Relevant STARD 2015 items were used to report participant selection, index measurement, reference assessment, and discrimination estimates [10]; relevant TRIPOD+AI items were considered for the secondary regression model [12]. These reporting frameworks did not establish that TAP sampling preceded diagnosis or treatment in every patient. All model assessments were internal, and no independent external validation cohort was avail-

able. The study was not prospectively registered, and no prospectively registered protocol was available.

2.2. Participants, Record Linkage and Analytic Cohort

The source TAP file contained 497 examination records, of which 496 were dated in 2024. These corresponded to 485 distinct complete hospitalization identifiers with 11 repeat records. The institutional hospitalization identifier remains unchanged across admissions of the same patient. After retaining the earliest examination per identifier, all 485 retained identifiers were nonmissing and distinct. In the original laboratory-linked cohort, 89 patients lacked a laboratory match, 11 had discordant matched names, and 81 exceeded the three-day laboratory-matching interval, leaving 304 patients. Because this restriction was not intrinsically required by the primary TAP–age–sex model, all 485 patients were reconsidered during revision without requiring general laboratory matching.

During revision, 2023–2025 institutional pathology exports were linked using concordant normalized names and the final six digits of hospitalization identifiers. One variant-name discrepancy was resolved separately using the original pathology accession number and concordant identifier and diagnosis information. Other linkage exceptions were documented individually rather than accepting name-only matches automatically. Reference eligibility required relevance to the target gastrointestinal condition in the TAP clinical episode. Following reference reconciliation and application of the ascertainment and eligibility criteria described in Section 2.4, the principal analytical cohort contained 259 participants, with 63 malignant and 196 nonmalignant outcomes. Participants without a suitable reference were not treated as nonmalignant. The reference reconciliation was performed during revision and was not a prospectively planned sampling procedure.

TAP was generally sampled at hospital admission according to routine clinical practice, but this did not imply that every examination preceded biopsy, a previous malignant diagnosis, or treatment. Available pathology receipt and report dates were linked separately from specimen-collection dates. Chart review identified prior cancer surgery, chemotherapy, and stoma-closure presentations in some participants. For one participant, an external malignant pathology diagnosis preceded TAP, while the reported neoadjuvant XELOX course began after sampling; this participant was excluded from the principal analysis because treatment preceded the selected surgical reference. Previous malignancy, prior treatment, current clinical presentation, and the episode-specific pathological outcome were therefore recorded as distinct concepts. Unavailable timing or treatment details were retained as unknown rather than coded as absent. Overall, 226 of the 485 patients were not included in the principal analysis because an eligible reference could not be established or other eligibility criteria were not met, as detailed in the participant flow diagram. This did not establish whether all such patients had undergone pathological examination elsewhere. The principal analysis was therefore interpreted as an association between TAP and pathological status rather than diagnostic accuracy in a uniformly untreated, previously undiagnosed population.

2.3. Index Test and Prespecified Threshold

TAP testing was performed using an abnormal glycan glycoprotein (TAP) detection kit based on agglutination (Zhejiang Ruisheng Medical Technology Co., Ltd., Taizhou, China), following the institutional standard operating procedure. Agglomerates were examined using a Leica DM3000 LED microscope (Leica Microsystems, Wetzlar, Germany), and their images were acquired and areas measured using a medical microscopic image analysis system. Venous whole blood was mixed thoroughly, and two thick smears were prepared using 25 μL per slide. The operating procedure specified an ambient temperature of 18–25 °C

and relative humidity of 45–60%, with protection from dust, airflow, direct light, and vibration. Reagent was equilibrated to the working environment and prepared according to the venous-blood procedure. Three approximately 50-μL drops were applied to each slide, producing six reaction spots. Slides were examined after the reaction was complete and spots had dried; the specified reaction time was approximately 2 h, within a 2 ± 1 h operating range under the prescribed conditions. The spots were systematically scanned using a 4× or 10× objective, and suspected agglomerates were identified, photographed, and measured using a 40× objective. Typical agglomerates were differentiated morphologically from glass, squamous debris, pollen, and other contaminants; the procedure specified water-and-coverslip examination where differentiation was uncertain. TAP results were obtained retrospectively from reports generated during routine clinical care. The two slides and six reaction spots described above formed part of a single examination according to the institutional SOP and did not constitute independent duplicate assays. No additional testing or repeat reading was performed for this retrospective study.

The analysis used the reported TAP agglomerate-area value rather than assuming that every report represented the geometric area of a single imaged particle. The venous-blood procedure specified a twofold area correction. It also defined a valid-particle criterion at approximately 81 μm^2: a single valid particle contributed its measured value, whereas two or more valid particles were summarized using 0.62 times their cumulative area, with an overriding single-particle rule when a qualifying particle reached 225 μm^2. When no particle qualified, the maximum typical-particle area was reported. The representative report image was the maximum typical particle, which need not reproduce a multiple-particle composite report value. Values ≥121 μm^2 were classified as elevated for this study; this report-level threshold is distinct from the particle-selection criterion. Illustrative TAP agglomerate micrographs are provided in Figure S1.

2.4. Histopathological Reference Standard and Manual Adjudication

The reference classification distinguished malignant from nonmalignant gastrointestinal pathological outcomes relevant to the TAP-testing episode. Linked biopsy or resection reports were independently reviewed by two physicians at attending level or above without reference to TAP results. Disagreements were resolved by a third physician, and the adjudicated diagnostic category was used as the reference classification. The final revised pathological classifications, including newly added, excluded and reclassified records, were also independently checked by another physician without reference to TAP results. Nonmalignant findings comprised benign gastrointestinal neoplasms, polyps, or adenomas; inflammatory or infectious disease; hemorrhoids; and isolated other or precancerous diagnoses. For patients with multiple pathology reports, only reports relevant to the target gastrointestinal or abdominal condition during the clinical episode in which TAP was measured were considered eligible. The highest-grade lesion among these eligible reports was used to determine the final patient-level reference diagnosis, and each patient was counted only once.

For the principal analysis, eligible references were sought within 90 days after TAP sampling; 60- and 180-day windows were examined in sensitivity analyses. These windows were adopted during revision after some follow-up records had already been examined and were not prospectively prespecified. Temporal proximity alone did not establish eligibility: patient identity, correspondence to the target lesion or a clinically continuous episode, and intervening anticancer treatment were also considered. Historical malignancy alone was not used to define the current outcome, and stoma-related changes or unrelated benign findings were not interpreted as evidence excluding recurrent malignancy. Records with

insufficient diagnostic certainty or unconfirmed gastrointestinal origin were excluded from the principal analysis.

For the study's binary classification, intramucosal carcinoma, gastrointestinal stromal tumors, and neuroendocrine tumors or carcinomas were coded as malignant. GIST risk categories were not used to subdivide the binary outcome. The author-confirmed low-grade appendiceal mucinous neoplasm was coded as nonmalignant. Non-gastrointestinal primary tumors and the solitary fibrous tumor of intermediate biological behavior were excluded from the principal binary analysis. These coding decisions were distinguished from the separate assessment of whether each report was relevant to the TAP-testing episode.

For patients with tumor-related pathology, a structured manual review captured anatomical site, histological type, specimen type, differentiation, pT and pN categories, lymph-node counts, lymphovascular invasion, and perineural invasion when documented. Biopsy specimens and tumor types for which pathological TNM staging was not applicable were not assigned an artificial stage. Blank or inapplicable entries were retained as missing or not applicable. Formal pM status was unavailable and was not analyzed. TAP testing and histopathological diagnosis were performed by different personnel. TAP assays were performed by laboratory technical staff, whereas histopathological diagnoses were made by pathologists. The two services used separate reporting systems that could not directly retrieve results from each other. These arrangements provided operational separation, although they did not establish complete masking to information available through clinical histories or test requests.

