Am J Cancer Res 2026;16(8):3488-3503
www.ajcr.us /ISSN:2156-6976/ajcr0174052

Original Article

A seven-variable clinical prediction model for bone metastasis at initial diagnosis of prostate cancer

Yongkang Ni[1], Ning Tao[1], Hengqing An[2]

[1]*School of Public Health, Xinjiang Medical University, Urumqi 830017, Xinjiang Uygur Autonomous Region, China;* [2]*Department of Urology, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830013, Xinjiang Uygur Autonomous Region, China*

Received May 14, 2026; Accepted July 17, 2026; Epub August 15, 2026; Published August 30, 2026

Abstract: Decisions about bone-imaging workup at initial diagnosis of prostate cancer often have to be made before PSMA-PET/CT or other specialized investigations are available, particularly in resource-constrained settings. Many recent prediction models rely on advanced imaging, radiomics, specialized biomarkers, or large registry infrastructures and may therefore be difficult to apply in this earlier decision context. We retrospectively reviewed records of 291 consecutive men with newly diagnosed, histopathologically confirmed prostate adenocarcinoma admitted to the First Affiliated Hospital of Xinjiang Medical University between March 2011 and November 2023. Starting from 93 candidate predictors, we applied a five-stage hybrid selection procedure (univariable screening, exploratory LASSO, clinical prescreening, data-quality review, and confirmatory LASSO) to derive the final variable set. The retained predictors were entered into a multivariable logistic regression and visualized as a nomogram. Model behavior was characterized through discrimination, calibration, decision curve analysis, and 1,000-iteration bootstrap optimism correction. Of the 291 patients, 113 (38.8%) had bone metastasis. The final 7-variable model included anemia status, alanine aminotransferase, lactate dehydrogenase, alkaline phosphatase, age, Gleason score (three groups), and total prostate-specific antigen (tPSA). The apparent AUC was 0.736 (95% CI 0.674-0.799); the bootstrap bias-corrected AUC was 0.713 with a calibration slope of 0.933. This seven-variable model uses routinely available variables and showed moderate discrimination with stable calibration; external validation in independent samples is required before clinical use.

Keywords: Prostate cancer, bone metastasis, prediction model, nomogram, logistic regression, risk stratification

Introduction

Prostate cancer is prevalent, and the bone is the most common location of metastasis. Although PSMA-based PET/CT improves detection, it is not routinely available at the time of diagnosis [1, 2]. Data from various studies indicate that approximately ten percent of patients already exhibit distant metastases at diagnosis, especially in the context of limited opportunities for early diagnosis [3]. Imaging at baseline is important because bone metastasis at diagnosis alters staging and treatment [4]. Clinicians tend to depend on single cutoffs or their overall judgment in practice.

Recent prediction models have used radiomics from multiparametric MRI [5, 6], serum sialic acid [7], and population-based risk tools. Though these approaches show strong discrimination, studies often require advanced imaging or specialized assays and are not easily applicable at first diagnosis, especially in resource-poor settings. Routine-variable models remain quite limited, and the data from western China is sparse.

We aimed to develop and internally validate a clinical prediction model for bone metastasis at initial diagnosis of prostate cancer using routine clinical and laboratory variables only. Our objective was to construct a parsimonious and interpretable first-line risk-stratifying tool for situations where advanced imaging or specialized biomarkers are not readily available. With a consecutive single-center sample in Xinjiang, China, we applied a structured five-stage variable selection process with explicit data quality

https://doi.org/10.62347/CIEN9343

review [8]; the completed TRIPOD+AI checklist is available in Supplementary 4 [9].

Materials and methods

Study design and participants

We conducted a single-center, retrospective observational study using electronic medical records from the Department of Urology, First Affiliated Hospital of Xinjiang Medical University. All consecutive male patients with histopathologically confirmed prostate cancer admitted between March 2011 and November 2023 were screened for inclusion.

Eligible patients were men aged ≥18 years with histopathologically confirmed prostate adenocarcinoma, newly diagnosed during the index hospitalization. All had baseline clinical, laboratory, and imaging data obtained before any cancer-directed therapy, and had undergone at least one imaging modality capable of assessing bone metastasis during the index hospitalization (mp-MRI, CT, or whole-body bone scintigraphy [ECT]).

Patients were excluded if they had a non-prostate-cancer or non-malignant pathological diagnosis, had received prior cancer-directed therapy before baseline assessment, had a second primary malignancy that could confound attribution of bone metastasis, had missing primary outcome data, or represented duplicate hospitalizations, in which case only the earliest diagnostic admission was retained. A total of 291 patients were included.

This study has been reviewed and approved by the Ethics Committee of Xinjiang Medical University (approval No. XIYKDXR202305-10011).

Data sources and variable definitions

The sources of data included four institutional systems, the hospital information system (HIS), the laboratory information system (LIS), the pathology system and the picture archiving and communication system (PACS). Two trained abstractors independently inserted records into a predefined form, with disagreements adjudicated against the primary documentation.

Outcome variable: The bone metastases at initial diagnosis, as recorded using mp-MRI, CT, and whole-body bone scintigraphy (ECT) during the index hospitalization (up to three modalities), were the primary outcome. Patients were classified as having bone metastasis if any imaging modality indicated the presence of bone metastasis. Patients were considered negative for bone metastasis if all evaluated modalities showed no evidence of bone involvement, with at least one modality explicitly confirming the absence of metastasis.

Outcome-wise, we treated patients with non-interpretable results from any of the three modalities as missing. We applied this any-positive rule to minimize false-negative classification as ECT, CT, and mp-MRI capture osteoblastic lesions, cortical bone destruction, and marrow infiltration, respectively. Simultaneously, the aforementioned composite any-positive endpoint may exhibit a reduced specificity; consequently, certain benign, degenerative, traumatic or inflammatory bone lesions may have been classified as metastases when standard clinical reports were not verified to a uniform reference standard.

Candidate predictor variables: In the past decade, new findings about prostate cancer bone metastasis have emerged and are listed as candidate predictors. The domains included demographics; comorbidities and medical history; PSA and derivatives; pathology; imaging-based local invasion; complete blood count; and biochemistry/coagulation variables.

Demographics: age at admission, body mass index, educational level, marital status, place of residence, and health insurance type.

Comorbidities and medical history: hypertension, diabetes mellitus, dyslipidemia, coronary heart disease, stroke or transient ischemic attack, and family history of prostate cancer.

PSA and derivatives: total PSA (tPSA), free PSA (fPSA), and the free-to-total PSA ratio (f/tPSA).

Pathology: number of biopsy cores, number of positive cores, Gleason sum score, and the highest Gleason grade.

Local invasion on imaging: extracapsular extension (ECE), seminal vesicle invasion (SVI), blad-

der or rectal invasion, and suspected pelvic lymph node involvement.

Complete blood count: white blood cell count, red blood cell count, hemoglobin, platelet count, and differential percentages.

Biochemistry and coagulation: alkaline phosphatase (ALP), lactate dehydrogenase (LDH), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin, direct bilirubin, indirect bilirubin, delta bilirubin, serum calcium, magnesium, plasma fibrinogen, prothrombin time activity, and the international normalized ratio.

For each laboratory variable, the measurement closest to pathological confirmation or bone-metastasis assessment was used; when multiple values were available during the index hospitalization, the first measurement after admission was selected. The raw database initially contained 291 patients and 119 candidate variables.

Data preprocessing and quality control

Raw data were cleaned to address duplicate HIS/LIS entries, zero-coded values that represented unmeasured laboratory results, occasional integer underflow, inconsistent categorical coding across years, and near-zero-variance variables. Where this could be reliably adjudicated against the original laboratory reports on a case-by-case basis, the affected zeros were converted to missing values. For LDH, however, a substantial number of zero-coded records (57/291, 19.6%) could not be reliably adjudicated from the source database; LDH was therefore retained with its original values, and the potential impact of these uncertain zeros was addressed in the Limitations section of the main manuscript. Variables with >50% missingness were removed, and remaining missing values were handled using multiple imputation by chained equations (MICE; m = 5). All subsequent analyses were performed on each of the five imputed datasets and pooled using Rubin's rules. Continuous variables with established clinical cut-offs were dichotomized where appropriate, and redundant or derived variables were excluded to reduce collinearity. Two protein variables with uncertain field identity and one volume variable with 57.7% missingness were also excluded. After preprocessing, 93 candidate predictors remained. Full details are provided in the Supplementary 1.

Statistical analysis

All analyses were performed in R version 4.4.2 (R Foundation for Statistical Computing, Vienna, Austria). A two-sided statistical test was performed with a $P<0.05$ threshold for significance.

Descriptive statistics: Continuous variables are summarized as median (Q1, Q3) and compared between the bone-metastasis and non-metastatic groups with the Mann-Whitney U test. Categorical variables are reported as counts (percentages) and compared with Pearson's chi-squared test, or Fisher's exact test when an expected cell count fell below five.

Variable selection: Due to the limited occurrence of 113 bone metastasis events for the 93 candidate and predictor variable, a hybrid variable selection method was utilized that includes statistical screening with clinical prior knowledge.

Stage 1: Univariable screening. Each of the 93 candidate variables was fitted in a univariable logistic regression model. Variables with $P<0.10$ were retained, yielding 25 candidates.

Stage 2: Exploratory unrestricted LASSO (sensitivity analysis). We first applied LASSO logistic regression with 10-fold cross-validation and lambda.1se to all 25 variables significant in univariable analysis. This left 21 variables in the model, but several categorical levels showed complete separation, and the events-per-variable ratio was about 3.0, below conventional stability recommendations [10-15]. It was therefore not suitable as the final model, although the guideline-endorsed core predictors, T stage, Gleason sum score, and tPSA, were retained.

Stage 3: Preclinical assessment. According to the NCCN and EAU guidelines and the biological literature, we prescreened 10 clinically meaningful variables, including those of tumor burden (ALP, LDH), grade and stage (T stage, Gleason sum score, tPSA), hematological status (red blood cell count, anemia), hepatic function (ALT), comorbidity (history of cardiac disease), and demographics (age).

Stage 4: Data quality review. We then reviewed the 10 clinically prescreened variables for database-specific coding artifacts before confirmatory modeling.

History of cardiac disease. The default code that most frequently occurred accounted for a total of 251/291 records and had a bone-metastasis rate of 109/251 = 43.4%. Meanwhile, the explicitly recorded 'none' entry variant A accounted for 37/291 records with a rate of 4/37 = 10.8%. Finally, the single 'none' entry variant B record was also metastasis-negative. This pattern does not make clinical sense for semantically identical negative entries and appears to strongly suggest default-value contamination. This variable was excluded from the analysis because negative true outcomes could not be reliably separated from records that were not filled.

Clinical T stage. Bone metastasis was noted at a frequency of 27.3% in T1 (6/22), 46.6% in T2 (83/178), 30.4% in T3 (14/46), and 22.2% in T4 (10/45). Clinically, metastatic risk would be expected to rise with advancing local stage, but the T2 group was both unreasonably large and had the most elevated event rate. When stage was not explicitly recorded, this pattern most closely resembled default T2 coding, so the variable was excluded rather than post hoc recoded.

In addition, Gleason sum score was collapsed into three ordered groups (≤7, 8, and ≥9) with ≤7 as the reference to improve stability and reflect commonly used clinical risk categories.

Stage 5: Confirmation of LASSO Selection. We repeated LASSO logistic regression with 10-fold cross-validation and lambda.1se on the 8 clinically prescreened variables. Red blood cell count was retained at lambda.min but not at lambda.1se. The final predictor set included anemia status, ALT, LDH, ALP, age at admission, Gleason sum score (3 groups), and tPSA alone.

Multivariable model and non-linearity assessment: The final multivariable logistic regression model was fitted using the lrm function of the rms package (version 6.8) in R. Linearity on the log-odds scale was assessed for age, ALP, LDH, ALT, and tPSA using restricted cubic splines with three knots; all tests yielded P>0.10, so all continuous predictors were entered linearly. Adjusted odds ratios (aORs) and 95% confidence intervals are reported. Multicollinearity was assessed using *GVIF^(1/(2·df))* for categorical terms and *VIF* for continuous terms; all values were <2.

The event-to-predictor ratio was 113/7, or 16.1, and the event-to-non-intercept parameter ratio was 113/8, or 14.1, both exceeding conventional recommendations for model stability [8, 15].

Model performance: The area under the receiver operating characteristic curve (AUC) was computed for discrimination with DeLong based 95% confidence intervals. Calibration is characterized by a smoothed calibration curve and by the corresponding slope, intercept, and maximum absolute deviation (Emax). The Brier Score summarized the overall accuracy. The net clinical benefit was determined through decision curve analysis (DCA) based on both the treat-all and the treat-none defaults.

Nomogram construction: Ultimately, a nomogram was developed from the final multivariable logistic regression, allowing point values for each predictor to be summed and converted into an estimated probability of bone metastasis at the point of initial diagnosis.

Results

Patient characteristics

From March 2011 to November 2023, 3,362 hospitalization records from the Urology Department were evaluated. After exclusions, 291 patients remained in the final analysis, including 113 (38.8%) with bone metastasis at diagnosis and 178 (61.2%) without bone metastasis (**Figure 1**).

As shown in **Table 1**, the average age of patients with bone metastasis was about three years older than those without bone metastasis; Bone metastasis in prostate cancer patients was frequently associated with higher Gleason scores (8 or ≥9); Bone metastasis also occurred more frequently with extracapsular extension, seminal vesicle invasion, and suspected lymph node involvement, but less frequently with hypertension.

