

Development and Internal Validation of a Prediction Model for Ureteral Access Sheath Insertion Depth During Flexible Ureteroscopy Based on Patient Anthropometric and Anatomical Parameters

hengda Hu

h1912d@outlook.com

Shanghai Jinshan District Central Hospital

Weixiong Ma

Shanghai Jinshan District Central Hospital

Wenpu Chen

Shanghai Jinshan District Central Hospital

Guofeng Yu

Shanghai Jinshan District Central Hospital

Article

Keywords: ureteral access sheath, flexible ureteroscopy, insertion depth, machine learning, L1–pubic symphysis distance

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10952303/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background:Optimal placement of UAS during fURS) relies largely on surgeon experience. Insertion that is too deep may cause ureteral injury, whereas insertion that is too shallow necessitates repeated fluoroscopic adjustment. This study aimed to develop and internally validate a machine learning model that predicts UAS insertion depth using the L1 vertebra–pubic symphysis distance and readily available anthropometric parameters.

Methods

This single-center cohort study enrolled 124 consecutive patients who underwent fURS with UAS placement between April 2025 and July 2026. The L1–pubic symphysis distance was measured on preoperative plain radiographs or CT scout images. Candidate predictors included height, age, sex, side, BMI and the L1–pubic symphysis distance. The outcome was the actual UAS insertion depth (cm), measured intraoperatively from the perineal skin to the ureteropelvic junction. The dataset was divided into training (70%, $n = 86$) and test (30%, $n = 38$) sets using stratified random sampling by sex. Six regression algorithms were compared, including Ridge regression, least absolute shrinkage and selection operator (Lasso), Bayesian ridge, random forest, support vector regression, and multilayer perceptron. Hyperparameters were tuned by five-fold cross-validation. Model performance was assessed using R^2 , mean absolute error (MAE), root mean square error (RMSE), and the proportion of predictions within ± 2 cm of the actual depth. The final model was refitted on the entire cohort, and bootstrap 95% confidence intervals (CIs) were obtained for all coefficients. The study followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guideline.

Results

The mean UAS insertion depth was 33.7 ± 2.8 cm (range, 26.5–40.0 cm). Height ($r = 0.749$), male sex ($r = 0.661$), L1–pubic symphysis distance ($r = 0.563$), BMI ($r = 0.251$), and younger age ($r = -0.498$) were correlated with insertion depth. Ridge regression with regularization parameter $\alpha = 10$ achieved the best balance of accuracy and interpretability, with a test-set R^2 of 0.661, MAE of 1.24 cm, RMSE of 1.56 cm, and 78.9% of predictions falling within ± 2 cm. Five-fold cross-validation yielded an R^2 of 0.642 ± 0.067 . Leave-one-out cross-validation across all 124 patients produced an R^2 of 0.649, MAE of 1.39 cm, and 77.4% accuracy within ± 2 cm, substantially outperforming an L1-distance-only model ($R^2 = 0.291$; MAE = 1.86 cm). Bootstrap internal validation (1,000 resamples) confirmed minimal overfitting (optimism = 0.026), with an optimism-corrected R^2 of 0.659, MAE of 1.39 cm, and a calibration slope of 1.033. Standardized coefficients identified male sex ($\beta = 0.927$; 95% CI, 0.651–1.205), height ($\beta = 0.811$; 95% CI, 0.444–1.148), and L1–pubic symphysis distance ($\beta = 0.740$; 95% CI, 0.396–1.086) as the strongest predictors, with BMI also contributing independently ($\beta = 0.410$; 95% CI, 0.164–0.674). A freely accessible online calculator was developed for clinical use.

Conclusion

A multivariable Ridge regression model incorporating the L1–pubic symphysis distance and anthropometric parameters can predict UAS insertion depth with clinically acceptable accuracy. The novel L1–pubic symphysis radiographic measurement, obtainable from routine preoperative imaging without additional examination, contributes independently to depth prediction. External validation in larger and more diverse cohorts is warranted before broad clinical implementation.