2.5. Covariates and Missing-Data Handling

Candidate clinical variables included age, sex, discharge-diagnosis category, C-reactive protein (CRP), albumin, hemoglobin, neutrophil-to-lymphocyte ratio (NLR), and length of stay. Discharge diagnoses were obtained from hospital discharge records and inpatient medical-record face sheets. Clinical categories were taken directly from the existing disease-category field in the source dataset, rather than reassigned during statistical analysis. Each participant retained the category associated with the linked admission record; multiple diagnoses were not separately ranked or reselected during analysis. The eight categories are presented in Table S1. These discharge-based categories were distinct from the episode-specific histopathological reference classification and could incorporate information obtained during pathological assessment. Accordingly, discharge-diagnosis category was included only in an additional selection-weighting sensitivity analysis restricted to participants with available discharge-category data. Age, sex, and TAP were complete in the pathology-verified cohort and constituted the primary multivariable model. Missingness was defined from blank or explicitly unavailable fields after analyte-specific review; no clinically plausible numeric measurement was recoded as missing solely because it equaled a particular number. CRP was available from the original laboratory-linked dataset for 123 of the 259 participants in the principal analysis and was used only in a labeled complete-case exploratory analysis. No values were imputed for the principal model or pathology subgroups. No index-test result was classified as indeterminate; patients without an eligible reference-standard result were excluded from the principal association analysis but retained in descriptions of the source population and reference availability.

2.6. Statistical Analysis

Continuous variables were summarized as median and interquartile range (IQR), and categorical variables as count and percentage. Pathological categories with small sample sizes were summarized descriptively. Subgroup and sensitivity analyses were considered exploratory.

Following reconciliation of the reference-standard records, the principal association analysis compared 63 patients with histopathologically confirmed gastrointestinal malignancy with 196 patients with nonmalignant pathology relevant to the evaluated clinical episode. An exploratory stricter comparison included the same 63 malignancies and 70 patients with benign gastrointestinal neoplasms, polyps, or adenomas; inflammatory and infectious conditions were excluded from this comparator group. Discrimination was assessed using receiver operating characteristic curves and the area under the curve (AUC), with percentile 95% confidence intervals estimated from 1500 patient-level nonparametric bootstrap resamples. Sensitivity, specificity, positive predictive value, and negative predictive value were calculated at the institutional elevated-TAP threshold of ≥121 μm^2, with Wilson 95% confidence intervals. This threshold was not optimized using the study outcomes. Because all eligible 2024 records were included retrospectively, no prospective sample-size calculation was performed.

Two logistic-regression models were fitted in the same histopathologically verified participants: a baseline model containing age and sex, and an extended model additionally containing TAP. TAP and age were entered as continuous variables per 10 μm^2 and per 10 years, respectively. Model complexity was restricted because only 63 malignant events were available. Odds ratios, Wald 95% confidence intervals, and p values were reported. The complete coefficients and prediction equation for the extended model are provided in Table S2.

The incremental discrimination associated with TAP was assessed as the difference in AUC between the extended and baseline models. Paired patient-level bootstrap resampling with 3000 resamples was used to estimate a percentile 95% confidence interval for the AUC difference while holding the originally fitted prediction scores fixed. This interval is conditional on the fitted scores and does not fully capture model-development uncertainty or establish independently validated incremental benefit. Internal optimism was estimated separately for both models using 3000 bootstrap resamples, comparing each bootstrap-fitted model's AUC in its bootstrap development sample with its AUC in the original sample. Mean optimism was subtracted from the corresponding apparent AUC. Calibration was assessed using the Brier score, calibration intercept and slope, and observed versus predicted malignancy proportions. Calibration-in-the-large was estimated using the logit of the predicted probability as an offset, whereas the calibration slope was estimated by regressing the outcome on this logit with an intercept included.

Chronological internal assessment used patients tested through 31 August 2024 for model development and later patients for evaluation. Because this cutoff was selected after considering event counts, sensitivity analyses repeated the assessment using alternative month-end cutoffs from June through September 2024. Confidence intervals for later-period model AUCs were estimated using 2000 bootstrap resamples, independently resampling the development and later-period cohorts and refitting the model in each development resample. At each chronological split, confidence intervals for later-period calibration measures were estimated using 1000 bootstrap resamples, holding the earlier-period fitted model fixed. These conditional intervals do not incorporate all uncertainty arising from model development. All chronological analyses were post hoc assessments within the same center rather than independent external validation.

Selection and reference availability were examined by comparing patients included in and excluded from the principal analysis. The principal selection-weighting analysis estimated the probability of inclusion in the analytical cohort using penalized logistic regression containing TAP, age, and sex. A separate sensitivity analysis additionally included discharge-diagnosis category and compared models with and without this variable within the same subset with available discharge-category data, because discharge diagnosis could

incorporate information obtained after pathological assessment. Imaging and endoscopic information suitable for characterizing the preceding clinical assessment was unavailable and could not be included. Both models used a ridge penalty of 1 on the non-intercept coefficients, with non-intercept predictors standardized before fitting. Stabilized inverse-probability weights were calculated, with estimated probabilities restricted to 0.05–0.95 and weights truncated at the 1st and 99th percentiles among included participants. Weighted AUC, threshold sensitivity and specificity, and effective sample size were calculated. For the principal weighted analysis, percentile 95% confidence intervals for AUC were estimated using 2000 bootstrap resamples of the 482 candidate patients remaining after exclusion of the three non-target or intermediate tumors, re-estimating the selection model, stabilization, and weight truncation within each resample. The discharge-category sensitivity analysis was reported as exploratory point estimates without bootstrap confidence intervals. The weighting analysis relied on conditional exchangeability given the measured covariates [13] and could not eliminate bias from unmeasured selection factors. Because inclusion depended on both reference availability and eligibility, these weights were not interpreted as estimating a pure probability of pathological verification. Patients without suitable histopathology were not classified as nonmalignant.

Exploratory complete-case analyses assessed adjustment for inflammation by adding $\log(1 + \mathrm{CRP})$ to the TAP–age–sex model. Models with and without CRP were compared within the same complete-case subset. Additional exploratory comparisons used benign gastrointestinal neoplasms, polyps, or adenomas and inflammatory or infectious disease as separate comparator groups. The adenocarcinoma-restricted analysis included gastric signet-ring cell carcinoma; alternative histological definitions were examined in sensitivity analyses. Anatomical distributions were summarized descriptively. Sensitivity analyses also examined 60- and 180-day reference-ascertainment windows and the inclusion of the metastatic adenocarcinoma with suggested but unconfirmed intestinal origin. Incomplete ascertainment of previous diagnoses and treatment prevented these analyses from establishing a uniformly prediagnostic and untreated population. Statistical analyses were performed using Python 3.12.14.

3. Results

3.1. Study Population and Pathological Spectrum

Among the 485 patients assessed for eligibility, 259 were included in the principal analysis, of whom 146 (56.4%) were male and the median age was 60 years. The analytical cohort included 63 gastrointestinal malignancies (24.3%) and 196 nonmalignant diagnoses: 70 benign neoplasms/polyps/adenomas, 107 inflammatory or infectious conditions, 13 hemorrhoidal diagnoses, one high-grade intraepithelial neoplasia/precancerous lesion, and five other nonmalignant diagnoses (Table 1; Figure 1). The malignant group had a median age of 66 years and included 38 men; the nonmalignant group had a median age of 57 years and included 108 men.