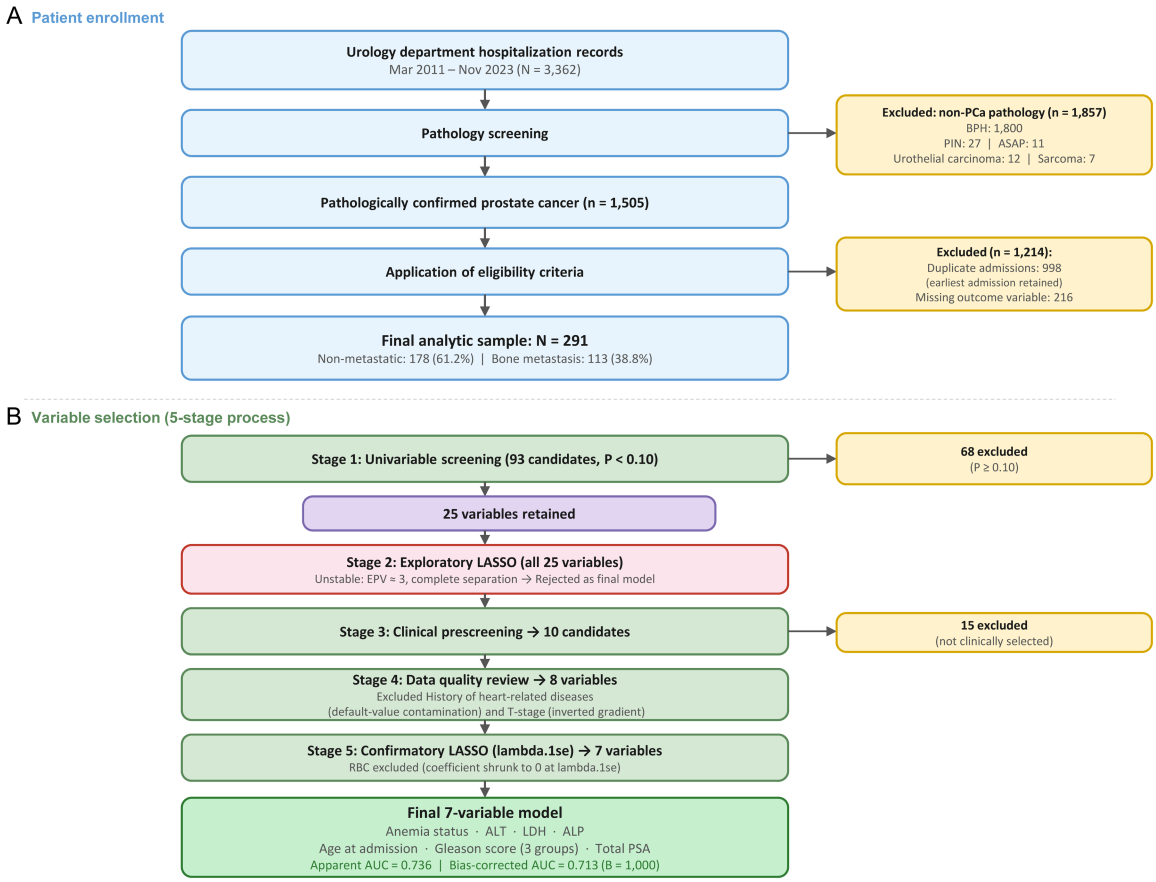


Figure 1. Study flowchart and five-stage variable selection process. Of 3,362 urology department hospitalization records screened, 291 patients were included in the final analytic sample (178 non-metastatic and 113 with bone metastasis). The candidate predictor set was reduced from 93 variables to a final 7-variable model through univariable screening, exploratory LASSO, clinical prescreening, data-quality review, and confirmatory LASSO. BPH, benign prostatic hyperplasia; PIN, prostatic intraepithelial neoplasia; ASAP, atypical small acinar proliferation.

Laboratory results showed a higher incidence of anemia, lower red blood cell count, and higher free PSA levels in bone-metastasis group. Alkaline phosphatase was elevated and alanine aminotransferase was decreased in the bone-metastasis group, while lactate dehydrogenase showed borderline elevation. The bivariable comparison did not suggest a significant difference for total PSA, which was numerically higher but not significant. Serum calcium was lower in the bone-metastasis group, whereas the white blood cell count, platelets, creatinine, and electrolytes showed no significant differences between groups. The details are shown in **Table 2**.

Univariable screening

Out of 93 candidate predictors, 25 predictors met the prespecified univariable threshold of P<0.10 and were retained for further assessment (**Table 3**; full results Supplementary Table 1). The strongest associations were seen with anemia status, extracapsular extension, and seminal vesicle invasion. Other notable predictors were alkaline phosphatase, age, red blood cell count, tPSA, lactate dehydrogenase and alanine aminotransferase.

Exploratory LASSO (Stage 2)

When the 25 univariable-significant variables were input into unrestricted LASSO logistic regression with 10-fold cross-validation, the lambda.1se model retained 21 variables but exhibited complete separation at several categorical levels and an events-per-variable ratio of about 3.0, below usual stability recommendations. Given the exploratory nature of the model, it was deemed unstable and was not

Table 1. Clinical characteristics of patients with and without bone metastasis at initial diagnosis

Variable	Overall (*N* = 291)	Non-metastatic (*N* = 178)	Bone metastasis (*N* = 113)	*P* value
Age at admission, years	74.0 (68.0, 78.0)	73.0 (65.2, 77.0)	76.0 (71.0, 79.0)	<0.001
Marital status†				0.075
Married	270 (92.8)	169 (94.9)	101 (89.4)	
Widowed	20 (6.9)	8 (4.5)	12 (10.6)	
Divorced	1 (0.3)	1 (0.6)	0 (0.0)	
Educational level				0.095
Illiterate	48 (16.5)	33 (18.5)	15 (13.3)	
Primary school	47 (16.2)	28 (15.7)	19 (16.8)	
Junior high school	32 (11.0)	25 (14.0)	7 (6.2)	
Senior high school	57 (19.6)	29 (16.3)	28 (24.8)	
Technical secondary school	36 (12.4)	18 (10.1)	18 (15.9)	
Junior college	35 (12.0)	20 (11.2)	15 (13.3)	
Bachelor's degree or above	36 (12.4)	25 (14.0)	11 (9.7)	
Residence†				0.758
Urban	244 (83.8)	149 (83.7)	95 (84.1)	
Rural	45 (15.5)	27 (15.2)	18 (15.9)	
Urban/town	2 (0.7)	2 (1.1)	0 (0.0)	
Health insurance type†				0.003
Urban employee basic medical insurance	225 (77.3)	137 (77.0)	88 (77.9)	
Urban resident basic medical insurance	18 (6.2)	12 (6.7)	6 (5.3)	
Other	35 (12.0)	27 (15.2)	8 (7.1)	
Other social insurance	12 (4.1)	2 (1.1)	10 (8.8)	
Urban employee plus urban resident basic medical insurance (dual enrollment)	1 (0.3)	0 (0.0)	1 (0.9)	
Hypertension				0.040
No	193 (66.3)	110 (61.8)	83 (73.5)	
Yes	98 (33.7)	68 (38.2)	30 (26.5)	
Diabetes mellitus				0.247
No	205 (70.4)	121 (68.0)	84 (74.3)	
Yes	86 (29.6)	57 (32.0)	29 (25.7)	
Dyslipidemia				0.757
No	216 (74.2)	131 (73.6)	85 (75.2)	
Yes	75 (25.8)	47 (26.4)	28 (24.8)	

Number of biopsy cores	13.0 (12.0, 13.0)	13.0 (12.0, 13.0)	13.0 (12.0, 13.0)	0.219
Number of positive cores	6.0 (3.0, 10.0)	6.0 (2.0, 10.0)	7.0 (4.0, 11.0)	0.018
Percentage of positive cores, %	53.8 (23.1, 84.6)	46.2 (16.7, 76.9)	66.7 (30.8, 92.3)	0.005
Gleason score (3 groups)				0.006
≤7 (reference)	134 (46.0)	95 (53.4)	39 (34.5)	
8	103 (35.4)	56 (31.5)	47 (41.6)	
≥9	54 (18.6)	27 (15.2)	27 (23.9)	
Extracapsular extension (ECE)				<0.001
No	197 (67.7)	137 (77.0)	60 (53.1)	
Yes	94 (32.3)	41 (23.0)	53 (46.9)	
Seminal vesicle invasion (SVI)†				<0.001
No	189 (64.9)	131 (73.6)	58 (51.3)	
Yes	93 (32.0)	43 (24.2)	50 (44.2)	
No (none)	3 (1.0)	2 (1.1)	1 (0.9)	
Yes (bilateral)	6 (2.1)	2 (1.1)	4 (3.5)	
Suspected lymph node involvement				0.012
No	212 (72.9)	139 (78.1)	73 (64.6)	
Yes	79 (27.1)	39 (21.9)	40 (35.4)	

Data are median (Q1, Q3) for continuous variables and n (%) for categorical variables. Continuous variables were compared using the Mann-Whitney U test. Categorical variables were compared using Pearson's chi-squared test or Fisher's exact test, as appropriate. †Fisher's exact test. ECE, extracapsular extension; SVI, seminal vesicle invasion; Q1/Q3, first/third quartile.

Table 2. Laboratory characteristics of patients with and without bone metastasis at initial diagnosis

Variable	Overall (*N* = 291)	Non-metastatic (*N* = 178)	Bone metastasis (*N* = 113)	*P* value
Total PSA (tPSA), ng/mL	19.6 (10.4, 59.0)	18.0 (10.9, 42.9)	27.7 (8.7, 88.0)	0.144
Free PSA (fPSA), ng/mL	2.4 (1.0, 7.0)	1.9 (0.9, 4.1)	4.4 (1.6, 23.8)	<0.001
Free-to-total PSA ratio (f/tPSA)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.300
Anemia status				<0.001
Normal (Hb≥120 g/L)	239 (82.1)	160 (89.9)	79 (69.9)	
Anemia (Hb<120 g/L)	52 (17.9)	18 (10.1)	34 (30.1)	
Red blood cell count (RBC), $\times10^{12}$/L	4.4 (4.0, 4.8)	4.5 (4.1, 4.9)	4.2 (3.8, 4.7)	0.001
White blood cell count (WBC), $\times10^{9}$/L	6.1 (5.2, 7.3)	6.1 (5.1, 7.3)	6.1 (5.3, 7.3)	0.659
Neutrophil percentage, %	64.2 (57.7, 69.9)	64.0 (57.4, 69.6)	64.5 (58.3, 70.4)	0.497
Lymphocyte percentage, %	25.3 (20.2, 30.7)	25.4 (20.7, 30.8)	25.3 (20.1, 30.2)	0.583
Platelet count, $\times10^{9}$/L	205.0 (171.5, 253.0)	205.0 (173.2, 250.5)	203.0 (167.0, 257.0)	0.794
Alanine aminotransferase (ALT), U/L	20.0 (14.0, 27.3)	21.5 (15.8, 28.0)	17.9 (12.3, 25.9)	0.013
Aspartate aminotransferase (AST), U/L	23.5 (17.6, 29.9)	23.8 (17.9, 31.5)	23.2 (16.9, 28.1)	0.090
Alkaline phosphatase (ALP), U/L	71.0 (58.0, 93.2)	70.0 (57.0, 84.4)	77.6 (59.7, 120.3)	0.005
Lactate dehydrogenase (LDH), U/L	154.8 (0.0, 192.3)	153.6 (0.0, 182.3)	167.0 (0.0, 201.0)	0.059
Creatinine, μmol/L	76.0 (63.8, 90.0)	77.3 (64.0, 89.8)	76.0 (62.4, 93.3)	0.844
Urea, mmol/L	6.1 (5.0, 7.6)	6.1 (5.1, 7.9)	6.1 (4.8, 7.3)	0.656
Calcium, mmol/L	2.2 (2.1, 2.3)	2.2 (2.1, 2.3)	2.2 (2.1, 2.2)	<0.001
Sodium, mmol/L	139.7 (137.5, 142.3)	140.0 (137.6, 142.5)	139.2 (137.5, 142.1)	0.225
Potassium, mmol/L	3.8 (3.5, 4.1)	3.7 (3.5, 4.1)	3.9 (3.5, 4.1)	0.287
Plasma fibrinogen, g/L	3.3 (2.8, 3.9)	3.2 (2.8, 3.9)	3.5 (2.8, 4.4)	0.016
Prothrombin time activity, %	104.4 (88.5, 114.2)	106.9 (90.0, 117.3)	100.0 (85.3, 109.8)	0.012
International normalized ratio (INR)	1.0 (1.0, 1.1)	1.0 (0.9, 1.0)	1.0 (1.0, 1.1)	0.007

pursued as the final prediction model. The clinical T stage, Gleason score, and tPSA core predictors were all retained.

Clinical prescreening and data quality review

Guided by clinical guidelines and the literature, 10 of the 25 univariable-significant variables were prescreened as clinically meaningful candidates representing tumor burden, stage/grade, hematological status, hepatic function, comorbidity, and demographics. Data quality review then excluded two variables. History of cardiac disease looked like default-value contamination: the dominant default code covered 251 of 291 records and had a bone-metastasis rate of 43.4%, whereas explicit 'none' entry variant A covered 37 records with a rate of 10.8%; the single 'none' entry variant B record was also metastasis-negative. Clinical T stage was also excluded because the rates were 27.3% for T1, 46.6% for T2, 30.4% for T3, and 22.2% for T4, with T2 both unusually large and directionally implausible. In both cases, we could not separate true clinical information from placeholder coding reliably.

Confirmatory LASSO (Stage 5)

Subsequently, logistic regression via LASSO with a 10-fold cross-validation was performed on the remaining eight clinically prescreened variables (Supplementary 2; Supplementary Table 3; Supplementary Figures 1 and 2). At lambda.1se, a subset of seven variables was retained. At this stage, we excluded only the red blood cell count variable. The coefficient for a Gleason score of ≥9 versus ≤7 was dropped from the model at lambda.1se, but since the 8 versus ≤7 contrast was non-zero, the variable was retained. The predictor set finally included anemia status, alanine aminotransferase, lactate dehydrogenase, alkaline phosphatase, age at admission, Gleason score (three groups), tPSA.