Introduction

Flexible ureteroscopy (fURS) has become a first-line treatment for renal and upper ureteral stones, and the ureteral access sheath (UAS) is an essential component of the procedure in most contemporary practices. The UAS facilitates repeated passage of the flexible ureteroscope, improves irrigation and visibility, reduces intrarenal pressure, and shortens operative time[1, 2]. Despite these advantages, UAS placement is not without risk. Traxer[3] reported ureteral wall injury in 46.5% of cases after UAS insertion, with severe injury involving the smooth muscle layer occurring in 13.3%. Male sex and older age were identified as risk factors for severe injury. Subsequent randomized and observational studies have confirmed that ureteral injury remains a clinically significant complication, particularly when the sheath is advanced against resistance or when the ureter is narrow[4–6].

The depth to which the UAS is inserted is a critical determinant of both safety and efficacy. When the sheath is advanced too deeply, its tip may go over the ureteropelvic junction (UPJ) or herniate into a renal calyx, potentially causing perforation, false passage, or mucosal avulsion. When the sheath is too shallow, its tip fails to reach the upper ureter, and the surgeon must repeatedly withdraw and reinsert the scope under fluoroscopic guidance to maintain access, increasing radiation exposure and operative time. At present, the appropriate insertion depth is estimated subjectively based on surgeon experience, body habitus, and intraoperative fluoroscopy. There is no validated preoperative tool to guide this decision.

The recent introduction of the flexible and navigable suction ureteral access sheath (FANS) has further heightened the importance of accurate depth selection. FANS features a flexible, steerable distal tip with integrated negative-pressure suction that must be positioned within or near the target calyx to effectively aspirate stone dust and debris, reduce intrarenal pressure, and maintain clear visibility, which caused an urgency for the accuracy of the insertion depth[7, 8].

Several investigators have explored methods for predicting ureteral length or ureteral stent length using anthropometric and imaging variables[9–11]. Hruby et al. [9] demonstrated that ureteric length correlates with height and other surface body habitus variables. Kawahara et al.[10] developed a preoperative predictive model for ureteral length based on CT measurements, and Hsu et al. [11] showed that natural ureteral length can be estimated from CT or plain radiography.

We hypothesized that the distance from the inferior endplate of the L1 vertebra to the superior margin of the pubic symphysis, measured on a preoperative plain radiograph or CT scout image, approximates the craniocaudal span of the retroperitoneal ureter and could serve as a novel predictor of UAS insertion

depth. This measurement is readily available from imaging already obtained for surgical planning and requires no additional radiation or cost. When combined with height, age, sex, side, and body mass index (BMI), it may enable accurate preoperative estimation of the optimal sheath depth.

Patients and Methods

Study Design and Population

This single-center retrospective cohort study was conducted at our hospital in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement[12]. The study protocol was approved by the Institutional Review Board of our hospital (approval number:LXKY202511).

Consecutive adult patients who underwent fURS with UAS placement for renal or proximal ureteral stones between April 2025 and July 2026 were screened for eligibility. Patients were included if they were aged 18 years or older, had a preoperative plain radiograph (kidney–ureter–bladder film) or CT scout image available for measurement, and had complete records of intraoperative UAS insertion depth. Exclusion criteria were ureteral stricture or congenital ureteral anomaly, history of ipsilateral ureteral surgery or ureteral reimplantation, spinal deformity (scoliosis, kyphosis, or prior spinal fusion), ectopic kidney, horseshoe kidney, pelvic kidney, and preoperative indwelling ureteral stent, as these conditions could alter ureteral anatomy and confound the depth measurement.

All procedures were performed with the patient under general anesthesia in the lithotomy position. A (11/13 or 12/14 Fr Cook or Boston scientific) UAS was used in all cases. The operative /technique was standardized: a guidewire was placed into the renal pelvis under fluoroscopic guidance, and the UAS was advanced over a second working wire under combined endoscopic and fluoroscopic visualization until the tip was positioned just distal to the UPJ, as confirmed by direct visualization through the sheath and fluoroscopy.

Predictor Variables

Six candidate predictors were selected a priori based on anatomical plausibility and availability in routine clinical practice: height (cm), age (years), sex (female = 0, male = 1), side (right = 0, left = 1), L1–pubic symphysis distance (mm), and BMI (kg/m²).