Of the 63 malignancies, 48 were colorectal, 12 were gastric, two involved the small bowel or duodenum, and one had author-confirmed gastrointestinal origin without a specified anatomical site. Fifty-five cases were classified within the broad adenocarcinoma group, including gastric signet-ring cell carcinoma. One additional intramucosal carcinoma without a specified histological subtype was retained in the malignant group but included in the adenocarcinoma-restricted comparison only in a sensitivity analysis. Reference-specimen receipt and report-issuance intervals are summarized in Table S3.

Table 1. Pathology-Defined Disease Spectrum and TAP Distribution.

Pathological Category	*n*	TAP Area, Median (IQR), μm²	Elevated TAP, *n* (%)
Gastrointestinal malignancy	63	133.1 (128.6–139.8)	52 (82.5%)
Benign neoplasm/polyp/adenoma	70	120.0 (115.5–132.4)	33 (47.1%)
Inflammatory/infectious disease	107	117.1 (107.6–128.6)	33 (30.8%)
Hemorrhoids	13	113.9 (106.9–116.5)	1 (7.7%)
Precancerous/high-grade intraepithelial lesion	1	120.4 (120.4–120.4)	0 (0.0%)
Other nonmalignant diagnoses	5	127.5 (100.4–133.1)	3 (60.0%)

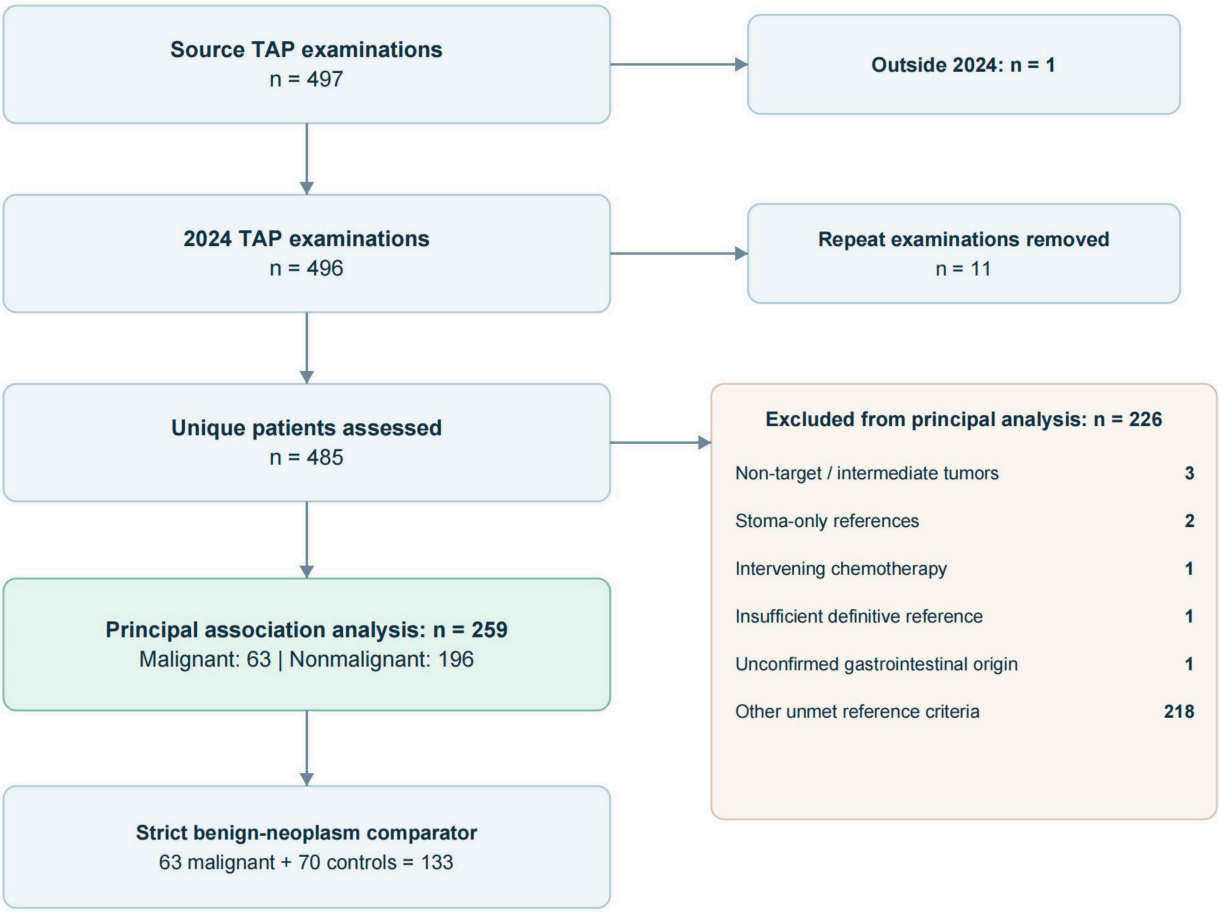

Figure 1. Participant selection and the principal analytical cohort. After restricting TAP records to 2024 and retaining the earliest examination per patient, 485 patients were assessed; 259 met eligibility criteria (63 malignant and 196 nonmalignant). General-laboratory matching was not required. Exclusion does not necessarily indicate absence of pathological examination.

Elevated TAP was defined as a reported TAP value ≥ 121 μm², using the institutional threshold established before analysis. Sparse categories were summarized descriptively. The "other nonmalignant diagnoses" category includes the author-confirmed low-grade appendiceal mucinous neoplasm coded as nonmalignant for this study.

3.2. Tap Distribution and Discrimination in the Principal Analysis

Median TAP was 133.1 (128.6–139.8) μm² in malignant disease and 118.2 (108.7–130.2) μm² in all nonmalignant disease. TAP distributions overlapped across malig-

nant and nonmalignant pathological categories, including benign neoplasms and inflammatory or infectious disease (Table 1; Figure 2).