Final multivariable model

The final multivariable logistic regression model included seven predictors and eight non-intercept parameters (**Table 4**). Anemia had the largest effect (adjusted OR 2.50; 95% CI 1.22 to 5.12; P = 0.012), followed by Gleason score

Table 3. Univariable logistic regression analysis of candidate predictors for bone metastasis at initial diagnosis

Variable	Crude *OR*	95% *CI*	*P* value	Final disposition
Extracapsular extension (ECE): Yes vs. No	2.95	1.78-4.91	<0.001	Not selected during clinical prescreening
Anemia status: Anemia (Hb<120 g/L) vs. Normal (Hb≥120 g/L)	3.83	2.03-7.20	<0.001	Retained in final model
Seminal vesicle invasion (SVI): Yes vs. No	2.63	1.57-4.38	<0.001	Not selected during clinical prescreening
Alkaline phosphatase (ALP), U/L	1.01	1.00-1.02	<0.001	Retained in final model
History of cardiac disease: explicit 'none' entry variant A vs. dominant default code	0.16	0.05-0.46	<0.001	Excluded in data quality review
Age at admission, years	1.06	1.02-1.09	<0.001	Retained in final model
Red blood cell count (RBC), $\times 10^{12}$/L	0.57	0.40-0.82	0.002	Excluded by LASSO penalization
Clinical T stage: T4 vs. T2	0.33	0.15-0.70	0.004	Excluded in data quality review
Percentage of positive cores, %	1.01	1.00-1.02	0.005	Not selected during clinical prescreening
Gleason score (3 groups): ≥9 vs. ≤7	2.44	1.27-4.67	0.007	Retained in final model
Health insurance type: Other social insurance vs. Urban employee basic medical insurance	7.78	1.67-36.37	0.009	Not selected during clinical prescreening
Gleason score (3 groups): 8 vs. ≤7	2.04	1.19-3.50	0.009	Retained in final model
Number of biopsy cores	0.77	0.63-0.94	0.009	Not selected during clinical prescreening
Suspected lymph node involvement: Yes vs. No	1.95	1.16-3.30	0.012	Not selected during clinical prescreening
Plasma fibrinogen, g/L	1.43	1.08-1.91	0.013	Not selected during clinical prescreening
Calcium, mmol/L	0.45	0.24-0.85	0.014	Not selected during clinical prescreening
Free PSA (fPSA), ng/mL	1.01	1.00-1.01	0.015	Not selected during clinical prescreening
Number of positive cores	1.07	1.01-1.13	0.019	Not selected during clinical prescreening
Total PSA (tPSA), ng/mL	1.00	1.00-1.00	0.021	Retained in final model
Prothrombin time activity, %	0.99	0.97-1.00	0.022	Not selected during clinical prescreening
Lactate dehydrogenase (LDH), U/L	1.00	1.00-1.00	0.035	Retained in final model
Alanine aminotransferase (ALT), U/L	0.98	0.96-1.00	0.040	Retained in final model
Hypertension: Yes vs. No	0.58	0.35-0.98	0.041	Not selected during clinical prescreening
Free-to-total PSA ratio	0.19	0.04-1.00	0.051	Not selected during clinical prescreening
International normalized ratio (INR)	6.22	1.00-38.82	0.051	Not selected during clinical prescreening
Marital status: Widowed vs. Married	2.51	0.99-6.35	0.052	Not selected during clinical prescreening
Educational level: Senior high school vs. Illiterate	2.12	0.95-4.73	0.065	Not selected during clinical prescreening
Magnesium, mmol/L	0.29	0.07-1.22	0.091	Not selected during clinical prescreening
Delta bilirubin, μmol/L	0.94	0.87-1.01	0.098	Not selected during clinical prescreening

Variables with P<0.10 in univariable logistic regression, sorted by *P* value. For Gleason score, both category comparisons are shown. Disposition: 'Retained in final model': included in the 7-variable multivariable model; 'Excluded in data quality review': suspected default-value coding artifacts (history of cardiac disease, clinical T stage); 'Excluded by LASSO penalization': retained at lambda.min but not at lambda.1se (red blood cell count); 'Not selected during clinical prescreening': significant but outside the clinical pre-screen set. Complete level-specific distributions and bone-metastasis rates for clinical T stage and history of cardiac disease are provided in Supplementary Table 10. The duplicate source field denotes a duplicate source column for the same free-to-total PSA ratio measurement. OR, odds ratio; CI, confidence interval.

Table 4. Final multivariable logistic regression model for bone metastasis at initial diagnosis

Variable	*Beta*	*SE*	*OR*	95% *CI*	*P* value	GVIF^(1/(2**df*))
Anemia status: Anemia (Hb<120 g/L) vs. Normal (Hb≥120 g/L)	0.918	0.365	2.50	1.22-5.12	0.012	1.02
Alanine aminotransferase (ALT), U/L	-0.022	0.010	0.98	0.96-1.00	0.024	1.04
Lactate dehydrogenase (LDH), U/L	0.003	0.001	1.00	1.00-1.01	0.069	1.01
Alkaline phosphatase (ALP), U/L	0.010	0.003	1.01	1.00-1.02	0.002	1.03
Age at admission, years	0.051	0.018	1.05	1.02-1.09	0.005	1.01
Gleason score (3 groups): 8 vs. ≤7	0.779	0.309	2.18	1.19-4.00	0.012	1.01
Gleason score (3 groups): ≥9 vs. ≤7	0.510	0.380	1.66	0.79-3.51	0.180	1.01
Total PSA (tPSA), ng/mL	0.001	0.000	1.00	1.00-1.00	0.140	1.01

Intercept = -5.594. OR, odds ratio; CI, confidence interval; SE, standard error. For categorical predictors, GVIF^(1/(2*df)) (generalized VIF) is shown. Reference categories: Anemia status: Normal (Hb≥120 g/L); Gleason score: Gleason ≤7. Model: binary logistic regression, N = 291 (events = 113). Coefficients shown are from the primary analysis fitted on the first MICE-imputed dataset; a sensitivity analysis using Rubin's rules pooled estimates across all five imputations is presented in Supplementary 3 (Supplementary Tables 6 and 7). For continuous predictors with small unit increments, per-unit ORs may round to 1.00; the regression coefficients should be used for prediction.

8 vs. ≤7 (aOR 2.18; 95% CI 1.19 to 4.00; P = 0.012). Older age and alkaline phosphatase were associated with higher risk, whereas alanine aminotransferase was associated with lower risk. LDH and tPSA stayed in the model because of LASSO selection and clinical relevance, even though they were not individually significant. All non-linearity tests were non-significant, and all *GVIF* values were below 2.

Model performance and internal validation

On the development sample, the observed AUC was 0.736 (95% CI 0.674 to 0.799), and the Brier score was 0.186 (**Figure 2A**; Supplementary Table 4). By applying the Youden cutoff of 0.436, it obtained a sensitivity of 58.4%, specificity of 81.5%, positive predictive value of 66.7% and negative predictive value of 75.5%.

Following 1,000 bootstrap resamples, the AUC was adjusted to 0.713 (Supplementary Table 5). The corrected calibration slope and intercept were 0.933 and -0.025, respectively. Emax was 0.050, and Brier score was 0.197. This shows limited overfitting and fairly constant calibration (**Figure 2B**).

Nomogram and clinical utility

A nomogram was created from the final multivariable logistic regression model to estimate individual risk (**Figure 3A**). The points allocated to the seven predictors for each patient are summed together and converted into the predicted probability of having bone metastasis at diagnosis. Alkaline phosphatase and tPSA displayed the broadest range of points, while anemia status and Gleason score resulted in fixed point increments.

The decision curve analysis shows greater net benefit of the model compared to treat-all and treat-none strategies across threshold probabilities ranging from approximately 0.1 to 0.7 (**Figure 3B**). This range is relevant to the de cision whether to pursue further imaging work-up for bone metastasis for a newly diagnosed patient.

Discussion

Bone metastasis at diagnosis continues to affect treatment decisions, and recent reviews and guidelines highlight this burden [1, 3, 4]. In this research, a routine-variable model was developed and internally validated in 291 consecutive newly diagnosed patients from the First Affiliated Hospital of Xinjiang Medical University. We identified seven predictors (anemia status, ALT, LDH, ALP, age, Gleason score, tPSA) through univariable screening, exploratory LASSO, clinical prescreening, data-quality review, and confirmatory LASSO. The model had moderate discrimination (apparent AUC 0.736; bias-corrected AUC 0.713) with stable calibration after bootstrap internal validation.

Individual predictor findings in context

Each predictor retained in the final model has a plausible biological or clinical rationale.

Well-established predictors. ALP (aOR 1.01 per U/L; P = 0.002) reflects osteoblastic activity

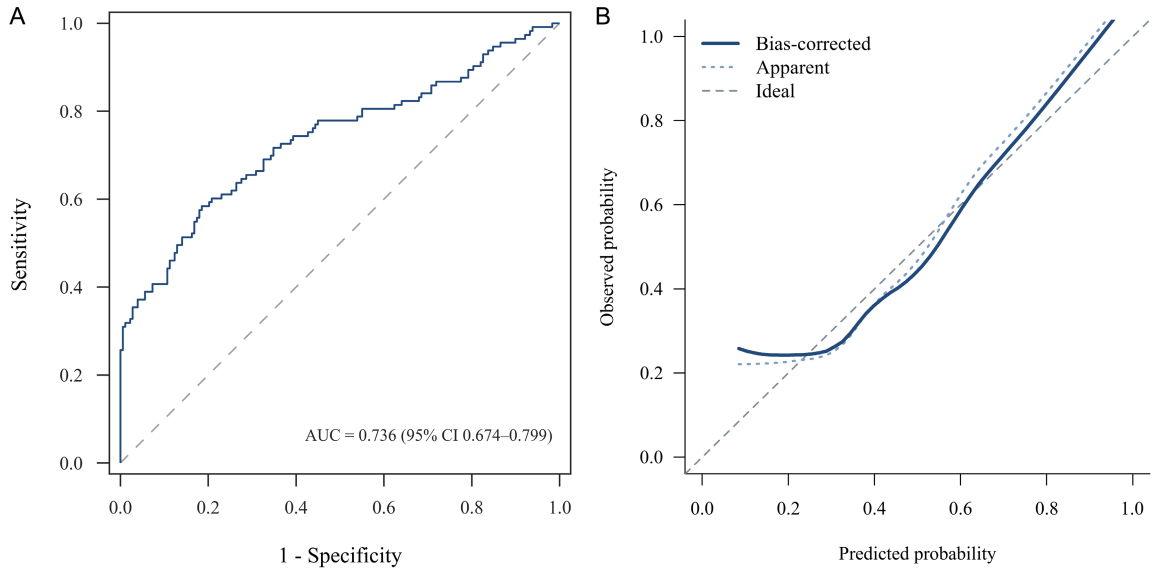


Figure 2. Model discrimination and calibration of the final 7-variable model. A. Receiver operating characteristic (ROC) curve of the final 7-variable logistic regression model on the development dataset (N = 291). The apparent area under the curve was 0.736 (95% CI 0.674 to 0.799), computed by the method of DeLong. The diagonal reference line represents a model with no discriminative ability. B. Calibration plot of the final prediction model. The dashed diagonal line indicates ideal calibration, the dotted line represents apparent calibration, and the solid line shows the bias-corrected calibration curve after bootstrap internal validation (B = 1,000). Gray lines indicate 95% confidence limits, and the rug plot shows the distribution of predicted probabilities.

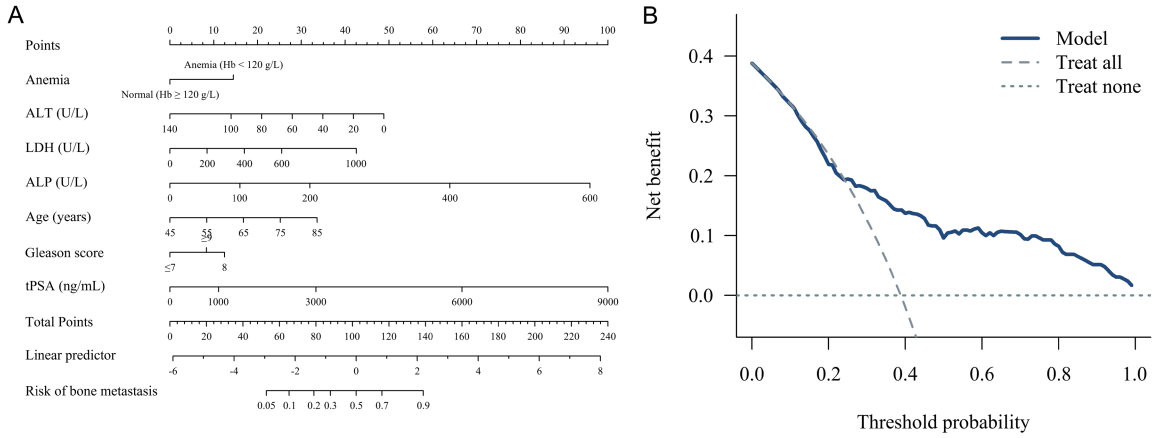


Figure 3. Nomogram and decision curve analysis for the final model. A. Nomogram for predicting the probability of bone metastasis at initial diagnosis of prostate cancer. Each predictor is assigned a point value on the top "Points" axis. The sum of all points is mapped to a linear predictor value and a predicted probability of bone metastasis on the bottom axes. ALT, alanine aminotransferase; LDH, lactate dehydrogenase; ALP, alkaline phosphatase; tPSA, total prostate-specific antigen. B. Decision curve analysis comparing the net benefit of the prediction model against the "treat all" and "treat none" strategies across a range of high-risk threshold probabilities. The model shows a positive net benefit over both default strategies from approximately threshold probability 0.1 to 0.7.

and remains clinically informative in metastatic prostate cancer [16]. tPSA was retained by LASSO despite a small per-unit odds ratio and non-significant multivariable *P* value, because contemporary guidelines and population-based evidence still support PSA in baseline imaging decisions [2, 4, 8, 10, 11]. Gleason score and older age were also associated with higher risk, consistent with recent studies of bone-metastatic presentation in prostate cancer [1, 3, 4].

Anemia. The final model had anemia (Hb<120 g/L) as the strongest predictor (aOR 2.50; P = 0.012). In a larger context of prostate cancer bone metastasis, low hemoglobin is most likely is a signal of an increased burden of disease and possible marrow involvement, which is consistent with a recent review of prostate cancer and bone-metastatic biology [3].

ALT and LDH. The study revealed that ALT had an inverse association with the presence of bone metastasis with an aOR 0.98 per U/L (P = 0.024). This finding is in line with other studies that showed low ALT levels are linked to sarcopenia, frailty, and reduced host reserve [17-20]. The evidence suggests that elevated LDH is associated with worse outcomes in patients with metastatic prostate cancer. Increased LDH levels may also reflect biological processes linked to the disease's bone-metastatic progression [21-23]. LDH showed borderline positive association (aOR 1.00; P = 0.069). We believe that the reduced effect of LDH in our model is partly explained by the high percentage of zero-coded values, which are likely due to unmeasured rather than true zero concentrations. Sensitivity analyses that dealt with zero-coded LDH as missing or left out LDH produced equivalent discrimination and calibration (bias-corrected AUC 0.709 and 0.710 respectively) (Supplementary Tables 8 and 9).

Methodological considerations

A methodological strength of this study was the five-stage hybrid strategy used for variable selection, combining univariable screening, exploratory LASSO, clinical prescreening, data-quality review, and confirmatory LASSO. The unrestricted exploratory LASSO retained more variables than the available events could support and showed separation problems, whereas the clinically prescreened confirmatory model converged stably and aligned with recent guidance on prediction-model tuning, stability, missing-data handling, and validation [11-14]. The broad agreement between the exploratory data-driven stage and the clinically prioritized variables also supports the robustness and face validity of the final predictor set.

A second strength is the explicit examination of data quality before final model development. We excluded cardiac disease history and clinical T stage because their data patterns strongly suggested default-value contamination in the electronic medical record. Had these variables been retained, they would have been more likely to artificially inflate apparent model performance at the expense of generating clinically misleading predictions. The seamless handling of such artifacts is of increasing importance in EMR-based prediction research and is aligned with TRIPOD reporting and recent guidance on missingness and model development in EHR data [9, 24].

Recent methodological studies showed that external validation sample size and missing data handling, data preparation choices, and model stability can materially impact the apparent performance of small clinical cohorts [11, 13, 14]. We therefore prespecified a selection pipeline with five stages, used MICE to deal with missing values, and limited the final model to routinely available variables.

A sensitivity analysis indicated that the apparent performance was similar if the two Gleason categories 8 and ≥9 were collapsed into a single category, ≥8, although there was some numerical gain to the three-group coding, as detailed in Supplementary 3; Supplementary Table 11.