The L1–pubic symphysis distance was defined as the craniocaudal distance from the inferior endplate of the L1 vertebral body to the superior margin of the pubic symphysis, measured in millimeters on the preoperative anteroposterior plain radiograph (or CT scout image) using the hospital picture archiving and communication system digital caliper tool(Fig. 1). Two independent observers performed the measurements on all images, and the mean of the two measurements was used for analysis. If the discrepancy between the two observers exceeded 5 mm, a third independent observer performed a

repeat measurement, and the outlying value was excluded; the final value was taken as the mean of the two closest measurements.

Outcome Measurement

The primary outcome was the actual UAS insertion depth, defined as the distance in centimeters from the perineal skin to UPJ along the instrumented urinary tract. In female patients, the depth was read directly from the centimeter markings on the UAS at the urethral meatus after final positioning. In male patients, the depth was calculated as the sheath insertion depth at the external urethral meatus minus the measured distance from the glans penis to the root of the penis, thereby normalizing the measurement to the perineal skin and eliminating the variable contribution of penile length. All depth measurements were independently recorded by two observers intraoperatively; the mean of the two readings was used for analysis. If the discrepancy between the two observers exceeded 5 mm, a third observer performed a repeat measurement and the outlying value was excluded, with the final value taken as the mean of the two closest readings.

Data Splitting and Preprocessing

The dataset was randomly divided into a training set (70%, $n = 86$) and a held-out test set (30%, $n = 38$) using stratified sampling by sex to preserve the sex distribution in both sets. The training set was used for model development, hyperparameter tuning, and cross-validation; the test set was used exclusively for final performance assessment and was not involved in any model selection decision.

Continuous variables were standardized (z-score transformation) within each cross-validation training fold to prevent data leakage. No imputation was required because the final dataset contained no missing values.

Model Development

Six regression algorithms were evaluated to capture both linear and nonlinear relationships between predictors and UAS insertion depth: Ridge regression (L2 regularization), Lasso regression (L1 regularization), Bayesian ridge regression, random forest, support vector regression with a radial basis function kernel (SVR), and a multilayer perceptron (MLP) neural network. Ridge and Lasso regularization strengths (α) were selected from $\{0.1, 1, 5, 10, 20, 50\}$; the random forest was tuned for maximum depth (2–6) and number of estimators (100–500); SVR was tuned over the cost parameter C (1, 10, 100) and kernel width γ ; the MLP architecture comprised two hidden layers (16 and 8 neurons) with ReLU activation and L2 penalty.

Hyperparameter tuning was performed by five-fold cross-validation on the training set, using MAE as the selection criterion. The best-performing model was then evaluated on the held-out test set. To provide a robust estimate of overall performance, leave-one-out cross-validation (LOOCV) was additionally conducted across the entire cohort of 124 patients. To quantify the incremental value of the multivariable approach over the novel radiographic measurement alone, a univariate linear regression

model using only the L1–pubic symphysis distance was also evaluated under the same cross-validation framework.

Model Evaluation

Model discrimination was assessed using the coefficient of determination (R^2), MAE, and RMSE. Clinical accuracy was defined as the proportion of predictions within ± 2 cm of the actual depth. This threshold was selected a priori based on the anatomical observation that the ureteropelvic junction (UPJ) typically lies approximately 2 cm proximal to the renal calyceal mucosa; consequently, a placement error within 2 cm is unlikely to cause the sheath tip to impinge on or injure the calyceal mucosa, and generally does not require intraoperative repositioning. The proportion within ± 1 cm was also reported as a stricter metric. Calibration was examined by plotting observed against predicted values in quintiles and calculating the calibration slope. Residuals (actual minus predicted) were inspected for systematic bias, and patients with absolute residuals exceeding 2 cm in the test set were individually characterized to identify patterns of prediction failure.