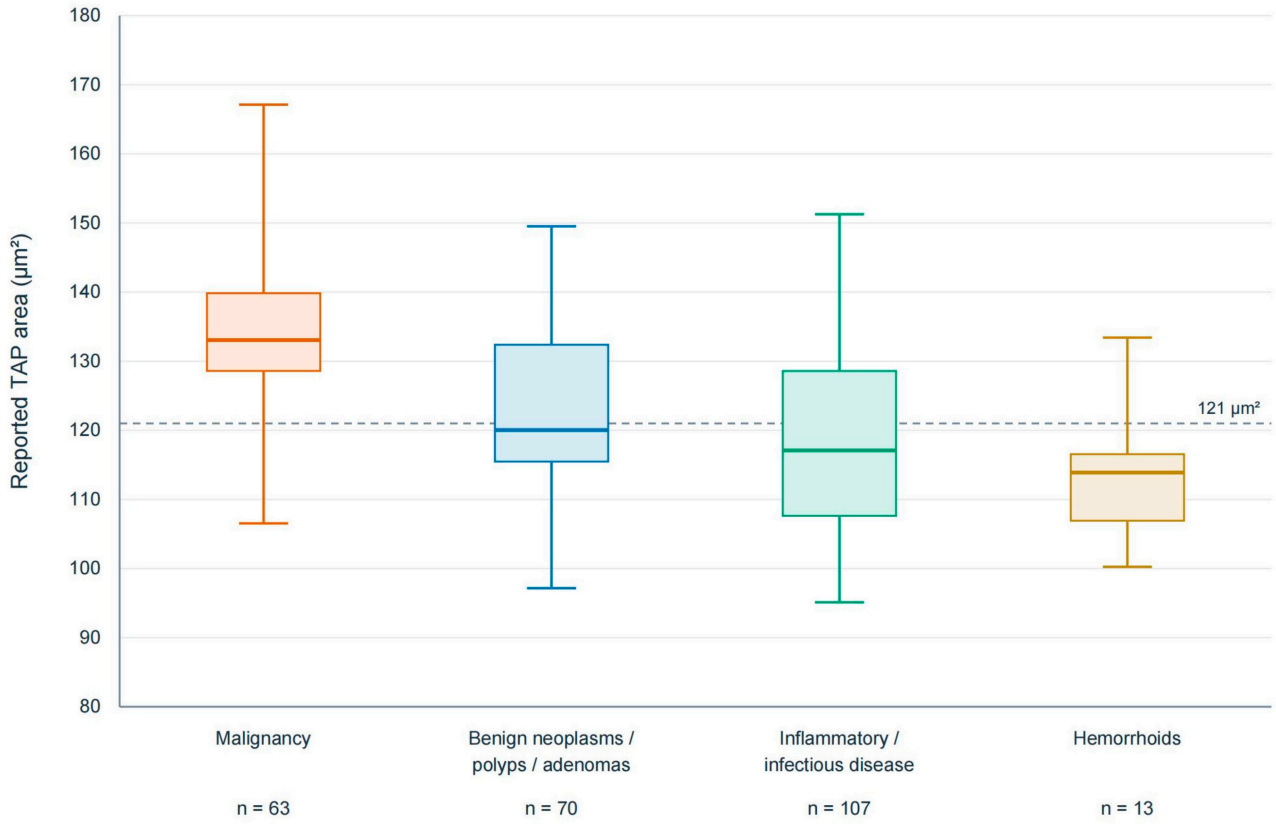

Figure 2. TAP distributions across pathological categories. Reported TAP values are shown for malignancy ($n = 63$), benign neoplasms/polyps/adenomas ($n = 70$), inflammatory or infectious disease ($n = 107$), and hemorrhoids ($n = 13$). Boxes indicate the interquartile range, central lines indicate medians, and whiskers extend to the observed minimum and maximum. The dashed line marks the fixed institutional threshold of 121 µm^2. One high-grade intraepithelial/precancerous lesion and five other nonmalignant diagnoses are summarized in Table 1 and omitted from this plot.

For malignancy versus all nonmalignant diagnoses, TAP yielded an AUC of 0.775 (95% CI 0.713–0.834). At 121 µm^2, there were 52 true positives, 126 true negatives, 70 false positives, and 11 false negatives. Sensitivity was 82.5% (95% CI 71.4–90.0%), specificity was 64.3% (95% CI 57.4–70.7%), PPV was 42.6%, and NPV was 92.0% (Table 2). Predictive values apply to the observed 24.3% malignancy prevalence within the selected analytical cohort and should not be extrapolated to population screening.

When the nonmalignant comparison was restricted to the 70 patients with benign gastrointestinal neoplasms, polyps, or adenomas, the AUC decreased to 0.716 (95% CI 0.628–0.806). Sensitivity remained 82.5%, whereas specificity fell to 52.9%. Elevated TAP was observed in 52 of 63 malignant cases (82.5%) and 33 of 70 benign-neoplasm controls (47.1%), demonstrating clinically important overlap between malignant and benign proliferative disease.

All threshold-based values correspond to the fixed institutional cutoff of ≥121 µm^2. The later temporal cohort comprised patients tested after 31 August 2024 and was used for chronological internal assessment, not independent external validation. The cutoff was not derived from this cohort. Confidence intervals for sensitivity, specificity, and predictive values were calculated using the Wilson method.

Table 2. TAP Discrimination and Fixed-Threshold Performance in the Principal, Strict-Comparator, and Chronological Cohorts.

Metric	Principal Analysis	Benign Comparator	Later Temporal Cohort
Population, *n*	259	133	122
Malignant/nonmalignant, *n*	63/196	63/70	28/94
TAP-only AUC (95% CI)	0.775 (0.713–0.834)	0.716 (0.628–0.806)	0.729 (0.629–0.820)
Sensitivity, % (95% CI)	82.5 (71.4–90.0)	82.5 (71.4–90.0)	71.4 (52.9–84.7)
Specificity, % (95% CI)	64.3 (57.4–70.7)	52.9 (41.3–64.1)	66.0 (55.9–74.7)
Positive predictive value, % (95% CI)	42.6 (34.2–51.5)	61.2 (50.5–70.8)	38.5 (26.5–52.0)
Negative predictive value, % (95% CI)	92.0 (86.2–95.5)	77.1 (63.5–86.7)	88.6 (79.0–94.1)

3.3. Age- and Sex-Adjusted Model and Bootstrap Internal Validation

The age-plus-sex model alone had an apparent AUC of 0.715 (95% CI 0.646–0.784), compared with 0.790 (95% CI 0.726–0.845) when TAP was added. The apparent AUC difference was 0.075 (paired fixed-score bootstrap 95% CI 0.024–0.125). This interval is conditional on the fitted scores and does not constitute independent validation of incremental benefit. In the three-predictor model, each 10-μm^2 increment in TAP was associated with an adjusted OR of 2.11 (95% CI 1.51–2.94; $p < 0.001$). Neither male sex (OR 1.70, 95% CI 0.87–3.30; $p = 0.120$) nor age per 10 years (OR 1.20, 95% CI 0.91–1.57; $p = 0.191$) reached statistical significance after simultaneous adjustment (Table 3; Figure 3; Table S4).

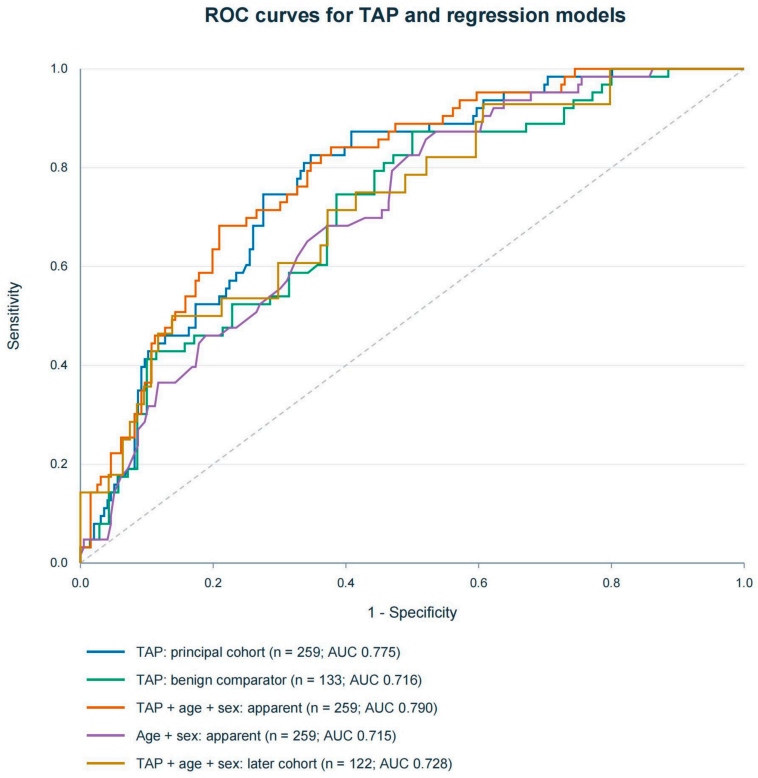

Figure 3. Receiver operating characteristic curves for TAP and regression models. Curves show TAP in the principal cohort (AUC 0.775), TAP against benign neoplasms/polyps/adenomas (0.716), the apparent TAP–age–sex model (0.790), the apparent age–sex model (0.715), and the later-period TAP–age–sex model (0.728). Apparent performance used 259 patients; chronological assessment used 137 earlier patients for development and 122 later patients for evaluation. The diagonal indicates chance discrimination.