Discrimination performance in the context of existing prediction models

Our model targets a more upstream point in the patient's pathway. It is not designed to refine risk in patients who already have access to advanced imaging or specialized assays. It is instead intended to assist in determining which newly diagnosed patients should undergo further workup. As only routine variables are available at this stage, discrimination cannot be expected to be on a par with models including imaging or specialized biomarkers. The value of the bias-corrected AUC is 0.713, which is moderate and could serve as a first-stage triage tool [2, 10]. Performance reported here also represents an internally validated analysis in a referral-based sample having a greater bone-metastasis prevalence than many registry-based or screening-based populations [23].

Recent studies that made use of biomarkers and imaging have tended to focus on specialized assays or imaging data rather than routine

clinical parameters [22, 25]. These strategies add to, rather than substitute for, a routine-variable first-line scenario.

When compared with recent MRI-radiomics, peri-prostatic adipose-tissue radiomics, exosomal-miRNA, and serum-biomarker models, our routine-variable model is more deployable as it does not require segmentation, slide digitization, or any specialized assay [26, 27]. Several recent models have emerged, including a translational model in 2021, a population-based tool based on 2023 data, and 2026 studies using bone sialoprotein, intermediate-risk cohorts, and PSMA-PET/CT. Many of them still rely on narrower and more dependent data sources; see Supplementary Table 2.

Clinical positioning and regional context

We view the model as an interpretable first-stage triage that could help prioritize adjunct imaging at initial diagnosis. This is not meant to supplant definitive imaging.

The study population was obtained from the First Affiliated Hospital of Xinjiang Medical University. There is no widespread population-based screening for prostate cancer and these patients often present with more advanced disease. Therefore, we see a higher bone-metastasis rate of 38.8% than in several registry-based/screening-prevalent populations [8]. Recent imaging and biomarker studies also support the significance of specialized workups if available [28].

The model is designed to assist with early staging decisions prior to the availability of PSMA PET/CT, a frequent occurrence in certain clinical settings. According to the latest guidelines and reviews, PSMA PET, if available, is the most sensitive option. Whole-body MRI and conventional CT/bone scintigraphy are still valid alternatives [29].

Because the outcome was defined by an any-positive composite of mpMRI, CT and ECT during the same hospitalization, some false-positive misclassification is also possible. Some reports have used the term metastases to label benign degenerative, traumatic, or inflammatory lesions. This occurs when routine clinical reports were not confirmed against a uniform reference standard. The requirement of at least one interpretable modality and conditioning on uninformative cases as missing did maintain internal consistency of the outcome ascertainment, but the residual misclassification could still have impacted discrimination and calibration.

Limitations

There are a few limitations to this study. The findings may not be generalizable due to the single-center and retrospective observational nature of the study at a tertiary referral hospital in Xinjiang. Internal bootstrap resampling was performed for validation, but no external dataset was used. Several variables that potentially contain valuable information (radiomics features, molecular markers and new biomarkers) were not routinely collected so could not be incorporated. The composite interpretation of mp-MRI, CT and bone scintigraphy resulted in an outcome that has imperfect accuracy; thus, misclassification cannot be excluded particularly during the era of PSMA-PET/CT. We could not completely verify the missingness mechanism, and despite retaining zero-coded LDH entries with their original values, the large amount of missingness might have weakened the observed association. After reviewing the data quality, we excluded clinical T stage and cardiac disease history as they appeared to be default-code artifacts. This likely resulted in a model that may have under-represented disease burden compared with a guideline-based model. Even if EPV was deemed acceptable, there is still a precision limitation due to the small sample size, especially among patients with Gleason ≥9. The $P<0.10$ screening threshold on the one hand and the lambda.1se criterion on the other are interesting modeling choices that could be explored in a sensitivity analysis. Benign degenerative, traumatic, or inflammatory bone lesions may also be misclassified under the any-positive framework. Sensitivity analyses where zero-coded LDH was treated as missing or excluded altogether yielded similar discrimination and calibration (bias-corrected AUC 0.709 and 0.710 vs. 0.713 in primary model), suggesting that the overall conclusions were not materially driven by this issue. The main model was fitted on one imputed dataset, although pooled estimates were consistent across imputations.

Conclusion

To sum up, we developed and internally validated a logistic regression-derived nomogram for

bone metastasis prediction at the time of initial prostate cancer diagnosis using routinely available clinical and laboratory variables. Discrimination was considered moderate and calibration stable after bootstrap correction. The tool is parsimonious and interpretable for early risk triage, but it is not ready for clinical use until more external validation in independent multi-center cohorts.

Acknowledgements

We sincerely thank all staff members of the Department of Urology at the First Affiliated Hospital of Xinjiang Medical University for their assistance. This work was supported by the Natural Science Foundation of the Xinjiang Uygur Autonomous Region (grant No. 2023-D01C206), the Regional Collaborative Innovation Special Project of the Autonomous Region, Science and Technology Support Program for Xinjiang (grant No. 2024E02054), the Excellence Youth Science Foundation of the Xinjiang Uygur Autonomous Region (grant No. 2023D01E05), and the National Natural Science Foundation of China (grant No. 82360476), and the 14th Five-Year Plan Autonomous Region Higher Education Institution Characteristic Discipline Program in Public Health and Preventive Medicine.

Disclosure of conflict of interest

None.

Abbreviations

ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the receiver operating characteristic curve; BMI, body mass index; CI, confidence interval; CT, computed tomography; DCA, decision curve analysis; EAU, European Association of Urology; ECT, emission computed tomography (whole-body bone scintigraphy); EPV, events per variable; f/tPSA, free-to-total prostate-specific antigen ratio; fPSA, free prostate-specific antigen; GVIF, generalized variance inflation factor; Hb, hemoglobin; HIS, hospital information system; LASSO, least absolute shrinkage and selection operator; LDH, lactate dehydrogenase; LIS, laboratory information system; MICE, multiple imputation by chained equations; mp-MRI, multiparametric magnetic resonance imaging; NCCN, National Comprehensive Cancer Network; OR, odds ratio; PACS, picture archiving and communication system; PCa, prostate cancer; PET, positron emission tomography; PSA, prostate-specific antigen; PSMA, prostate-specific membrane antigen; SVI, seminal vesicle invasion; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; tPSA, total prostate-specific antigen; VIF, variance inflation factor.

Address correspondence to: Hengqing An, Department of Urology, The First Affiliated Hospital of Xinjiang Medical University, No. 137 Liyushan South Road, High-Tech Zone, Urumqi 830013, Xinjiang Uygur Autonomous Region, China. Tel: +86-0991-4361343; E-mail: anhq-md@xjmu.edu.cn; Ning Tao, School of Public Health, Xinjiang Medical University, No. 567 Shangde North Road, Shuimogou District, Urumqi 830017, Xinjiang Uygur Autonomous Region, China. Tel: +86-0991-2110325; E-mail: 38518412@xjmu.edu.cn

References

[1] Archer Goode E, Wang N and Munkley J. Prostate cancer bone metastases biology and clinical management (Review). Oncol Lett 2023; 25: 163.

[2] Wilson ZJ, Xu G, Tewari SO and Lu Y. Comparison of PSMA-based (18)F-DCFPyL PET/CT and Tc-99m MDP bone scan in detection of bone metastasis in prostate cancer. Am J Nucl Med Mol Imaging 2023; 13: 1-10.

[3] Raychaudhuri R, Lin DW and Montgomery RB. Prostate cancer: a review. JAMA 2025; 333: 1433-1446.

[4] Cornford P, van den Bergh RCN, Briers E, Van den Broeck T, Brunckhorst O, Darraugh J, Eberli D, De Meerleer G, De Santis M, Farolfi A, Gandaglia G, Gillessen S, Grivas N, Henry AM, Lardas M, van Leenders GJLH, Liew M, Linares Espinos E, Oldenburg J, van Oort IM, Oprea-Lager DE, Ploussard G, Roberts MJ, Rouvière O, Schoots IG, Schouten N, Smith EJ, Stranne J, Wiegel T, Willemse PM and Tilki D. EAU-EANM-ESTRO-ESUR-ISUP-SIOG guidelines on prostate cancer-2024 update. Part I: screening, diagnosis, and local treatment with curative intent. Eur Urol 2024; 86: 148-163.

[5] Zhang YF, Zhou C, Guo S, Wang C, Yang J, Yang ZJ, Wang R, Zhang X and Zhou FH. Deep learning algorithm-based multimodal MRI radiomics and pathomics data improve prediction of bone metastases in primary prostate

cancer. J Cancer Res Clin Oncol 2024; 150: 78.
[6] Nie S, Fan B, Gui S, Zou H and Lan M. Predictive impact of T2-MRI radiomics model on initial diagnosis of bone metastasis in prostate cancer patients. BMC Med Imaging 2025; 25: 106.
[7] Sun J, Tian T, Wang N, Jing X, Qiu L, Cui H, Liu Z, Liu J, Yan L and Li D. Pretreatment level of serum sialic acid predicts both qualitative and quantitative bone metastases of prostate cancer. Front Endocrinol (Lausanne) 2024; 15: 1338420.
[8] Li S, Chen C, Zhu H, Lin Q and Yu Z. Risk evaluation of bone metastases and a simple tool for detecting bone metastases in prostate cancer: a population-based study. Comput Math Methods Med 2023; 2023: 9161763.
[9] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, Boulesteix AL, Camaradou JC, Celi LA, Denaxas S, Denniston AK, Glocker B, Golub RM, Harvey H, Heinze G, Hoffman MM, Kengne AP, Lam E, Lee N, Loder EW, Maier-Hein L, Mateen BA, McCradden MD, Oakden-Rayner L, Ordish J, Parnell R, Rose S, Singh K, Wynants L and Logullo P. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024; 385: q902.
[10] Pickersgill NA, Yu R, Vertosick EA, Fletcher SA, Liso N, Kluckert J, Akin O, Ehdaie B, Laudone VP, Eastham JA, Vickers AJ, Touijer KA and Gaffney CD. Diagnostic and prognostic utility of PSMA-PET/CT in node-metastatic prostate cancer. BJU Int 2026; 138: 250-256.
[11] Riley RD, Snell KIE, Archer L, Ensor J, Debray TPA, van Calster B, van Smeden M and Collins GS. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. BMJ 2024; 384: e074821.
[12] Lin S, Groenwold RHH, Mehta HB, Kim JS and Segal JB. Addressing missingness in predictive models that use electronic health record data. Ann Intern Med 2025; 178: 1451-1463.
[13] van Os HJA, Kanning JP, Wermer MJH, Chavannes NH, Numans ME, Ruigrok YM, van Zwet EW, Putter H, Steyerberg EW and Groenwold RHH. Developing clinical prediction models using primary care electronic health record data: the impact of data preparation choices on model performance. Front Epidemiol 2022; 2: 871630.
[14] Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J and Mallett S; PROBAST Group†. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019; 170: 51-58.
[15] Collins GS, Reitsma JB, Altman DG and Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015; 350: g7594.
[16] O'Sullivan JM, Heinrich D, Castro E, George S, Dizdarevic S, Baldari S, Essler M, Jong IJ, Lastoria S, Hammerer PG, Tombal B, James ND, Meltzer J, Sandström P and Sartor O. Alkaline phosphatase decline and pain response as predictors of overall survival benefit in patients treated with radium-223: a post hoc analysis of the REASSURE study. Br J Cancer 2025; 132: 354-360.
[17] Laufer M, Perelman M, Sarfaty M, Itelman E and Segal G. Low alanine aminotransferase, as a marker of sarcopenia and frailty, is associated with shorter survival among prostate cancer patients and survivors. a retrospective cohort analysis of 4064 patients. Eur Urol Open Sci 2023; 55: 38-44.
[18] Moshkovits Y, Chetrit A and Dankner R. The association between frailty biomarkers and 20-year all-cause and cardiovascular mortality among community-dwelling older adults. Postgrad Med 2024; 136: 641-650.
[19] Clayton-Chubb D, Majeed A, Roberts SK, Schneider HG, Commins I, Fitzpatrick J, Woods RL, Ryan J, Hussain SM, Tan N, Lubel JS, Tran C, Hodge AD, McNeil JJ and Kemp WW. Serum transaminases and older adults: distribution and associations with all-cause mortality. J Gerontol A Biol Sci Med Sci 2024; 79: glae203.
[20] Zeng J, Li N, Ge J, Ma H, Sun S, Jing Y, Qian C, Cui R, Qu S and Sheng H. Serum ALT activity and its isoenzymes as potential biomarkers for diagnosis of sarcopenia in older adults: a retrospective, cross-sectional study. BMC Geriatr 2025; 25: 53.
[21] Vauchier C, Thibault C, Velev M, Becker O, Guillaume Z, Ashton E, Zaibet S, Berthou H, Courtinat F, Gervais C, Rochand A, Robbrecht D, Oudard S and Auclin E. Integrated prognostic score in metastatic castration-resistant prostate cancer treated with cabazitaxel - a CABASTY post hoc analysis validated by two international prospective phase 3 trials. Eur Urol Oncol 2026; [Epub ahead of print].
[22] Kato T, Tohi Y, Takamoto A, Shimizu R, Daizumoto K, Katayama S, Kobayashi K, Morinaka H, Nishimura K, Fukuhara H, Morizane S, Furukawa J, Araki M, Shiraishi K, Miyaji Y, Miura N, Inoue K and Sugimoto M; Chu-Shikoku Japan Urological Consortium. Prognostic value of the modified Glasgow Prognostic Score and Prognostic Nutritional Index in prostate cancer

treated with cabazitaxel. Int J Clin Oncol 2026; 31: 834-842.
[23] Kim M, Cong H, Goettl R, Liu J and Fong KW. Bone metastatic progression of prostate cancer is regulated by TRIM28-LDHA-mediated metabolism. Mol Cancer Res 2026; 24: 93-105.
[24] Riley RD, Pate A, Dhiman P, Archer L, Martin GP and Collins GS. Clinical prediction models and the multiverse of madness. BMC Med 2023; 21: 502.
[25] Liu B, Cai Q, Zhao X, Su H, Lin Z, Wu J, Li X, Zhu W, Zou C and Luo Y. A nomogram based on radiomic features from peri-prostatic adipose tissue for predicting bone metastasis in first-time diagnosed prostate cancer patients. Adipocyte 2025; 14: 2517583.
[26] Wang C, Wei XP, Zhou C, Wang J, Zhang YF, He H, Zhang WB, Lv HX, Wang F and Zhou FH. Precise prediction of bone metastases and metastatic burden using exosomal mirnas and radiomics: a multi-center study. J Transl Med 2025; 23: 677.
[27] Liu T, Zhu CH, Yang K, Du Y and Cheng F. PI-RADS v2 score may be as a predictor of bone metastasis in prostate cancer. Indian J Cancer 2026; 63: 82-89.
[28] Zhu J, Shen C, Wang Z, Chen F and Wang T. Predictive risk model for bone metastasis in intermediate-risk prostate cancer: a single-center retrospective analysis. Sci Rep 2026; 16: 23032.
[29] Philippou C, Gloger S, Ubrig B, Bitterlich N, Naseva EK, Sommerfeld HJ, Wiedemann A, Theegarten D, Abdulla H and Philippou S. Estimating bone metastasis risk in prostate cancer: a three-parameter model using bone sialoprotein, ISUP grading, and tumor progression. Urol Int 2026; 110: 154-164.