The final model was refitted on all 124 patients using the selected algorithm and hyperparameters. Nonparametric bootstrap resampling (1,000 iterations) was used to estimate 95% CIs for the regression coefficients. To quantify model overfitting and obtain optimism-corrected performance estimates, we additionally performed bootstrap internal validation following the method described by Steyerberg et al. In each of 1,000 bootstrap resamples, the model was refitted and its performance was evaluated both on the bootstrap sample (apparent performance) and on the original full dataset (test performance); the mean difference between apparent and test performance across bootstraps was defined as the optimism. The optimism-corrected performance was calculated as the apparent performance minus the mean optimism. Calibration was assessed by the calibration slope (ideal value = 1) and intercept (ideal value = 0) from regressing observed depth on predicted depth. The final prediction equation was deployed as a web-based calculator using the original (unstandardized) coefficient scale.

Statistical Analysis

Statistical analysis was performed using Python (version 3.8) with the scikit-learn library (version 1.3). Continuous variables are summarized as mean $\pm$ standard deviation or median with interquartile range as appropriate; categorical variables are presented as counts and percentages. Between-group comparisons used the independent t test or Mann–Whitney U test for continuous variables and the chi-square or Fisher exact test for categorical variables. Pearson correlation coefficients were calculated to assess univariate associations between each predictor and the outcome. A two-sided p value $< .05$ was considered statistically significant.

Results

Baseline Characteristics

A total of 124 patients met the inclusion criteria and were included in the analysis. Table 1 summarizes the baseline characteristics of the overall cohort and the training and test sets. The mean age was 54.5 ± 13.9 years, and 90 patients (72.6%) were male. The mean height and BMI were 166.7 ± 7.1 cm and 24.7 ± 3.0 kg/m², respectively. The mean L1–pubic symphysis distance was 296.8 ± 20.4 mm (range, 237.9–350.4 mm). The mean UAS insertion depth was 33.7 ± 2.8 cm (range, 26.5–40.0 cm). Left-sided procedures accounted for 79 cases (63.7%). The training and test sets were well balanced with respect to all baseline characteristics, confirming adequate stratified randomization.

Table 1
Baseline Characteristics of the Study Population

Characteristic	Overall (*n* = 124)	Training (*n* = 86)	Test (*n* = 38)
Age (years), mean $\pm$ SD	54.5 ± 13.9	53.8 ± 14.3	56.0 ± 13.1
Sex, *n* (%)			
Female	34 (27.4)	24 (27.9)	10 (26.3)
Male	90 (72.6)	62 (72.1)	28 (73.7)
Height (cm), mean $\pm$ SD	166.7 ± 7.1	166.7 ± 7.2	166.6 ± 7.1
BMI (kg/m ²), mean $\pm$ SD	24.7 ± 3.0	24.9 ± 3.1	24.3 ± 2.6
Side, *n* (%)			
Left	79 (63.7)	55 (64.0)	24 (63.2)
Right	45 (36.3)	31 (36.0)	14 (36.8)
L1–pubic symphysis (mm), mean $\pm$ SD	296.8 ± 20.4	298.2 ± 19.4	293.7 ± 22.7
UAS insertion depth (cm), mean $\pm$ SD	33.7 ± 2.8	33.7 ± 2.8	33.6 ± 2.7

Note. BMI = body mass index; L1–pubic symphysis = distance from the inferior endplate of L1 to the superior margin of the pubic symphysis; SD = standard deviation; UAS = ureteral access sheath.

Univariate Correlations

All continuous predictors except side were significantly correlated with UAS insertion depth. Height showed the strongest correlation ($r = 0.749$, $p < .001$), followed by male sex ($r = 0.661$, $p < .001$), L1–pubic symphysis distance ($r = 0.563$, $p < .001$), and age ($r = -0.498$, $p < .001$). BMI was moderately correlated ($r = 0.251$, $p = .005$), and side showed no significant univariate association ($r = 0.030$, $p = .74$). UAS depth decreased progressively with age, from a mean of 35.6 ± 1.7 cm in patients aged 40 years or younger to 31.1 ± 2.7 cm in those older than 70 years.

Model Performance Comparison

Table 2 presents the performance of all evaluated algorithms on the training and test sets. Ridge regression with $\alpha = 10$ achieved the most favorable balance of discrimination, calibration, and simplicity. On the held-out test set, it produced an R^2 of 0.661, MAE of 1.24 cm, RMSE of 1.55 cm, with 78.9% of predictions within ± 2 cm and 52.6% within ± 1 cm. Five-fold cross-validation on the training set yielded an R^2 of 0.641 ± 0.066 and MAE of 1.38 ± 0.13 cm, indicating minimal overfitting.