Table 3. Multivariable Associations in the Primary and CRP-Complete Models.

Analysis	Predictor	OR	95% CI	*p*
Primary ($n = 259$)	TAP per 10 μm^2	2.11	1.51–2.94	<0.001
Primary ($n = 259$)	Age per 10 years	1.20	0.91–1.57	0.191
Primary ($n = 259$)	Male sex	1.70	0.87–3.30	0.120
CRP-complete ($n = 123$)	TAP per 10 μm^2	1.70	1.06–2.73	0.028
	Age per 10 years	1.29	0.87–1.92	0.205
	Male sex	2.25	0.87–5.82	0.095
	log(1 + CRP)	1.19	0.89–1.59	0.246

The CRP-complete analysis contained 123 participants and 30 malignant events and was exploratory. Within this same subset, the TAP association before adding CRP was OR 1.65 (95% CI 1.04–2.63; $p = 0.034$), compared with OR 1.70 (95% CI 1.06–2.73; $p = 0.028$) after adjustment. These findings do not exclude residual inflammatory confounding or bias from selective CRP availability.

Across 3000 bootstrap resamples, mean optimism in the adjusted-model AUC was 0.010, giving an optimism-corrected AUC of 0.780. The apparent Brier score was 0.150. These internal assessments do not establish external validity; chronological performance and calibration are reported below (Figures 3 and 4).

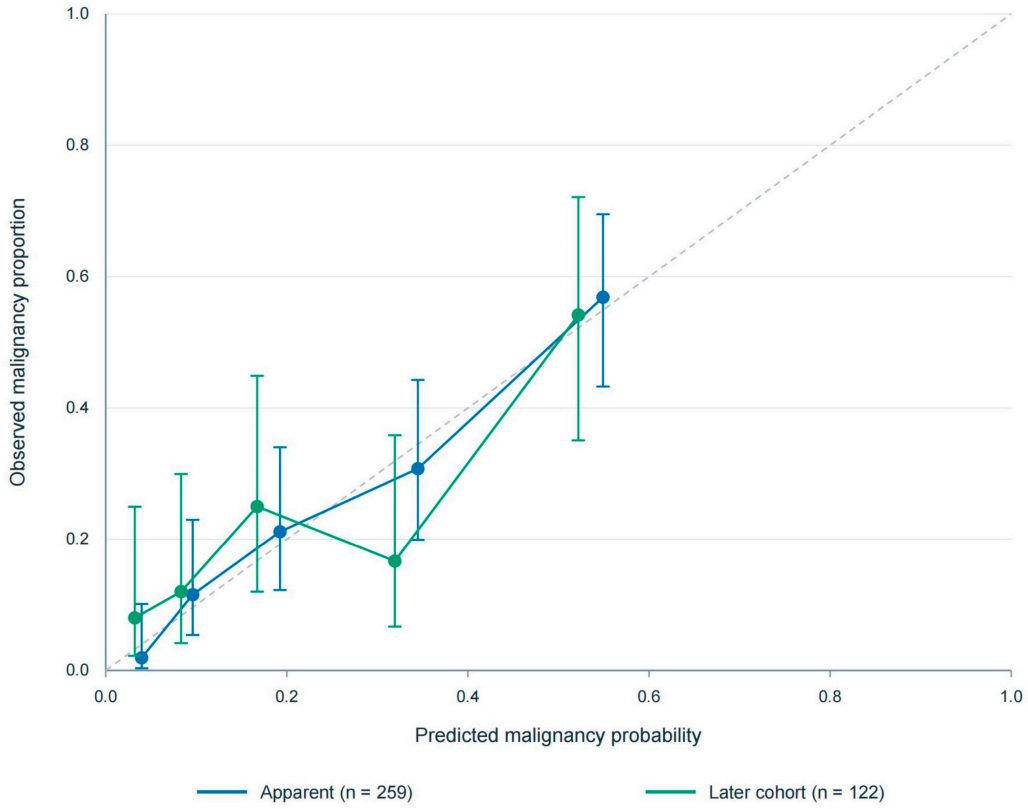

Figure 4. Calibration of the TAP–age–sex model. Points show mean predicted probabilities and observed malignant proportions in five approximately equal-sized groups in the principal ($n = 259$) and later ($n = 122$) cohorts. Bars show Wilson 95% confidence intervals; the diagonal indicates ideal calibration. Later-cohort predictions used the model fitted to 137 earlier patients. Connecting lines are visual guides.

3.4. Chronological Internal Assessment

Using the illustrative August 31 cutoff, the earlier cohort included 137 patients and 35 malignancies, and the later cohort included 122 patients and 28 malignancies. The three-predictor model achieved an AUC of 0.818 in development and 0.728 (95% CI 0.597–0.821) in the later cohort. Later-cohort TAP-only AUC was 0.729; threshold sensitivity and specificity were 71.4% and 66.0%, respectively. Temporal calibration slope was 0.704 (95% CI 0.376–1.155), calibration-in-the-large estimated using an offset was 0.053 (95% CI −0.463 to 0.525), and the Brier score was 0.158 (95% CI 0.120–0.197) (Figure 4). Across the four month-end cutoffs, later-period samples contained 92–165 patients and 16–40 malignancies; adjusted-model AUCs ranged from 0.677 to 0.728 and TAP-only AUCs from 0.691 to 0.753, with wide confidence intervals (Table S5). No cutoff was interpreted as confirmatory or external validation.

3.5. Selection into the Analytical Cohort and Inverse-Probability Weighting

Among the 485 patients assessed, 259 were included in the principal analysis and 226 were not included. Median TAP values were 120.4 and 120.2 μm^2, respectively; median ages were 60 and 61 years, and the proportions of men were 56.4% and 56.6% (Table 4). Eligible references for the principal analysis were available for 174 of 304 patients in the original laboratory-linked cohort, 52 of 89 without a laboratory match, two of 11 with discordant matched names, and 31 of 81 outside the original three-day laboratory-matching interval (Table S6). These proportions reflect both reference availability and analytical eligibility and should not be interpreted as pure pathological verification rates.

Table 4. Characteristics of Patients Included in and Excluded from the Principal Analysis.

Characteristic	Included (n = 259)	Excluded (n = 226)
Age, years	60.0 (47.0–69.0)	61.0 (46.0–72.0)
TAP area, μm^2	120.4 (113.5–133.1)	120.2 (110.7–134.1)
Male sex	146 (56.4%)	128 (56.6%)
Elevated TAP at 121 μm^2	122 (47.1%)	104 (46.0%)

Using a selection model containing TAP, age, and sex without discharge diagnosis, the weighted TAP AUC was 0.775 (95% CI 0.714–0.838), compared with 0.775 in the unweighted principal analysis. Weighted sensitivity and specificity at 121 μm^2 were 81.8% and 65.1%, respectively, with an effective sample size of 258.5. In the original 304-patient laboratory-linked subset, which contained 174 participants eligible for the principal analysis, weighted AUCs were 0.728 without discharge-diagnosis category and 0.737 with this category included. These subset analyses were exploratory point estimates without bootstrap confidence intervals and were not directly comparable with the full-cohort weighted estimate. The similarity of weighted and unweighted estimates does not exclude residual selection bias arising from unmeasured factors, including clinical suspicion, imaging, and endoscopic findings (Table S7).