Supplementary 1

Supplementary methods: data preprocessing details

Data cleaning

Several data quality issues in the raw database were addressed before analysis. First, duplicate columns (variables recorded in both the hospital information system and laboratory information system) were reconciled and reduced to a single entry. Second, some laboratory variables encoded 'not measured' as the numeric value 0 rather than as a missing value, producing spurious zeros. Where this could be reliably adjudicated against the original laboratory reports on a case-by-case basis, the affected zeros were converted to missing values. For LDH, however, a substantial number of zero-coded records (57/291, 19.6%) could not be reliably adjudicated from the source database; the primary model therefore retained the original values, and the impact of these uncertain zeros was further examined in supplementary sensitivity analyses and discussed in the main manuscript. Third, occasional variables suffered from integer underflow (for example, a negative prostate volume consistent with a 32-bit overflow), and such values were set to missing. Fourth, categorical levels that were inconsistently coded across years (for example, Rh blood group recorded as '+/?' in some years and 'positive/negative' in others) were harmonized against a standard dictionary. Fifth, ethnicity was excluded from the analysis because of near-zero variance (a single ethnic group accounted for more than 96% of cases), which would have prevented any meaningful inference for underrepresented subgroups.

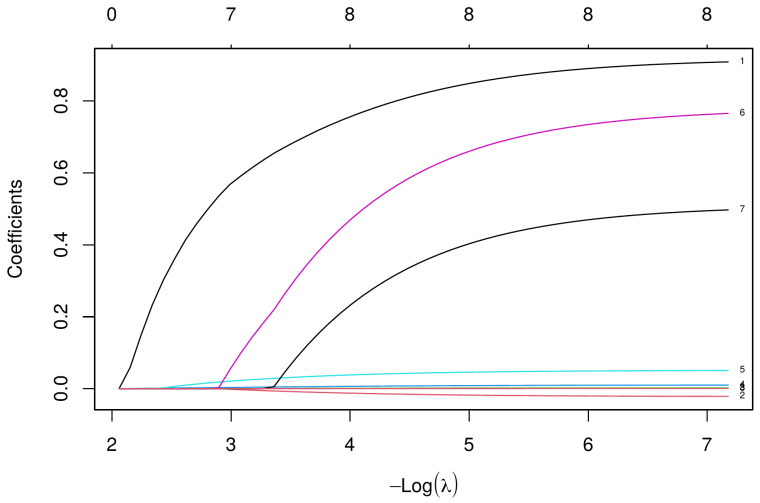
Missing data handling

Variables with more than 50% missingness were dropped, on the grounds that imputation at that level of missingness cannot be reliable. Remaining missing values were addressed using multiple imputation by chained equations (MICE) with the mice package in R (version 3.16), generating m = 5 imputed datasets. A ridge penalty of 0.001 was applied to avoid singularity issues arising from collinearity, and predictor matrix construction used the quickpred function with mincor = 0.2 and minpuc = 0.30. The primary analysis was performed on the first MICE-imputed dataset. As a sensitivity analysis, the prespecified final 7-variable model was refit across all five imputed datasets and estimates were combined using Rubin's rules; results are presented in Supplementary 3.

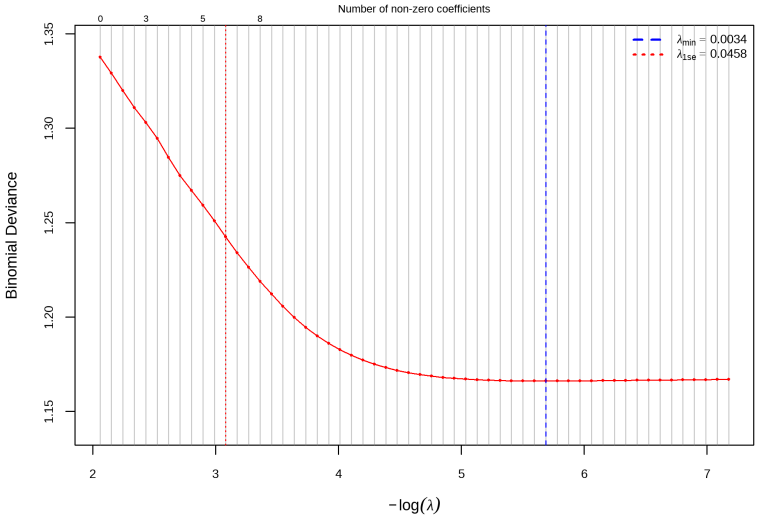
Variable encoding and collinearity management

Continuous variables with an established clinical cut-off were dichotomized to enhance clinical interpretability. Specifically, hemoglobin was dichotomized as anemic versus non-anemic at the standard male threshold of 120 g/L (Hb_anemia). Several variables were excluded to avoid redundancy with other candidate predictors: the sub-staged T category (redundant with the 4-level T stage), primary and secondary Gleason grades (in part-whole relationship with the Gleason sum score), the AST-to-ALT ratio (derived from ALT), and mean corpuscular hemoglobin concentration (correlated with hemoglobin). Two protein-related variables (total protein and albumin) were excluded because of uncertain field identity: their recorded values did not match the expected reference ranges in the data dictionary, suggesting possible field swaps in the source data. Globulin values were consistent with expected reference ranges, so the variable was retained in the candidate-screening set; it did not meet significance thresholds and was not selected at any subsequent stage. The merged prostate volume variable derived from mpMRI and transrectal ultrasound had a missing rate of 57.7% and was therefore not included in modeling. After these steps, 93 candidate predictors were retained for the variable-screening stage.

Supplementary 2



Supplementary Figure 1. LASSO coefficient paths for the confirmatory LASSO stage (Stage 5) applied to the 8 clinically prescreened variables. Each curve shows the regularized coefficient of one predictor as the penalty parameter -log(lambda) changes along the horizontal axis. Numbers at the top indicate how many non-zero coefficients remain at each value of lambda. Seven variables retained non-zero coefficients at lambda.1se and were carried forward to the final multivariable logistic regression model.



Supplementary Figure 2. Ten-fold cross-validation deviance curve for the confirmatory LASSO (Stage 5). The horizontal axis shows -log(lambda); the vertical axis shows the mean binomial deviance across folds with its standard error. Numbers at the top indicate the number of non-zero coefficients retained at each value of lambda. Vertical reference lines mark lambda.min (dashed line; minimum cross-validated deviance, lambda.min = 0.0034) and lambda.1se (dotted line; largest lambda within one standard error of lambda.min, lambda.1se = 0.0458). lambda.1se was used as the selection criterion for the final 7-variable model.

Supplementary Table 1. Descriptive statistics and univariable analysis of all 93 candidate predictors

Variable	Overall (N = 291)	Non-metastatic (N = 178)	Bone metastasis (N = 113)	Crude OR (95% CI)	*P* value
Anemia status	-	-	-	-	<0.001
Normal (Hb≥120 g/L)	239 (82.1)	160 (89.9)	79 (69.9)	-	
Anemia (Hb<120 g/L)	52 (17.9)	18 (10.1)	34 (30.1)	-	
Extracapsular extension (ECE)	-	-	-	-	<0.001
No	197 (67.7)	137 (77.0)	60 (53.1)	-	
Yes	94 (32.3)	41 (23.0)	53 (46.9)	-	
History of cardiac disease [Excluded: default-value contamination]	-	-	-	-	<0.001
Dominant default code (raw No)	251 (86.3)	142 (79.8)	109 (96.5)	-	
Explicit ‘none’ entry variant A	37 (12.7)	33 (18.5)	4 (3.5)	-	
Explicit ‘none’ entry variant B	1 (0.3)	1 (0.6)	0 (0.0)	-	
Yes - coronary artery disease	1 (0.3)	1 (0.6)	0 (0.0)	-	
Yes - other cardiac disease	1 (0.3)	1 (0.6)	0 (0.0)	-	
Seminal vesicle invasion (SVI)	-	-	-	-	<0.001
No	189 (64.9)	131 (73.6)	58 (51.3)	-	
Yes	93 (32.0)	43 (24.2)	50 (44.2)	-	
No (none)	3 (1.0)	2 (1.1)	1 (0.9)	-	
Yes (bilateral)	6 (2.1)	2 (1.1)	4 (3.5)	-	
Alkaline phosphatase (ALP), U/L	71.0 (58.0, 93.2)	70.0 (57.0, 84.4)	77.6 (59.7, 120.3)	1.01 (1.00-1.02)	<0.001
Age at admission, years	74.0 (68.0, 78.0)	73.0 (65.2, 77.0)	76.0 (71.0, 79.0)	1.06 (1.02-1.09)	<0.001
Health insurance type	-	-	-	-	0.002
Urban employee basic medical insurance	225 (77.3)	137 (77.0)	88 (77.9)	-	
Urban resident basic medical insurance	18 (6.2)	12 (6.7)	6 (5.3)	-	
Other	35 (12.0)	27 (15.2)	8 (7.1)	-	
Other social insurance	12 (4.1)	2 (1.1)	10 (8.8)	-	
Urban employee plus urban resident basic medical insurance (dual enrollment)	1 (0.3)	0 (0.0)	1 (0.9)	-	
Red blood cell count (RBC), $\times 10^{12}$/L [Not retained by LASSO at lambda.1se]	4.4 (4.0, 4.8)	4.5 (4.1, 4.9)	4.2 (3.8, 4.7)	0.57 (0.40-0.82)	0.002
Percentage of positive cores, %	53.8 (23.1, 84.6)	46.2 (16.7, 76.9)	66.7 (30.8, 92.3)	1.01 (1.00-1.02)	0.005
Gleason score (3 groups)	-	-	-	-	0.006
≤7	134 (46.0)	95 (53.4)	39 (34.5)	-	
8	103 (35.4)	56 (31.5)	47 (41.6)	-	
≥9	54 (18.6)	27 (15.2)	27 (23.9)	-	
Clinical T stage [Excluded: default-value contamination]	-	-	-	-	0.006
T1	22 (7.6)	16 (9.0)	6 (5.3)	-	
T2	178 (61.2)	95 (53.4)	83 (73.5)	-	
T3	46 (15.8)	32 (18.0)	14 (12.4)	-	
T4	45 (15.5)	35 (19.7)	10 (8.8)	-	

Seven-variable bone metastasis model

Number of biopsy cores	13.0 (12.0, 13.0)	13.0 (12.0, 13.0)	13.0 (12.0, 13.0)	0.77 (0.63-0.94)	0.009
Suspected lymph node involvement	-	-	-	-	0.012
No	212 (72.9)	139 (78.1)	73 (64.6)	-	
Yes	79 (27.1)	39 (21.9)	40 (35.4)	-	
Plasma fibrinogen, g/L	3.3 (2.8, 3.9)	3.2 (2.8, 3.9)	3.5 (2.8, 4.4)	1.43 (1.08-1.91)	0.013
Calcium, mmol/L	2.2 (2.1, 2.3)	2.2 (2.1, 2.3)	2.2 (2.1, 2.2)	0.45 (0.24-0.85)	0.014
Free PSA (fPSA), ng/mL	2.4 (1.0, 7.0)	1.9 (0.9, 4.1)	4.4 (1.6, 23.8)	1.01 (1.00-1.01)	0.015
Number of positive cores	6.0 (3.0, 10.0)	6.0 (2.0, 10.0)	7.0 (4.0, 11.0)	1.07 (1.01-1.13)	0.019
Total PSA (tPSA), ng/mL	19.6 (10.4, 59.0)	18.0 (10.9, 42.9)	27.7 (8.7, 88.0)	1.00 (1.00-1.00)	0.021
Prothrombin time activity, %	104.4 (88.5, 114.2)	106.9 (90.0, 117.3)	100.0 (85.3, 109.8)	0.99 (0.97-1.00)	0.022
Lactate dehydrogenase (LDH), U/L	154.8 (0.0, 192.3)	153.6 (0.0, 182.3)	167.0 (0.0, 201.0)	1.00 (1.00-1.00)	0.035
Alanine aminotransferase (ALT), U/L	20.0 (14.0, 27.3)	21.5 (15.8, 28.0)	17.9 (12.3, 25.9)	0.98 (0.96-1.00)	0.040
Hypertension	-	-	-	-	0.040
No	193 (66.3)	110 (61.8)	83 (73.5)	-	
Yes	98 (33.7)	68 (38.2)	30 (26.5)	-	
Free-to-total PSA ratio (duplicate source field)	0.2 (0.1, 0.3)	0.2 (0.1, 0.3)	0.1 (0.1, 0.2)	0.19 (0.04-1.00)	0.051
International normalized ratio (INR)	1.0 (1.0, 1.1)	1.0 (0.9, 1.0)	1.0 (1.0, 1.1)	6.22 (1.00-38.82)	0.051
Marital status	-	-	-	-	0.071
Married	270 (92.8)	169 (94.9)	101 (89.4)	-	
Widowed	20 (6.9)	8 (4.5)	12 (10.6)	-	
Divorced	1 (0.3)	1 (0.6)	0 (0.0)	-	
Gleason secondary grade	-	-	-	-	0.077
Gleason grade 2	1 (0.3)	1 (0.6)	0 (0.0)	-	
Gleason grade 3	105 (36.1)	71 (39.9)	34 (30.1)	-	
Gleason grade 4	146 (50.2)	88 (49.4)	58 (51.3)	-	
Gleason grade 5	39 (13.4)	18 (10.1)	21 (18.6)	-	
Magnesium, mmol/L	0.9 (0.8, 0.9)	0.9 (0.8, 0.9)	0.9 (0.8, 0.9)	0.29 (0.07-1.22)	0.091
Educational level	-	-	-	-	0.095
Illiterate	48 (16.5)	33 (18.5)	15 (13.3)	-	
Primary school	47 (16.2)	28 (15.7)	19 (16.8)	-	
Junior high school	32 (11.0)	25 (14.0)	7 (6.2)	-	
Senior high school	57 (19.6)	29 (16.3)	28 (24.8)	-	
Technical secondary school	36 (12.4)	18 (10.1)	18 (15.9)	-	
Junior college	35 (12.0)	20 (11.2)	15 (13.3)	-	
Bachelor's degree or above	36 (12.4)	25 (14.0)	11 (9.7)	-	
Delta bilirubin, μmol/L	0.0 (0.0, 5.2)	1.6 (0.0, 5.7)	0.0 (0.0, 4.2)	0.94 (0.87-1.01)	0.098
Total cholesterol, mmol/L	3.9 (3.0, 4.7)	4.0 (3.2, 4.8)	3.6 (2.7, 4.6)	0.90 (0.79-1.03)	0.115