Table 2
Performance of Multiple Machine Learning Models on the Training and Test Sets

Model	Train *R ²	Train MAE	Test *R ²	Test MAE	Test RMSE	Within ± 2 cm (%)	Within ± 1 cm (%)
Ridge ($\alpha = 10$)	0.687	1.35	0.661	1.24	1.56	78.9	52.6
Lasso ($\alpha = 0.01$)	0.690	1.33	0.656	1.22	1.57	78.9	52.6
Bayesian ridge	0.689	1.34	0.660	1.23	1.56	78.9	52.6
Random forest (max depth = 3)	0.797	1.03	0.652	1.24	1.58	78.9	50.0
SVR (RBF kernel)	0.846	0.64	0.471	1.48	1.94	68.4	42.1
MLP (16, 8)	0.755	1.11	0.028	2.07	2.64	57.9	34.2
<i>Note. MAE = mean absolute error; RMSE = root mean square error; Lasso = least absolute shrinkage and selection operator; SVR = support vector regression; RBF = radial basis function; MLP = multilayer perceptron. All metrics are reported on the original centimeter scale.</i>							

The random forest achieved a test-set R^2 comparable to Ridge (0.652) but exhibited a pronounced gap between training and test performance (training $R^2 = 0.797$), indicative of mild overfitting despite depth restriction. SVR and the MLP showed clear overfitting, with training R^2 exceeding 0.75 but test R^2 dropping below 0.48; the MLP was particularly unstable, with a near-zero test R^2 of 0.028. Ridge, Lasso, and Bayesian ridge regression provided the most consistent performance across training, cross-validation, and test evaluations, and Ridge was selected as the final model because of its explicit regularization path, straightforward coefficient interpretation, and direct translation into a bedside prediction formula.

Final Model and Coefficients

The final Ridge model ($\alpha = 10$) was refitted on all 124 patients. Table 3 presents the standardized regression coefficients with bootstrap 95% CIs, alongside the original-scale coefficients used in the prediction formula. Male sex was the strongest standardized predictor ($\beta = 0.927$; 95% CI, 0.651–1.205), followed by height ($\beta = 0.811$; 95% CI, 0.444–1.148), the L1–pubic symphysis distance ($\beta = 0.740$; 95% CI, 0.396–1.086), and BMI ($\beta = 0.410$; 95% CI, 0.164–0.674). BMI emerged as a significant independent

positive predictor, with its 95% CI excluding zero. The 95% CIs for age ($\beta = -0.157$; 95% CI, -0.481 to 0.153) and side ($\beta = 0.136$; 95% CI, -0.131 to 0.380) included zero, indicating that these variables did not contribute independently at the 0.05 significance level after adjustment for the remaining predictors.

Table 3
Final Ridge Regression Model Coefficients ($\alpha = 10$)

Predictor	Standardized β	95% CI	Original-scale coefficient
Intercept	—	—	-0.5895
Height (cm)	0.811	0.444 to 1.148	0.1143
Age (years)	-0.157	-0.481 to 0.153	-0.0114
Sex (male = 1)	0.927	0.651 to 1.205	2.0772
Side (left = 1)	0.136	-0.131 to 0.380	0.2824
L1–pubic symphysis (mm)	0.740	0.396 to 1.086	0.0363
BMI (kg/m ²)	0.410	0.164 to 0.674	0.1376

Note. BMI = body mass index; CI = confidence interval. Bootstrap 95% CIs based on 1,000 resamples.

Internal Validation and Incremental Value

LOOCV across all 124 patients yielded an R^2 of 0.649, MAE of 1.39 cm, RMSE of 1.63 cm, and 77.4% of predictions within ± 2 cm, closely matching the held-out test results and confirming the stability of the model. In contrast, a univariate model using only the L1–pubic symphysis distance achieved a LOOCV R^2 of 0.291, MAE of 1.86 cm, and 62.1% accuracy within ± 2 cm, demonstrating that the multivariable model substantially improves prediction over the radiographic measurement alone. The calibration slope was close to unity, and residuals were approximately normally distributed around zero without evidence of systematic over- or under-prediction across the range of observed depths (Fig. 2).