Continuous variables are presented as median (IQR), and categorical variables as n (%). Exclusion from the principal analysis does not imply that no pathological examination was performed. General laboratory measurements were not required for inclusion; CRP was examined separately in the available complete-case subset.

3.6. Sensitivity and Exploratory Clinicopathological Analyses

Against inflammatory or infectious pathology alone, TAP achieved an AUC of 0.802 (95% CI 0.738–0.865; 63 malignant and 107 nonmalignant cases). Restricting malignant disease to broad adenocarcinoma, including gastric signet-ring cell carcinoma, yielded an

AUC of 0.720 (95% CI 0.625–0.808; 55 malignant cases and 70 benign-neoplasm, polyp, or adenoma controls). Adding the untyped intramucosal carcinoma to this restricted analysis yielded an AUC of 0.724 (95% CI 0.634–0.810). The lower discrimination against benign proliferative lesions reinforces the importance of comparator composition rather than supporting a universal diagnostic cutoff (Table 5; Figure 5).

Table 5. TAP Discrimination across Comparator and Reference-Ascertainment Sensitivity Analyses.

Analysis	Malignant/Nonmalignant, *n*	TAP AUC (95% CI)
Malignancy versus benign neoplasms, polyps, or adenomas	63/70	0.716 (0.628–0.806)
Malignancy versus inflammatory or infectious disease	63/107	0.802 (0.738–0.865)
Broad adenocarcinoma, including gastric signet-ring cell carcinoma, versus benign neoplasms, polyps, or adenomas	55/70	0.720 (0.625–0.808)
Same comparison additionally including untyped intramucosal carcinoma	56/70	0.724 (0.634–0.810)
60-day reference-ascertainment window	62/195	0.775 (0.712–0.831)
180-day reference-ascertainment window *	63/196	0.775 (0.713–0.834)
Including metastatic adenocarcinoma with suggested intestinal origin	64/196	0.778 (0.714–0.838)

AUC confidence intervals were estimated by nonparametric bootstrap resampling. * The 180-day window yielded the same participants as the principal 90-day analysis; the principal-analysis estimate and confidence interval are therefore shown.

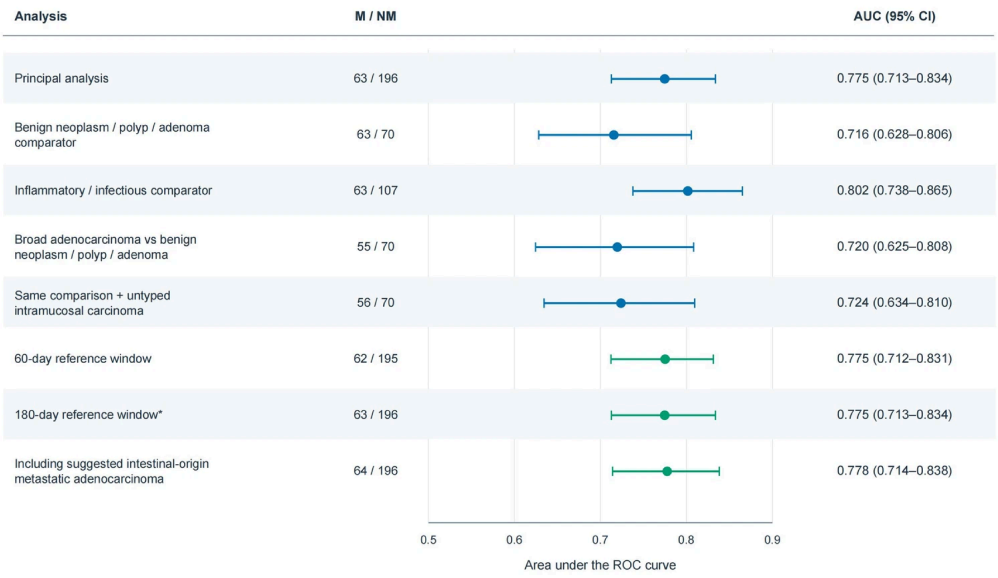

Figure 5. TAP discrimination across comparator and reference-ascertainment sensitivity analyses. Points show AUCs; bars show percentile 95% confidence intervals from 1500 patient-level bootstrap resamples. Malignant and nonmalignant counts are shown separately. Broad adenocarcinoma includes gastric signet-ring cell carcinoma. The 180-day analysis retains the principal cohort; the final comparison adds one metastatic adenocarcinoma with uncertain intestinal origin. All analyses are exploratory. * The 180-day reference-ascertainment analysis retained the same patients as the principal analysis.

Using a 60-day reference-ascertainment window yielded 257 participants, including 62 malignancies, with a TAP AUC of 0.775 (95% CI 0.712–0.831). Extending the window to 180 days did not change the principal analytical cohort. Including the additional metastatic adenocarcinoma with suggested but unconfirmed intestinal origin yielded 260 participants, including 64 malignancies, and an AUC of 0.778 (95% CI 0.714–0.838). These analyses assessed sensitivity to reference ascertainment and source eligibility; they did not establish uniformly prediagnostic or pretreatment TAP sampling.

CRP was available for 123 participants in the principal analysis, including 30 malignancies. Within this same complete-case subset, the TAP–age–sex model had an AUC of 0.724; adding log(1 + CRP) yielded an apparent AUC of 0.728. The TAP association was OR 1.65 (95% CI 1.04–2.63; $p = 0.034$) before CRP adjustment and OR 1.70 (95% CI 1.06–2.73; $p = 0.028$) after adjustment (Table 3). These findings do not establish that TAP is independent of inflammatory effects, given selective CRP availability, the limited number of events, and possible residual inflammatory confounding.

4. Discussion

4.1. Principal Findings

In this pathology-referenced gastrointestinal-surgery cohort, TAP showed modest discrimination between malignant and nonmalignant disease. Three observations are central. First, TAP values and the proportion with elevated TAP were higher in histologically confirmed malignancy than across a clinically relevant nonmalignant disease spectrum. Second, TAP remained associated with malignancy after adjustment for age and sex, but the uncertainty around its incremental discrimination and the absence of independent validation warrant caution. Third, discrimination was lower when the comparison was restricted to benign neoplasms, polyps, or adenomas, and specificity at the institutional 121 μm^2 threshold remained limited. Accordingly, these findings support an association between TAP and gastrointestinal pathological status but do not establish its value as a rule-in, rule-out, screening, or replacement test.

4.2. Relationship to Previous Tap and Glycoprotein Studies

The findings are biologically compatible with aberrant glycosylation in gastrointestinal cancer [14]. Prior studies linked truncated O-glycans and altered Cosmc/T-synthase activity to colorectal tumor biology [3], identified whole-serum lectin signatures [5], and reported discriminatory O-glycoforms [4,15]. However, those technologies, populations, and molecular targets differ from a routine agglomeration-based TAP readout, so their performance estimates are not directly interchangeable.

Gu et al. reported an AUC of 0.935 for gastric cancer at a study-derived threshold of 110.128 μm^2 [6]. Our lower AUC may partly reflect differences in case mix, including the use of symptomatic inflammatory and benign-neoplastic controls [16]. Only 12 malignancies were gastric; therefore, the present data cannot provide a reliable gastric-cancer-specific estimate.