Neuroendocrine differentiation	-	-	-	-	0.136
No	287 (98.6)	176 (98.9)	111 (98.2)	-	
No (entry variant)	2 (0.7)	2 (1.1)	0 (0.0)	-	
Yes	2 (0.7)	0 (0.0)	2 (1.8)	-	
Lymphocyte count, $\times 10^9$/L	1.5 (1.1, 1.8)	1.5 (1.2, 1.8)	1.5 (1.1, 1.8)	0.74 (0.50-1.11)	0.144
Aspartate aminotransferase (AST), U/L	23.5 (17.6, 29.9)	23.8 (17.9, 31.5)	23.2 (16.9, 28.1)	0.98 (0.96-1.01)	0.144
Indirect bilirubin, μmol/L	0.0 (0.0, 7.7)	0.4 (0.0, 8.0)	0.0 (0.0, 6.0)	0.97 (0.93-1.01)	0.152
Prothrombin time (PT), s	11.2 (10.8, 12.0)	11.1 (10.7, 11.8)	11.4 (11.0, 12.2)	1.12 (0.95-1.32)	0.171
Normal control PT, s	11.5 (11.2, 11.5)	11.5 (11.2, 11.5)	11.5 (11.2, 11.5)	1.52 (0.82-2.82)	0.179
HDL cholesterol, mmol/L	0.9 (0.7, 1.2)	0.9 (0.7, 1.2)	0.9 (0.7, 1.2)	0.73 (0.45-1.17)	0.188
Red cell distribution width (RDW), %	13.1 (12.6, 13.9)	13.1 (12.5, 13.6)	13.2 (12.7, 14.1)	1.10 (0.95-1.28)	0.194
Neutrophil percentage, %	64.2 (57.7, 69.9)	64.0 (57.4, 69.6)	64.5 (58.3, 70.4)	1.02 (0.99-1.04)	0.207
Gamma-glutamyl transferase (GGT), U/L	22.9 (17.2, 33.7)	23.0 (17.6, 32.8)	22.4 (16.3, 34.7)	1.00 (0.99-1.00)	0.216
Globulin, g/L	38.2 (35.2, 41.8)	38.6 (35.4, 41.8)	37.7 (34.8, 41.8)	0.97 (0.93-1.02)	0.223
Mean corpuscular hemoglobin (MCH), pg	30.8 (29.6, 31.9)	31.0 (29.7, 31.9)	30.7 (29.6, 31.8)	0.93 (0.83-1.04)	0.231
Perineural invasion	-	-	-	-	0.236
No	234 (80.4)	147 (82.6)	87 (77.0)	-	
Yes	55 (18.9)	29 (16.3)	26 (23.0)	-	
No (entry variant)	2 (0.7)	2 (1.1)	0 (0.0)	-	
Eosinophil percentage, %	2.0 (1.0, 3.2)	2.0 (1.1, 3.0)	1.9 (0.9, 3.6)	1.08 (0.95-1.22)	0.245
Diabetes mellitus	-	-	-	-	0.247
No	205 (70.4)	121 (68.0)	84 (74.3)	-	
Yes	86 (29.6)	57 (32.0)	29 (25.7)	-	
Rh blood group	-	-	-	-	0.256
Rh positive	278 (95.5)	172 (96.6)	106 (93.8)	-	
Rh negative	13 (4.5)	6 (3.4)	7 (6.2)	-	
Basophil percentage, %	0.4 (0.2, 0.7)	0.4 (0.2, 0.7)	0.4 (0.2, 0.6)	0.63 (0.28-1.41)	0.262
Serum cholinesterase (ChE), U/L	5504.6 (0.0, 7061.6)	5568.7 (0.0, 7241.7)	5491.8 (0.0, 6796.2)	1.00 (1.00-1.00)	0.317
LDL cholesterol, mmol/L	2.4 (1.6, 3.0)	2.4 (1.9, 3.0)	2.3 (1.3, 3.0)	0.93 (0.80-1.08)	0.339
Lymphocyte percentage, %	25.3 (20.2, 30.7)	25.4 (20.7, 30.8)	25.3 (20.1, 30.2)	0.99 (0.96-1.01)	0.340
Clinical N stage	-	-	-	-	0.342
N0	205 (70.4)	129 (72.5)	76 (67.3)	-	
N1	86 (29.6)	49 (27.5)	37 (32.7)	-	
Sodium, mmol/L	139.7 (137.5, 142.3)	140.0 (137.6, 142.5)	139.2 (137.5, 142.1)	0.99 (0.98-1.01)	0.354
Rectal invasion	-	-	-	-	0.373
No	256 (88.0)	159 (89.3)	97 (85.8)	-	
Yes	35 (12.0)	19 (10.7)	16 (14.2)	-	
Normal control TT, s	20.3 (20.0, 20.5)	20.3 (20.0, 20.5)	20.3 (20.0, 20.5)	0.98 (0.93-1.03)	0.384

Bladder invasion	-	-	-	-	0.401
No	246 (84.5)	153 (86.0)	93 (82.3)	-	
Yes	45 (15.5)	25 (14.0)	20 (17.7)	-	
Chloride, mmol/L	104.9 (102.6, 107.6)	104.6 (102.9, 107.4)	104.9 (102.0, 107.9)	0.99 (0.98-1.01)	0.401
Occupation	-	-	-	-	0.410
Retired	221 (75.9)	133 (74.7)	88 (77.9)	-	
Farmer	33 (11.3)	22 (12.4)	11 (9.7)	-	
Other	17 (5.8)	8 (4.5)	9 (8.0)	-	
Technical staff	12 (4.1)	8 (4.5)	4 (3.5)	-	
Civil servant	5 (1.7)	5 (2.8)	0 (0.0)	-	
Self-employed	2 (0.7)	1 (0.6)	1 (0.9)	-	
Unemployed	1 (0.3)	1 (0.6)	0 (0.0)	-	
Mean corpuscular volume (MCV), fL	92.4 (89.3, 95.8)	92.2 (89.1, 95.4)	92.6 (89.9, 96.3)	1.02 (0.97-1.07)	0.434
Eosinophil count, $\times 10^9$/L	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	2.11 (0.32-14.01)	0.441
Potassium, mmol/L	3.8 (3.5, 4.1)	3.7 (3.5, 4.1)	3.9 (3.5, 4.1)	0.98 (0.92-1.04)	0.443
Basophil count, $\times 10^9$/L	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.01 (0.00-3012.72)	0.469
Creatine kinase (CK), U/L	58.7 (30.5, 97.8)	59.2 (33.2, 97.1)	58.0 (26.0, 98.0)	1.00 (1.00-1.00)	0.485
Glucose, mmol/L	5.4 (4.7, 6.4)	5.4 (4.8, 6.8)	5.2 (4.7, 6.3)	0.97 (0.87-1.07)	0.503
Monocyte percentage, %	7.9 (6.4, 9.1)	7.8 (6.2, 9.1)	8.1 (6.6, 9.3)	1.03 (0.94-1.13)	0.544
Total bilirubin, μmol/L	12.2 (10.0, 15.9)	12.8 (10.1, 15.9)	11.3 (9.8, 15.6)	1.01 (0.99-1.02)	0.545
Plateletcrit	0.2 (0.2, 0.3)	0.2 (0.2, 0.2)	0.2 (0.2, 0.3)	2.88 (0.08-108.46)	0.568
White blood cell count (WBC), $\times 10^9$/L	6.1 (5.2, 7.3)	6.1 (5.1, 7.3)	6.1 (5.3, 7.3)	1.03 (0.93-1.15)	0.580
APTT ratio	1.0 (1.0, 1.1)	1.0 (1.0, 1.1)	1.1 (1.0, 1.1)	0.91 (0.63-1.30)	0.592
Creatinine, μmol/L	76.0 (63.8, 90.0)	77.3 (64.0, 89.8)	76.0 (62.4, 93.3)	1.00 (1.00-1.00)	0.622
Triglycerides, mmol/L	1.2 (0.9, 1.7)	1.2 (0.9, 1.8)	1.1 (0.9, 1.6)	0.94 (0.74-1.20)	0.631
Free-to-total PSA ratio (f/tPSA)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	1.26 (0.47-3.37)	0.640
Thrombin time (TT), s	19.5 (18.6, 20.6)	19.5 (18.6, 20.5)	19.5 (18.5, 21.1)	1.01 (0.96-1.07)	0.656
Coronary heart disease	-	-	-	-	0.679
No	215 (73.9)	130 (73.0)	85 (75.2)	-	
Yes	76 (26.1)	48 (27.0)	28 (24.8)	-	
Platelet distribution width (PDW), fL	12.3 (10.3, 15.6)	12.3 (10.3, 15.6)	12.1 (10.2, 15.7)	0.99 (0.91-1.07)	0.755
Dyslipidemia	-	-	-	-	0.757
No	216 (74.2)	131 (73.6)	85 (75.2)	-	
Yes	75 (25.8)	47 (26.4)	28 (24.8)	-	
Residence	-	-	-	-	0.766
Urban	244 (83.8)	149 (83.7)	95 (84.1)	-	
Rural	45 (15.5)	27 (15.2)	18 (15.9)	-	
Urban/town	2 (0.7)	2 (1.1)	0 (0.0)	-	

Intraductal carcinoma	-	-	-	-	0.776
No	287 (98.6)	175 (98.3)	112 (99.1)	-	
No (entry variant)	2 (0.7)	2 (1.1)	0 (0.0)	-	
Yes	2 (0.7)	1 (0.6)	1 (0.9)	-	
Hematocrit, %	39.3 (28.8, 42.9)	39.7 (26.6, 43.5)	38.4 (31.1, 41.5)	1.00 (0.98-1.01)	0.778
Mean platelet volume (MPV), fL	10.2 (9.5, 10.8)	10.2 (9.6, 10.8)	10.3 (9.4, 10.8)	0.97 (0.78-1.21)	0.789
Platelet count, $\times 10^9$/L	205.0 (171.5, 253.0)	205.0 (173.2, 250.5)	203.0 (167.0, 257.0)	1.00 (1.00-1.00)	0.812
Urea, mmol/L	6.1 (5.0, 7.6)	6.1 (5.1, 7.9)	6.1 (4.8, 7.3)	1.00 (0.99-1.01)	0.828
History of stroke or TIA	-	-	-	-	0.839
No	217 (74.6)	132 (74.2)	85 (75.2)	-	
Yes	74 (25.4)	46 (25.8)	28 (24.8)	-	
History of cardiovascular disease	-	-	-	-	0.856
No	218 (74.9)	134 (75.3)	84 (74.3)	-	
Yes	73 (25.1)	44 (24.7)	29 (25.7)	-	
Neutrophil count, $\times 10^9$/L	3.8 (3.1, 4.8)	3.9 (3.1, 4.8)	3.6 (3.1, 4.6)	0.99 (0.88-1.11)	0.879
Normal control APTT, s	30.2 (30.0, 30.7)	30.2 (30.0, 30.8)	30.2 (30.0, 30.6)	0.98 (0.78-1.23)	0.880
Inorganic phosphorus, mmol/L	1.1 (1.0, 1.2)	1.1 (1.0, 1.2)	1.1 (1.0, 1.2)	0.96 (0.38-2.42)	0.926
Osmolality, mOsm/kg	280.2 (0.0, 292.1)	282.3 (0.0, 292.7)	277.1 (0.0, 290.6)	1.00 (1.00-1.00)	0.947
Activated partial thromboplastin time (APTT), s	31.4 (29.9, 33.5)	31.1 (29.7, 33.4)	31.6 (30.4, 33.7)	1.00 (0.93-1.08)	0.956
Uric acid, μmol/L	330.0 (265.9, 379.9)	322.0 (264.0, 362.7)	341.0 (276.4, 399.5)	1.00 (1.00-1.00)	0.956
Serum bicarbonate, mmol/L	24.8 (22.8, 27.1)	24.5 (22.8, 27.1)	25.4 (22.7, 27.4)	1.00 (0.94-1.06)	0.962
Direct bilirubin, μmol/L	-	-	-	-	0.963
0 μmol/L	146 (50.2)	88 (49.4)	58 (51.3)	-	
0.3 μmol/L	143 (49.1)	88 (49.4)	55 (48.7)	-	
13.1 μmol/L	1 (0.3)	1 (0.6)	0 (0.0)	-	
80.53 μmol/L	1 (0.3)	1 (0.6)	0 (0.0)	-	
Biopsy approach	-	-	-	-	1.000
Transrectal (TRUS-guided)	288 (99.0)	176 (98.9)	112 (99.1)	-	
Combined approach	3 (1.0)	2 (1.1)	1 (0.9)	-	
Anti-D antibody	-	-	-	-	1.000
Anti-D positive	289 (99.3)	177 (99.4)	112 (99.1)	-	
Anti-D negative	2 (0.7)	1 (0.6)	1 (0.9)	-	

All 93 candidate predictors from initial univariable screening, sorted by *P* value ascending. Continuous variables: median (Q1, Q3), OR and P from univariable logistic regression. Categorical variables: P from chi-squared or Fisher's exact test; sub-rows show n (%). For history of cardiac disease, the dominant raw code is labeled 'dominant default code', and explicitly recorded negative entries are labeled 'explicit 'none' entry variant A' and 'explicit 'none' entry variant B' to reflect likely default-value contamination. History of cardiac disease and clinical T stage were excluded from modeling because of likely default-value contamination. Red blood cell count was retained at lambda.min but not at lambda.1se. Direct bilirubin was imported as categorical because only four unique measurement values were present. The duplicate source field denotes a duplicate source column for the same free-to-total PSA ratio measurement. OR, odds ratio; CI, confidence interval; Q1/Q3, first/third quartile.