Bootstrap Internal Validation

Bootstrap internal validation (1,000 resamples) was performed to quantify overfitting and obtain optimism-corrected performance estimates for the final Ridge model refitted on all 124 patients (Table 4). The apparent R^2 was 0.685, with a mean optimism of 0.026, yielding an optimism-corrected R^2 of 0.659. The optimism-corrected MAE was 1.39 cm and RMSE was 1.63 cm, with 76.7% of predictions within ± 2 cm. The calibration slope was 1.033 (ideal = 1) and the calibration intercept was -1.11 (ideal = 0), indicating good calibration with minimal systematic bias. The small magnitude of the optimism ($\Delta R^2 = 0.026$) confirms that the L2 regularization in Ridge regression effectively controlled overfitting, and the optimism-corrected estimates are closely consistent with both the held-out test set ($R^2 = 0.661$) and LOOCV ($R^2 = 0.649$), providing convergent evidence of model stability.

Table 4
Bootstrap Internal Validation of the Final Ridge Model (n = 124, 1,000 Resamples)

Performance Metric	Apparent	Optimism	Optimism-Corrected
R ²	0.685	0.026	0.659
MAE (cm)	1.319	-0.067	1.386
RMSE (cm)	1.547	-0.083	1.630
Within ± 2 cm (%)	79.8	3.18	76.7
Calibration slope	1.046	0.013	1.033
Calibration intercept	-1.563	-0.451	-1.112

Note. MAE = mean absolute error; RMSE = root mean square error. Apparent performance was obtained by fitting and evaluating the model on the full dataset. Optimism was calculated as the mean difference between bootstrap-sample apparent performance and original-dataset test performance across 1,000 bootstrap resamples. Optimism-corrected performance = apparent - mean optimism.

Nested Model Comparison

To quantify the incremental value of the L1–pubic symphysis distance beyond readily available anthropometric parameters, we constructed two nested Ridge regression models and compared them using leave-one-out cross-validation (Table 5). Model 1 included only conventional anthropometric variables—height, BMI, age, sex, and side—and achieved a LOOCV R² of 0.606 (RMSE = 1.73 cm; MAE = 1.46 cm; 73.4% within ± 2 cm). Addition of the L1–pubic symphysis distance in Model 2 increased the LOOCV R² to 0.649, corresponding to a ΔR^2 of + 0.043, a reduction in RMSE from 1.73 to 1.63 cm, and an improvement in the proportion of predictions within ± 2 cm from 73.4% to 77.4%. The incremental contribution of the L1–pubic symphysis distance was statistically significant by the nested-model F-test ($F(1, 117) = 14.19, p < .001$). These results demonstrate that the L1–pubic symphysis distance provides statistically significant and clinically meaningful incremental predictive value beyond conventional anthropometric parameters alone.

Table 5
Nested Model Comparison Showing Incremental Value of the L1–Pubic Symphysis Distance

Model	Predictors	LOO R ²	Adjusted R ²	RMSE (cm)	MAE (cm)	Within ± 2 cm (%)	ΔR^2
Model 1	Height, BMI, age, sex, side	0.606	0.590	1.73	1.46	73.4	—
Model 2	Model 1 + L1–pubic symphysis	0.649	0.631	1.63	1.39	77.4	+ 0.043

Note. LOO = leave-one-out cross-validation; RMSE = root mean square error; MAE = mean absolute error. Both models used Ridge regression with $\alpha = 10$. The nested-model F-test for Model 1 versus Model 2 was $F(1, 117) = 14.19, p < .001$.