A 2025 study combined TAP with fecal syndecan-2 methylation in colorectal cancer, adenoma, and healthy controls [7]. In our more clinically challenging malignant-versus-benign-neoplasm comparison, nearly half of benign controls had elevated TAP. This overlap limits stand-alone specificity. In colorectal lesions, the biological continuity of the adenoma–carcinoma sequence, involving progressively altered molecular regulatory networks, may partly contribute to biomarker overlap between benign neoplastic and malignant lesions; however, this interpretation remains speculative for TAP [17]. Contemporary cell-free-DNA screening tests were evaluated in different populations and cannot be directly compared with this retrospective surgical disease-spectrum study [2].

4.3. Clinical Interpretation of the Fixed Institutional Threshold

At 121 μm^2, the observed sensitivity of 82.5% was accompanied by a specificity of 64.3% against all nonmalignant pathology and 52.9% against benign neoplasms, polyps, or adenomas. Thus, 11 of 63 malignancies fell below the threshold, while 70 of 196 nonmalignant cases met or exceeded it. A value below 121 μm^2 cannot reliably exclude gastrointestinal malignancy, and an elevated result cannot establish malignancy without clinical assessment and definitive investigation. The NPV of 92.0% arose within a selected analytical cohort and depends on disease prevalence; it should not be generalized to asymptomatic populations or interpreted as a screening guarantee.

The adjusted odds ratio indicates an association between TAP and malignancy conditional on age and sex but does not establish clinical utility. Clinical decision-curve benefit was not assessed [18]. Any potential adjunctive role requires prospective evaluation; the present findings do not support using TAP alone to trigger invasive procedures or as a substitute for endoscopy and histopathology.

4.4. Internal Validation and Selection Bias

The distinction between apparent, optimism-corrected, and chronological performance is methodologically important [19]. The apparent adjusted-model AUC was 0.790, and bootstrap correction reduced it to 0.780. The illustrative later-period AUC was 0.728 with wide confidence intervals. Later-period malignant event counts ranged from 16 to 40. All temporal assessments remained within one institution and one year and therefore do not establish generalizability [20].

Entry into the analytical cohort depended on both reference availability and eligibility. Removing the general-laboratory matching requirement reduced an avoidable restriction but did not eliminate selection arising from clinical investigation and reference ascertainment. Weighting without discharge diagnosis yielded an AUC of 0.775, compared with 0.775 in the unweighted principal analysis. Within the original laboratory-linked subset, corresponding weighted AUCs were 0.728 without discharge diagnosis and 0.737 with it included. These subset estimates should not be directly compared with the full-cohort estimate. Excluding discharge diagnosis from the principal weighting model addressed the concern that this variable could incorporate pathological findings. However, neither approach eliminated potential bias from unmeasured clinical suspicion, imaging, endoscopy, frailty, or treatment decisions, and both relied on assumptions about selection conditional on measured covariates.

4.5. Inflammation and Clinicopathological Heterogeneity

Inflammatory and infectious gastrointestinal disease generated substantial overlap with malignant TAP values, emphasizing that an abnormal glycoprotein-related agglomeration signal should not be considered cancer-specific [21]. In the CRP-complete subset, the TAP association remained positive before and after CRP adjustment, with adjusted odds ratios of 1.65 and 1.70, respectively. The apparent model AUC changed only slightly, from 0.724 to 0.728. These findings do not establish independence from inflammatory effects: CRP was available for only 123 participants, including 30 malignant events, and a single inflammatory marker cannot capture all relevant inflammatory processes. Selective CRP availability and residual inflammatory confounding therefore remain important limitations.

The dataset supports exploratory disease-spectrum comparisons but does not support claims about staging, metastasis prediction, postoperative prognosis, recurrence, or survival.

4.6. Strengths and Limitations

Strengths include patient-level histopathology, consolidation of repeat reports, manually confirmed diagnostic categories, a fixed local threshold, clinically relevant comparators, parsimonious modeling, bootstrap correction, calibration, chronological robustness analyses, and explicit assessment of reference availability and selection into the analytical cohort.

Several limitations are equally important. First, the study was retrospective, single-center, and based on gastrointestinal-surgery patients; referral patterns and case mix limit generalizability. Second, 226 of 485 patients assessed were not included in the principal analysis because an eligible reference could not be established or other eligibility criteria were not met. Removing the general-laboratory matching requirement did not eliminate selection related to clinical investigation and reference eligibility, and inverse-probability weighting cannot recover unknown outcomes or correct unmeasured selection mechanisms. Third, the chronological partition was selected after inspection of event counts; alternative-cutoff analyses showed wide uncertainty and do not constitute external validation. Fourth, the malignant sample contained only 63 cases, including 30 CRP-complete cases, limiting precision and the complexity of credible multivariable models.

Fifth, the reference population was anatomically heterogeneous and dominated by colorectal disease, whereas gastric and small-bowel malignancies were too infrequent for reliable site-specific estimates. Sixth, although pathology receipt and report dates were recovered, these did not consistently establish specimen-collection times, whether TAP preceded biopsy or surgery, or whether malignancy was already known when testing was requested. TAP was generally sampled on admission, but this did not establish an untreated or prediagnostic setting. Some participants had documented prior malignancy or anticancer treatment, and incomplete history ascertainment prevented identification of a uniformly untreated, previously undiagnosed population. The findings were therefore interpreted as associations with pathological status. Although pathology ascertainment was extended into 2025, the institutional exports could not establish complete follow-up outside the hospital. The reference-ascertainment windows were adopted during revision after some follow-up information had been examined and were not prospectively prespecified.

Analytical repeatability and agreement between readers were not evaluated in this retrospective study, limiting assessment of measurement variability relative to the observed between-group differences. Finally, CRP was incompletely observed, pathological staging was structurally unavailable for some biopsy specimens, and no independent cohort, prospective follow-up, head-to-head comparison with established tumor markers, or decision-analytic utility evaluation was available. The study therefore cannot support claims of population-screening performance, diagnostic replacement, superiority to existing markers, early-stage sensitivity, prognostic value, or treatment-response monitoring. Independent prospective cohorts with prespecified verification pathways, systematic inflammatory-marker measurement, complete time stamps, assay-standardization data, and site-specific recruitment are needed before clinical implementation.

5. Conclusions

In a single-center gastrointestinal-surgery cohort, peripheral-blood TAP showed modest discrimination between malignant and nonmalignant pathological outcomes and remained associated with malignancy after adjustment for age and sex. Bootstrap correction, chronological internal assessment, and selection-weighting analyses did not establish external validity or eliminate uncertainty from retrospective selection and incomplete clinical histories. The apparent improvement over age and sex alone did not establish clinical utility. The fixed institutional threshold generated substantial false-positive and false-negative results, particularly when malignant lesions were compared with benign neoplasms. These

findings support further prospective, independent evaluation of TAP but do not establish its performance in a uniformly untreated, previously undiagnosed population or support stand-alone diagnostic, staging, or screening use.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/diagnostics16233967/s1, Table S1, Analytical inclusion by discharge diagnosis category; Table S2, Full specification of the primary logistic regression model; Table S3, Reference timing and ascertainment window analyses; Table S4, Incremental discrimination after adding TAP; Table S5, Chronological internal assessment by partition date; Table S6, Reference availability by original laboratory matching group; Table S7, Selection weighting sensitivity analyses; and Figure S1, Illustrative TAP agglomerate micrographs.

Author Contributions: Conceptualization, Z.L.; methodology, Z.L. and J.H.; data curation, J.H., N.W. and M.W.; resources, N.W. and M.W.; writing—original draft preparation, Z.L.; writing—review and editing, all authors; supervision, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Health Commission of Hunan Province (grant numbers 202201042974 and 20254352), the Changsha Science and Technology Project (grant number kq2004157), the Changsha Municipal Health Commission (grant numbers KJ-202501), the Natural Science Foundation of Hunan Province (grant number 2026JJ82440), and the Science Foundation of the First Hospital of Changsha (grant numbers Y2025-01 and Y2025-02). The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the First Hospital of Changsha (protocol code 2025-40; approval date: 10 March 2025).