Supplementary Table 2. Landscape comparison of published prostate cancer bone-metastasis prediction models

Study (Year)	Country/ Setting	Sample size	BM rate (%)	Predictor domain	N predictors	Method	AUC (internal)	External validation	Calibration reported	TRIPOD compliant	Key limitation
Chen et al. [2019]	China, two-center retrospective	308 development (+51 cross-validation; +43 external validation)	40.6 (125/308)	Clinical + pathology + laboratory + optional MRI PI-RADS v2	4 (BGS, cTx, tPSA, ALP); +PI-RADS in 80-patient subset	Multivariable logistic nomogram; PI-RADS incremental analysis	0.910 (clinical-lab model); 0.934 (+PI-RADS subset; base subset 0.884)	Yes (independent hospital dataset, n = 43; sensitivity/specificity reported)	Yes (calibration curve; mean absolute error 0.03)	Not stated	Small external validation set; PI-RADS-enhanced model tested in only 80/308 patients
Lu et al. [2021]	USA, SEER registry	168,414 (84,207 train/84,207 split-validation)	3.2 (5,354/ 168,414)	Registry clinical + demographic + distant-metastasis status	6 in nomogram 1; 9 in nomogram 2 (adds brain, liver, lung metastasis)	Multivariable logistic nomogram	AUC 0.911 train; 0.910 split-validation (nomogram 2); C-index 0.917/0.915	No (random split-sample only)	Yes (calibration plots; DCA/CIC reported)	Yes (TRIPOD checklist stated)	Coarse SEER variable granularity; split-sample validation only; no routine laboratory or imaging markers
Liu WC et al. [2021]	USA SEER + China Nanchang external	207,137 (SEER) + 644 (external, Nanchang)	3.25 (SEER); 18.2 external	Registry clinical + demographic	8 ML inputs; 4 LR-independent factors (T, N, PSA, Gleason)	XGBoost compared with DT, RF, MLP, LR, and naive Bayes	0.951 (10-fold CV); 0.955 (SEER internal test, XGBoost)	Yes (644 Chinese patients; external AUC 0.962)	Not reported	Not stated	SEER categorical coding and class imbalance; external validation from one Chinese center only
Li et al. [2023]	USA, SEER registry	75,698	5.07 (3,835/ 75,698)	Registry clinical + demographic	2 dominant tree variables (PSA and lymph-node score); broader logistic covariates	Multivariable logistic/Cox regression + classification tree	AUC not reported (classification-tree accuracy 95.90%)	No	Not reported	Not stated	Discrimination and calibration metrics not fully reported; low BM prevalence unlike referral cohorts
Zhang W et al. [2020]	China, single-center	116 (81 train/35 split-validation)	47.4 (55/116)	mpMRI radiomics + clinical	12 radiomic features + tPSA in nomogram	LASSO-logistic mpMRI radiomics nomogram	0.93 (training); 0.92 (split-validation)	No (split-sample only)	Yes (calibration and DCA reported)	Not stated	Small sample and no independent external validation; requires MRI tumor segmentation
Zhang YF et al. [2024]	China, single-center (Gansu)	211 (169 train/42 split-validation)	50.2 (106/211)	mpMRI radiomics + pathomics + deep transfer learning	44 radiomics + 23 DTL + 13 pathomics features selected by LASSO	ResNet-50 feature extraction + machine-learning fusion model	0.93 (combined radiomics + DTL + pathomics, validation set)	No (split-sample only)	Yes (calibration curves and DCA reported)	Not stated	Single-center enriched cohort; requires MRI segmentation and digitised histopathology slides
Liu BH et al. [2024]	China, single-center	156	26.9 (42/156)	MRI-measured PPAT + clinical	3 independent predictors (normalized PPAT, cN, ALP); 6 in nomogram	Multivariable logistic nomogram	C-index 0.856; AUC 0.775 (normalized PPAT alone)	No	Yes (calibration curve and DCA reported)	Not stated	Requires MRI PPAT volumetric segmentation; no independent validation set
Liu B et al. [2025]	China, single-center	151	26.5 (40/151)	PPAT T2-MRI radiomics + clinical	3 independent predictors (Radscore, ALP, cN); Radscore from 16 LASSO-selected features	LASSO-logistic radiomics nomogram; internal bootstrap B = 1,000	C-index 0.908 (development); 0.903 (bootstrap-validated); Radscore AUC 0.888	No (internal bootstrap only)	Yes (calibration and DCA reported)	Not stated	Requires PPAT segmentation and radiomics feature extraction; small sample and no external validation
Yu et al. [2020]	China, single-center (Xinjiang)	150	40.0 (60/150)	Serum biomarkers	3 biomarkers (T-PSA, F-PSA, ProGRP)	Multivariable logistic regression + ROC biomarker combinations	0.752 (ProGRP); 0.885 (T-PSA); 0.919 (F-PSA); 0.941 combined	No	Not reported	Not stated	Biomarker-only study without external validation; ProGRP is not routinely available in many institutions

Sun et al. [2024]	China, single-center (Qilu Hospital)	1,202 screened; 461 selected; 224 definitive BM/non-BM for ROC	13.9 (64/461); 28.6 among definitive ECT groups	Serum biomarker (sialic acid) + clinical/laboratory covariates	1 primary biomarker (SA); multivariable analyses adjusted for AKP/PSA/Gleason etc.	ROC cut-off + multivariable logistic regression; dynamic nomograms	0.767 (SA for BM); 0.766 (SA for HVD)	No	Yes (nomogram calibration curves and DCA reported)	Not stated	Single-center retrospective design; definitive BM comparison used a selected ECT subset; no external validation
This study [2026]	China, single-center retrospective (Xinjiang)	291	38.8 (113/291)	Routine clinical + laboratory + pathology	7 (Anemia, ALT, LDH, ALP, Age, Gleason 3-group, tPSA)	Multivariable logistic regression (hybrid LASSO, five-stage selection)	Apparent 0.736 (95% CI 0.674-0.799); bootstrap-corrected 0.713	No (internal bootstrap B = 1,000)	Yes (bias-corrected slope 0.933, intercept -0.025, Emax 0.050, Brier 0.197)	Yes	External validation pending

Selected published bone-metastasis prediction studies summarized in Supplementary Table 2 are also cited in the main manuscript. Abbreviations: AUC, area under the receiver operating characteristic curve; ALP, alkaline phosphatase; ALT, alanine aminotransferase; BGS, biopsy Gleason score; BM, bone metastasis; cN, clinical nodal stage; cT/cTx, clinical tumour stage; DCA, decision curve analysis; DT, decision tree; DTL, deep transfer learning; fPSA, free prostate-specific antigen; HVD, high-volume disease; LASSO, least absolute shrinkage and selection operator; LDH, lactate dehydrogenase; LR, logistic regression; LVD, low-volume disease; MLP, multilayer perceptron; mpMRI, multiparametric magnetic resonance imaging; NBC, naive Bayes classifier; PCa, prostate cancer; PI-RADS, Prostate Imaging-Reporting and Data System; PPAT, peri-prostatic adipose tissue; ProGRP, pro-gastrin-releasing peptide; PSA, prostate-specific antigen; Radscore, radiomics score; RF, random forest; SA, sialic acid; SEER, Surveillance, Epidemiology, and End Results; TRIPOD+AI, Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis, updated with artificial-intelligence extensions; tPSA, total prostate-specific antigen; XGBoost/XGB, extreme gradient boosting. Notes on validation terminology. “External validation” in this table refers to evaluation in a geographically and institutionally independent dataset that was not used for model development. Split-sample approaches (random or chronological partitioning of a single dataset) and internal bootstrap resampling are treated as internal validation only, in accordance with TRIPOD+AI recommendations. Notes on missing cells. Values of “Not reported” indicate that the corresponding information was not recoverable from the primary publication. For studies reporting performance as a C-index rather than an AUC, the original metric label has been preserved. Source of data. Numeric values - sample size, bone-metastasis rate, AUC/C-index, calibration, and validation status - were transcribed directly from the corresponding primary publications, using abstracts and full-text methods/results tables where required.

Supplementary Table 3. LASSO variable selection (10-fold CV)

Variable	Coefficient at lambda.min	Coefficient at lambda.1se	Retained at lambda.1se
Anemia status: Anemia (Hb<120 g/L) vs. Normal (Hb≥120 g/L)	0.8809	0.5917	Yes
Gleason score (3 groups): 8 vs. ≤7	0.7184	0.0996	Yes
Gleason score (3 groups): ≥9 vs. ≤7	0.4553	0.0000	Shrunk to 0 (variable retained)
Age at admission, years	0.0485	0.0231	Yes
Alanine aminotransferase (ALT), U/L	-0.0197	-0.0028	Yes
Alkaline phosphatase (ALP), U/L	0.0094	0.0038	Yes
Lactate dehydrogenase (LDH), U/L	0.0024	0.0002	Yes
Total PSA (tPSA), ng/mL	0.0006	0.0001	Yes
Red blood cell count (RBC), $\times 10^{12}$/L	-0.0932	-	No

LASSO penalized logistic regression with 10-fold cross-validation. Clinical T stage and history of cardiac disease excluded prior to LASSO (data quality concerns). Each category comparison of multi-level predictors is shown as a separate row. Gleason ≥9 vs. ≤7: coefficient regularized to zero at lambda.1se, but the Gleason variable was retained because the 8 vs. ≤7 contrast survived. ‘Shrunk to 0 (variable retained)’: coefficient reached zero while the parent variable remained selected. Red blood cell count: retained at lambda.min only. LASSO, least absolute shrinkage and selection operator.

Supplementary Table 4. Apparent performance of the final model

Metric	Value
AUC (95% CI)	0.736 (0.674-0.799)
Brier score	0.186
Optimal threshold (Youden)	0.436
Sensitivity, %	58.4
Specificity, %	81.5
Positive predictive value, %	66.7
Negative predictive value, %	75.5

Model performance on training dataset (N = 291). AUC, area under the ROC curve; 95% CI by DeLong's method. Optimal threshold maximises Youden's index. PPV, positive predictive value; NPV, negative predictive value.

Supplementary Table 5. Bootstrap internal validation (B = 1,000)

Metric	Apparent	Optimism	Bias-corrected
AUC	0.736	0.024	0.713
Calibration slope	1.000	0.067	0.933
Calibration intercept	0.000	-	-0.025
Brier score	0.186	-0.010	0.197
Emax	0.000	-	0.050

Bootstrap internal validation (B = 1,000 resamples) was performed using the rms package in R. AUC derived from Somers' Dxy as AUC = (1 + Dxy)/2. Calibration slope and intercept: ideal values 1.0 and 0.0 respectively. Emax: maximum absolute calibration error. Em dash (-): no meaningful optimism estimate. Brier score: lower is better. Values were rounded independently; therefore, apparent minus optimism may not exactly equal the displayed bias-corrected value.

Supplementary 3

Sensitivity analyses

Pooled estimates across imputations

As a sensitivity analysis, we refit the prespecified final 7-variable logistic regression model across all five MICE-imputed datasets without repeating imputation or variable selection. Regression coefficients were combined across imputations using Rubin's rules. For model discrimination, the apparent area under the receiver operating characteristic curve (AUC) was calculated separately within each imputed dataset and summarized as the mean and range across imputations. Bootstrap optimism-correction was not repeated for this sensitivity analysis. The aim was to assess whether reliance on a single imputed dataset for the primary analysis materially affected coefficient estimates or model discrimination.

Supplementary Table 6. Pooled coefficient estimates compared with the primary single-imputation analysis

Variable	Primary β	Pooled β	Pooled 95% CI	Pooled P	Δβ
Intercept	-5.594	-5.213	-8.068 to -2.358	<0.001	+0.381
Anemia (Hb<120 g/L)	0.918	0.677	-0.125 to 1.479	0.097	-0.241
ALT	-0.022	-0.024	-0.045 to -0.003	0.023	-0.002
LDH	0.003	0.002	-0.003 to 0.007	0.403	-0.001
ALP	0.010	0.009	0.002 to 0.016	0.017	-0.001
Age	0.051	0.050	0.014 to 0.086	0.007	-0.001
Gleason 8 vs. ≤7	0.779	0.791	0.173 to 1.409	0.012	+0.012
Gleason ≥9 vs. ≤7	0.510	0.698	-0.122 to 1.517	0.094	+0.188
tPSA	0.001	0.001	-0.001 to 0.002	0.355	-0.000

Primary β: coefficients from logistic regression fitted on the first MICE-imputed dataset (same values as main manuscript **Table 3**). Pooled β: coefficients combined across m = 5 MICE imputations using Rubin's rules. Δβ: pooled minus primary. Reference categories: Anemia status - Normal (Hb≥120 g/L); Gleason score - Gleason ≤7. CI, confidence interval.

Supplementary Table 7. Pooled discrimination compared with the primary single-imputation analysis

Metric	Primary (single)	Pooled (Rubin)
Apparent AUC	0.736 (95% CI 0.674-0.799)	0.731 (range 0.715-0.754)

Primary AUC and 95% CI: from the model fitted on the first MICE-imputed dataset (apparent performance). Pooled AUC: mean and observed range across the five imputed datasets. CI, confidence interval.

Interpretation

Pooled estimates were broadly consistent with the primary single-imputation analysis in terms of model discrimination (mean pooled AUC 0.731 vs. primary apparent AUC 0.736) and the direction and clinical interpretation of all predictors. Differences in individual coefficient magnitudes - most notably anemia (pooled 0.677 vs. primary 0.918) and Gleason ≥9 vs. ≤7 (pooled 0.698 vs. primary 0.510) - reflect across-imputation variability typical of small-sample logistic regression. The signs of all coefficients were preserved and the qualitative conclusions of the primary model were unchanged. Calibration metrics from logistic recalibration on the same imputed datasets used for fitting are uninformative by construction (slope ≡ 1, intercept ≡ 0); the bootstrap optimism-corrected calibration reported in the main manuscript should be interpreted as the primary calibration assessment.

Sensitivity analysis for zero-coded LDH

For both sensitivity models, optimism correction was repeated using the same performance metrics as in the primary analysis. Because these analyses were intended as sensitivity checks rather than primary validation, 300 bootstrap resamples were used. The primary model retained the original LDH values. Reclassifying zero-coded LDH entries as missing and repeating MICE with refitting of the same 7-variable model yielded similar performance (apparent AUC 0.735, bias-corrected AUC 0.709, apparent Brier 0.188, bias-corrected Brier 0.200). A 6-variable model excluding LDH also performed similarly (apparent AUC 0.730, bias-corrected AUC 0.710, apparent Brier 0.190, bias-corrected Brier 0.199). These results suggest that the primary model was not materially dependent on the handling of zero-coded LDH values.

Supplementary Table 8. Sensitivity analysis for zero-coded LDH: core performance metrics

Model	Apparent *AUC*	Bias-corrected *AUC*	Apparent *Brier*	Bias-corrected *Brier*	Bias-corrected slope	Bias-corrected intercept	Corrected *Emax*
Primary model	0.736	0.713	0.186	0.197	0.933	-0.025	0.050
LDH zero coded as missing (7-variable)	0.735	0.709	0.188	0.200	0.916	-0.027	0.052
LDH excluded (6-variable)	0.730	0.710	0.190	0.199	0.946	-0.016	0.051

Primary model used 1,000 resamples.