Informed Consent Statement: Written informed consent for the use of clinical data for research purposes was obtained from all participants.

Data Availability Statement: The patient-level data are not publicly available because they contain potentially sensitive clinical information. De-identified data may be available from the corresponding author on reasonable request and subject to institutional and ethics approval.

Acknowledgments: The authors gratefully acknowledge all staff members of the Department of Pathology for their valuable contributions to the collection, organization, and verification of pathological records used in this study. During manuscript preparation, OpenAI ChatGPT (GPT-6.1) was used to assist with English-language editing and document formatting.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sung, H.; Filho, A.M.; Laversanne, M.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. *CA A Cancer J. Clin.* **2026**, *76*, e70090. [CrossRef] [PubMed]
2. Chung, D.C.; Gray, D.M.; Singh, H.; Issaka, R.B.; Raymond, V.M.; Eagle, C.; Hu, S.; Chudova, D.I.; Talasaz, A.; Greenson, J.K.; et al. A Cell-free DNA Blood-Based Test for Colorectal Cancer Screening. *N. Engl. J. Med.* **2024**, *390*, 973–983. [CrossRef] [PubMed]
3. Jiang, Y.; Liu, Z.; Xu, F.; Dong, X.; Cheng, Y.; Hu, Y.; Gao, T.; Liu, J.; Yang, L.; Jia, X.; et al. Aberrant O-glycosylation contributes to tumorigenesis in human colorectal cancer. *J. Cell. Mol. Med.* **2018**, *22*, 4875–4885. [CrossRef] [PubMed]
4. Takakura, D.; Ohashi, S.; Kobayashi, N.; Tokuhisa, M.; Ichikawa, Y.; Kawasaki, N. Targeted O-glycoproteomics for the development of diagnostic markers for advanced colorectal cancer. *Front. Oncol.* **2023**, *13*, 1104936. [CrossRef] [PubMed]
5. Bertok, T.; Bertokova, A.; Jane, E.; Hires, M.; Aguedo, J.; Potocarova, M.; Lukac, L.; Vikartovska, A.; Kasak, P.; Borsig, L.; et al. Identification of Whole-Serum Glycobiomarkers for Colorectal Carcinoma Using Reverse-Phase Lectin Microarray. *Front. Oncol.* **2021**, *11*, 735338. [CrossRef] [PubMed]

6. Gu, C.; Xie, L.; Li, B.; Zhang, L.; Li, F.; Wang, W.; Su, J.; Xu, Z. Quantification of Tumor Abnormal Proteins in the Diagnosis and Postoperative Prognostic Evaluation of Gastric Cancer. *Clin. Med. Insights Oncol.* **2022**, *16*, 11795549221104440. [CrossRef] [PubMed]
7. Liu, X.; Zhao, G.; Mao, W.; Li, Q.; Liao, J.; He, G. Clinical significance of fecal Syndecan-2 gene methylation combined with blood tumor abnormal protein detection in the diagnosis of colorectal cancer and precancerous lesions. *Clin. Biochem.* **2025**, *136*, 110887. [CrossRef] [PubMed]
8. Lan, F.; Zhu, M.; Qi, Q.; Zhang, Y.; Liu, Y. Prognostic value of serum tumor abnormal protein in gastric cancer patients. *Mol. Clin. Oncol.* **2016**, *5*, 216–220. [CrossRef] [PubMed]
9. Li, L.-X.; Zhang, B.; Gong, R.-Z. Insights into the role of tumor abnormal protein in early diagnosis of cancer: A prospective cohort study. *Medicine* **2020**, *99*, e19382. [CrossRef] [PubMed]
10. Bossuyt, P.M.; Reitsma, J.B.; Bruns, D.E.; Gatsonis, C.A.; Glasziou, P.P.; Irwig, L.; Lijmer, J.G.; Moher, D.; Rennie, D.; de Vet, H.C.W.; et al. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *Clin. Chem.* **2015**, *61*, 1446–1452. [CrossRef] [PubMed]
11. Day, E.; Eldred-Evans, D.; Prevost, A.T.; Ahmed, H.U.; Fiorentino, F. Adjusting for verification bias in diagnostic accuracy measures when comparing multiple screening tests: An application to the IP1-PROSTAGRAM study. *BMC Med. Res. Methodol.* **2022**, *22*, 70. [CrossRef] [PubMed]
12. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; Riley, R.D.; Beam, A.L.; Van Calster, B.; Ghassemi, M.; Liu, X.; Reitsma, J.B.; van Smeden, M.; et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **2024**, *385*, e078378. Correction in *BMJ* **2024**, *385*, q902. https://doi.org/10.1136/bmj-2023-078378. [CrossRef] [PubMed]
13. Arifin, W.N.; Yusof, U.K. Correcting for partial verification bias in diagnostic accuracy studies: A tutorial using R. *Stat. Med.* **2022**, *41*, 1709–1727. [CrossRef] [PubMed]
14. Pinho, S.S.; Reis, C.A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **2015**, *15*, 540–555. [CrossRef] [PubMed]
15. Grzesik, K.; Janik, M.; Hoja-Łukowicz, D. The hidden potential of glycomarkers: Glycosylation studies in the service of cancer diagnosis and treatment. *Biochim. Biophys. Acta Rev. Cancer* **2023**, *1878*, 188889. [CrossRef] [PubMed]
16. Usher-Smith, J.A.; Sharp, S.J.; Griffin, S.J. The spectrum effect in tests for risk prediction, screening, and diagnosis. *BMJ* **2016**, *353*, i3139. [CrossRef] [PubMed]
17. Li, J.; Yan, X.; Jiao, H.; Chen, J.; Lin, Y.; Zhou, M.; Jin, F.; Xu, Q.; Zhang, Z. Identification of target gene–microribonucleic acid–transcription factor regulatory networks in colorectal adenoma-carcinoma sequence. *Oncol. Transl. Med.* **2025**, *11*, 118–137. [CrossRef]
18. Van Calster, B.; Collins, G.S.; Vickers, A.J.; Wynants, L.; Kerr, K.F.; Barreñada, L.; Varoquaux, G.; Singh, K.; Moons, K.G.; Hernandez-Boussard, T.; et al. Evaluation of performance measures in predictive artificial intelligence models to support medical decisions: Overview and guidance. *Lancet Digit. Health* **2025**, *7*, 100916. [CrossRef] [PubMed]
19. Collins, G.S.; Dhiman, P.; Ma, J.; Schlussel, M.M.; Archer, L.; Van Calster, B.; Harrell, F.E., Jr.; Martin, G.P.; Moons, K.G.M.; van Smeden, M.; et al. Evaluation of clinical prediction models (part 1): From development to external validation. *BMJ* **2024**, *384*, e074819. [CrossRef] [PubMed]
20. Riley, R.D.; Archer, L.; Snell, K.I.E.; Ensor, J.; Dhiman, P.; Martin, G.P.; Bonnett, L.J.; Collins, G.S. Evaluation of clinical prediction models (part 2): How to undertake an external validation study. *BMJ* **2024**, *384*, e074820. [CrossRef] [PubMed]
21. He, M.; Zhou, X.; Wang, X. Glycosylation: Mechanisms, biological functions and clinical implications. *Signal Transduct. Target. Ther.* **2024**, *9*, 194. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.