Supplementary Table 9. Sensitivity analysis for zero-coded LDH: coefficient comparison across models

Term	Primary *beta*	LDH zero as missing *beta*	LDH excluded *beta*
Model intercept	-5.594	-5.388	-5.155
Anemia status: Anemia (Hb<120 g/L) vs. Normal (Hb≥120 g/L)	0.918	0.518	0.955
Alanine aminotransferase (ALT), U/L	-0.022	-0.027	-0.021
Lactate dehydrogenase (LDH), U/L	0.003	0.007	
Alkaline phosphatase (ALP), U/L	0.010	0.010	0.010
Age at admission, years	0.051	0.039	0.050
Gleason score (3 groups): 8 vs. ≤7	0.779	0.696	0.753
Gleason score (3 groups): ≥9 vs. ≤7	0.510	0.516	0.518
Total PSA (tPSA), ng/mL	0.001	0.002	0.001

Supplementary Table 10. Data-quality assessment of excluded clinically important variables

Variable	Level	Bone metastasis	Total	Rate, %
Clinical T stage	T1	6	22	27.3
Clinical T stage	T2	83	178	46.6
Clinical T stage	T3	14	46	30.4
Clinical T stage	T4	10	45	22.2
History of cardiac disease	Dominant default code	109	251	43.4
History of cardiac disease	Explicit ‘none’ entry variant A	4	37	10.8
History of cardiac disease	Explicit ‘none’ entry variant B	0	1	0.0
History of cardiac disease	Yes - coronary artery disease	0	1	0.0
History of cardiac disease	Yes - other cardiac disease	0	1	0.0

Sensitivity analysis for Gleason coding

Using the same model structure and preprocessing pipeline as the primary analysis, the collapsed ≥8 coding performed slightly worse than the three-group coding.

Supplementary Table 11. Sensitivity analysis for Gleason coding: apparent performance metrics

Model	Apparent AUC	Apparent Brier
Three-group Gleason coding	0.736	0.186
Collapsed ≥8 Gleason coding	0.734	0.187

The collapsed ≥8 coding gave a slightly lower apparent AUC (0.734 vs. 0.736) and a slightly higher Brier score (0.187 vs. 0.186), so the three-group coding was retained.

Supplementary 4

TRIPOD+AI checklist

Applicable to: Model Development and Internal Validation (bootstrap, B = 1,000).

This checklist follows the TRIPOD+AI statement [Collins et al., BMJ 2024] [Ref. 9 in the main manuscript]. The item descriptions below are paraphrased short forms of the 27 reporting items (with subitems where applicable); for the full original wording, readers are referred to the TRIPOD+AI statement. The present study developed a multivariable logistic regression model with LASSO-based variable selection; it is not a deep-learning or machine-learning model, so AI-specific subitems that do not apply are marked "N/A" with a brief explanation. "M" refers to the main manuscript and "S" to this supplementary material.

Section/ Topic	Item	Topic	Checklist item (paraphrased)	Reported on (section/table/figure)
Title and abstract				
Title and abstract	1	Title	Identify the study as developing and/ or evaluating a prediction model, the target population, and the outcome to be predicted.	Main manuscript - Title page (development and internal validation; prostate cancer patients at initial diagnosis; outcome: bone metastasis).
Title and abstract	2	Abstract	Provide a structured summary describing objective, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	Main manuscript - Abstract (Background, Methods, Results, Conclusions).
Introduction				
Introduction	3a	Background	Explain the medical context and rationale for developing or evaluating the model, including references to existing models.	Main manuscript - §1 Introduction (paragraphs 1-3).
Introduction	3b	Objectives	Specify the study objectives, including whether the study describes model development, validation, or both.	Main manuscript - §1 Introduction (final paragraph) and Abstract Objective/Methods.
Methods				
Methods	4a	Data - sources	Describe the source of data separately for development and validation data sets (e.g., registry, cohort).	Main manuscript - §2.1 Study Design and Participants; §2.2 Data Sources (HIS/LIS/PACS). Internal validation by bootstrap (no separate validation dataset).
Methods	4b	Data - dates/ eligibility	Specify the key study dates (start of enrolment, end of follow-up) and eligibility criteria for the development and validation data.	Main manuscript - §2.1 Study Design and Participants (March 2011 - November 2023; inclusion/ exclusion criteria); §3.1 and Figure 1 for screening and exclusions.
Methods	5a	Participants - setting	Describe the setting, including the number and location of centres where data were collected.	Main manuscript - §2.1 (single-center, First Affiliated Hospital of Xinjiang Medical University, Urumqi, China); §4.4 regional context.
Methods	5b	Participants - eligibility	Describe eligibility criteria for participants.	Main manuscript - §2.1 (inclusion and exclusion criteria enumerated); Figure 1 flowchart.
Methods	5c	Participants - treatments received	Give details of any treatments received, and how they were handled during model development.	Main manuscript - §2.1 (baseline data obtained before any cancer-directed therapy; patients with prior cancer-directed therapy excluded). Not applicable as predictor; outcome defined at baseline pre-treatment.
Methods	6a	Outcome - definition	Clearly define the outcome that is being predicted, including how and when it was measured.	Main manuscript - §2.2.1 Outcome Variable (any-positive rule across mpMRI, CT, ECT during index hospitalization).

Methods	6b	Outcome - blinded assessment	Report any actions to blind outcome assessment from predictor information.	Main manuscript - §2.2 Data Sources (outcome extracted from imaging reports by trained abstractors with discrepancies resolved against primary records). Radiologists' imaging reports were generated during routine clinical care, independent of model development.
Methods	7a	Predictors - definition and measurement	Clearly define all predictors, including how and when they were measured.	Main manuscript - §2.2.2 Candidate Predictor Variables (93 predictors across 7 domains; measurement timing specified - closest laboratory value to pathological confirmation); Table 3 for final 7 predictors with units.
Methods	7b	Predictors - blinded assessment	Report any actions to blind assessment of predictors for the outcome and other predictors.	Main manuscript - §2.2 (independent abstraction by two trained researchers with discrepancy resolution). Predictors were extracted from structured EMR/LIS/pathology fields generated during routine care, not re-adjudicated.
Methods	8	Sample size	Explain how the study size was arrived at.	Main manuscript - §2.4.3 (EPV 113/7 = 16.1 for predictors, 113/8 = 14.1 for parameters, both >10); §2.4.2 Stage 2 documenting why the 25-variable exploratory LASSO was rejected (EPV ≈ 3.0); §4.5 Limitations acknowledges sample-size constraint for Gleason ≥9 subgroup.
Methods	9	Missing data	Describe how missing data were handled, with details of any imputation method.	Main manuscript - §2.3 Data Preprocessing (variables with >50% missingness removed; MICE m = 5; primary analysis on first imputation, Rubin's rules pooled estimates as sensitivity analysis). Supplementary §S1.2 Missing Data Handling (mice package R v3.16; ridge penalty 0.001; quickpred mincor = 0.2, minpuc = 0.30); §S3 Sensitivity Analyses (Rubin's rules pooled estimates across m = 5 imputations).
Methods	10a	Statistical analysis - predictor handling	Describe how predictors were handled in the analyses (continuous, dichotomized, transformations).	Main manuscript - §2.3 (dichotomization where clinically indicated); §2.4.3 (restricted cubic splines with 3 knots to test non-linearity for age, ALP, LDH, ALT, tPSA; all P>0.10 so entered linearly); Supplementary §S1.3 Variable Encoding (Hb dichotomized at 120 g/L; Gleason collapsed into 3 groups).
Methods	10b	Statistical analysis - model building	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Main manuscript - §2.4.2 Variable Selection (five-stage hybrid: univariable screening P<0.10, exploratory LASSO, clinical prescreening, data-quality review, confirmatory LASSO with lambda.1se); §2.4.3 final logistic regression fitted with the rms package; §2.4.4 Model Performance (bootstrap B = 1,000 via the rms package).
Methods	10c	Statistical analysis - model performance	For validation, describe how the predictions were calculated and how model performance was assessed and, if relevant, compared.	Main manuscript - §2.4.4 Model Performance (AUC with DeLong 95% CI; calibration slope/intercept/Emax; Brier score; DCA vs. treat-all and treat-none); §2.4.5 Nomogram Construction.
Methods	10d	Statistical analysis - model updating	Describe any model updating arising from the validation.	N/A - No external validation dataset was available; bootstrap internal validation does not involve model updating. See §4.5 Limitations and §4.6 Conclusion for the recommendation that external validation and, if warranted, recalibration precede clinical use.
Methods	11	Risk groups	Provide details on how risk groups were created, if done.	Main manuscript - §3.7 Model Performance reports the Youden-index threshold (0.436) for a single binary risk classification. No multi-level risk strata were defined; the nomogram (Figure 3A) provides continuous individualised probability estimates.

Methods	12	Development vs. validation	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	N/A - Internal validation by bootstrap resampling on the same development dataset; no external validation dataset with potentially different setting, eligibility, outcome, or predictors.
Results				
Results	13a	Participants - flow	Describe the flow of participants through the study, including the number with and without the outcome and, if applicable, a summary of follow-up time.	Main manuscript - §3.1 Patient Characteristics (3,362 screened → 291 analyzed; 113 bone-metastasis-positive, 178 negative); Figure 1 flowchart.
Results	13b	Participants - characteristics	Describe the characteristics of the participants (including number of participants with and without missing data for each predictor and outcome).	Main manuscript - §3.1 and Tables 1a/1b (demographics, clinicopathological, and laboratory characteristics by outcome group). Supplementary Table 1 reports descriptive statistics and univariable analysis for all 93 candidate predictors, including per-variable missingness.
Results	13c	Participants - comparison (validation)	For validation, show a comparison with the development data of the distribution of important variables.	N/A - Internal bootstrap validation; no external validation dataset for cross-dataset comparison.
Results	14a	Model development - participants	Specify the number of participants and outcome events in each analysis.	Main manuscript - §3.1 (N = 291, 113 events). Table 3 footnote confirms N = 291 and events = 113 for the final multivariable model. Supplementary Table 1 reports N per candidate predictor.
Results	14b	Model development - unadjusted associations	If done, report the unadjusted association between each candidate predictor and outcome.	Main manuscript - §3.2 Univariable Screening; Table 2 (univariable logistic regression for predictors reaching P<0.10). Supplementary Table 1 provides unadjusted comparisons for all 93 candidate predictors.
Results	15a	Model specification - final form	Present the full prediction model to allow predictions for individuals (e.g., regression coefficients, intercept, baseline survival, selected features).	Main manuscript - Table 3 Multivariable Logistic Regression - Final 7-Variable Model (Beta, SE, OR, 95% CI, P for all 8 non-intercept parameters); Figure 3A Nomogram provides the graphical equivalent for individual-level prediction.
Results	15b	Model specification - use	Explain how to use the prediction model.	Main manuscript - §3.8 Nomogram and Clinical Utility (procedure for summing predictor points and mapping to predicted probability); Figure 3A legend describes use of the Points axis and linear-predictor-to-probability mapping.
Results	16	Model performance	Report performance measures (with CIs) for the prediction model.	Main manuscript - §3.7 Model Performance and Internal Validation (apparent AUC 0.736, 95% CI 0.674-0.799; Brier 0.186; sensitivity/specificity/PPV/NPV at Youden threshold; bias-corrected AUC 0.713; calibration slope 0.933; intercept -0.025; Emax 0.050; bias-corrected Brier 0.197). Figure 3B (DCA), Figure 2B (calibration), and Figure 2A (ROC). Supplementary Tables 4 (apparent) and 5 (bootstrap bias-corrected).
Results	17	Model updating	If done, report the results from any model updating (including the updated model equation and subsequent performance).	N/A - No model updating performed (no external validation dataset).
Discussion				
Discussion	18	Limitations	Discuss any limitations of the study (such as non-representative sample, few events per predictor, missing data).	Main manuscript - §4.5 Limitations (single-center retrospective design; internal validation only; unavailable radiomics/molecular predictors; imperfect composite imaging outcome; incompletely verified missingness mechanisms, particularly 19.6% zero-coded LDH; modest sample size for Gleason ≥9 subgroup; modeling-choice sensitivity).

Discussion	19a	Interpretation - validation	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Main manuscript - §3.7 and §4.3 Discrimination Performance (apparent AUC 0.736 vs. bias-corrected 0.713, optimism 0.024, stable calibration); §4.3 compares discrimination with previously published bone-metastasis prediction models in Supplementary Table 2.
Discussion	19b	Interpretation - overall	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	Main manuscript - §4 Discussion (§4.1 predictor interpretation in biological/clinical context; §4.2 methodological considerations; §4.3 comparison with radiomics, biomarker, and registry models; §4.4 clinical positioning and regional context).
Discussion	20	Implications	Discuss the potential clinical use of the model and implications for future research.	Main manuscript - §4.4 Clinical Positioning (first-line risk-stratification aid; DCA-supported threshold range 0.1-0.7 for imaging-workup decisions); §4.5 and §4.6 identify external multi-center validation as the required next step before clinical implementation.
Open science				
Open science	21	Funding	Give the source of funding and the role of the funders for the present study.	Main manuscript - Funding section (after Acknowledgements).
Open science	22	Conflicts of interest	Declare any conflicts of interest and financial disclosures for all authors.	Main manuscript - Conflicts of Interest section.
Open science	23	Protocol	Indicate where the study protocol can be accessed, or state that a protocol was not prepared.	Main manuscript - §2 Materials and Methods describes the analysis plan. No separate registered protocol was prepared for this retrospective EMR-based study; the five-stage variable-selection strategy was pre-specified and fully described in §2.4.2.
Open science	24	Registration	Give registration information, including register name and registration number, or state that the study was not registered.	This retrospective observational study of existing medical records was not prospectively registered. Ethics approval was obtained - see Main manuscript Institutional Review Board Statement.
Open science	25	Data sharing	Provide details of the availability of the study data.	Main manuscript - Data Availability Statement. Individual-level data are not publicly available owing to patient-privacy considerations; anonymized data are available from the corresponding author on reasonable request, subject to institutional review board approval.
Open science	26	Code sharing	Provide details of the availability of the analytical code.	Analytical code is not deposited in a public repository. Core R packages and versions used are specified in the Main manuscript §2.4 (R v4.4.2), §2.3 and Supplementary §S1.2 (mice v3.16), and §2.4.3 (rms v6.8), enabling reproduction of the analytical pipeline. Code is available from the corresponding author on reasonable request.
Patient and public involvement				
Patient and public involvement	27	Patient involvement	Provide details of patient and public involvement during the design, conduct, reporting, interpretation, or dissemination of the study, or state no involvement.	No patient or public involvement in study design, conduct, or reporting. This was a retrospective analysis of retrospectively collected, de-identified clinical data, approved by the institutional ethics committee (see Main manuscript Institutional Review Board Statement).

Abbreviations: AI, artificial intelligence; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CT, computed tomography; DCA, decision curve analysis; ECT, emission computed tomography (whole-body bone scintigraphy); EMR, electronic medical record; EPV, events per variable; HIS, hospital information system; LASSO, least absolute shrinkage and selection operator; LDH, lactate dehydrogenase; LIS, laboratory information system; MICE, multiple imputation by chained equations; mpMRI, multiparametric magnetic resonance imaging; N/A, not applicable; PACS, picture archiving and communication system; ProGRP, pro-gastrin-releasing peptide; PSA, prostate-specific antigen; ROC, receiver operating characteristic; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis. Reference: Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024; 385: e078378 (PMID 38626948).