L1–BMI Interaction Analysis

We hypothesized that the L1–pubic symphysis distance (a longitudinal bony measurement) and BMI (a measure of body habitus and soft-tissue thickness) may jointly determine UAS insertion depth in a manner analogous to the two legs of a right triangle, with the instrumented urinary tract path corresponding to the hypotenuse. To test this hypothesis, we examined the L1×BMI product term and its incremental value. In univariate analysis, the L1×BMI interaction term correlated with depth at $r = 0.465$, which was substantially higher than BMI alone ($r = 0.251$) but lower than L1 alone ($r = 0.563$). In reduced models containing only L1 and/or BMI, the addition of the L1×BMI product term did not improve LOOCV performance: L1 + BMI achieved $R^2 = 0.310$, whereas L1 + BMI+L1×BMI also achieved $R^2 = 0.310$. In the full six-predictor model, inclusion of the L1×BMI term likewise yielded no meaningful improvement (LOOCV $R^2 = 0.648$ versus 0.649 without the interaction; standardized β for the interaction term = 0.324). These findings suggest that although L1 and BMI each contribute independently to depth prediction, their joint effect is approximately additive rather than synergistic, and the right-triangle analogy—while conceptually appealing—does not capture additional variance beyond what the two main effects already provide.

Online Calculator

The final Ridge model was implemented as a freely accessible from **the supplement file**. The interface requires input of height, age, sex, side, BMI, and L1–pubic symphysis distance; the predicted depth.

Discussion

In this study, we developed and internally validated a machine learning model for predicting UAS insertion depth during fURS based on six preoperative variables. The final Ridge regression model achieved a test-set R^2 of 0.661 and correctly predicted 78.9% of cases within ± 2 cm, with LOOCV confirming stable performance across the entire cohort. Male sex, height, and the novel L1–pubic symphysis distance were the three strongest predictors, and the multivariable model more than doubled the variance explained compared with the L1 distance alone. To our knowledge, this is the first study to formally model UAS insertion depth and to propose the L1–pubic symphysis distance as a preoperative radiographic predictor.

The anatomical rationale for the L1–pubic symphysis distance is straightforward. The ureter descends from the UPJ, which lies near the L1–L2 level, along the anterior surface of the psoas muscle, crosses the pelvic brim, and enters the bladder at the ureteral orifice near the pubic symphysis. The craniocaudal distance between these two bony landmarks therefore approximates the vertical span of the ureter's retroperitoneal course. Although the ureter is not a straight vertical structure and its length exceeds this linear distance because of its medial and anterior curves, the L1–pubic symphysis distance captures a

major component of the anatomical variation that determines how far a sheath must travel to reach the UPJ. Its measurement requires only a plain radiograph, both of which are routinely obtained before fURS for stone localization and surgical planning, so it introduces no additional cost or radiation.

The predominance of male sex as the strongest predictor ($\beta = 0.927$) is consistent with urethral anatomy. Even after normalizing the male measurement to the perineal skin by subtracting penile length, adjustment for height and body habitus revealed an independent sex effect of approximately 2.0 cm—likely attributable to structural differences in the perineal region, as the male urethra courses anteriorly.

Height emerged as the second-strongest predictor ($\beta = 0.811$), corroborating previous work on ureteral length. It was found that ureteric length correlated with anthropometric variables of body habitus [9, 13, 14], and Kawahara et al. [10] incorporated height into their ureteral length prediction formula. The L1–pubic symphysis distance ($\beta = 0.740$) and height likely capture complementary anatomical information: height reflects overall body scale, whereas the L1–pubic symphysis distance reflects the individual's truncal proportions, which vary independently of stature. BMI emerged as a significant independent positive predictor ($\beta = 0.410$), suggesting that adiposity and soft-tissue thickness contribute to the effective insertion depth beyond what is captured by height alone. The negative coefficient for age is consistent with age-related changes in body composition and ureteral course; older patients in our cohort had progressively shallower insertion depths, possibly reflecting reduced truncal height from vertebral compression or changes in pelvic organ position.

The nonsignificant independent contribution of side is expected, because the left and right ureters are of similar length despite their slightly different anatomical relationships. We additionally explored a potential L1×BMI interaction, motivated by the conceptual model in which the longitudinal bony distance (L1–pubic symphysis) and the transverse soft-tissue dimension (reflected by BMI) form the two legs of a right triangle, with the instrumented urinary tract path corresponding to the hypotenuse. Although the L1×BMI product term correlated with depth at $r = 0.465$ —higher than BMI alone ($r = 0.251$)—its inclusion in either reduced or full models yielded no incremental improvement in LOOCV R^2 , and its standardized coefficient in the full model was modest ($\beta = 0.324$). These findings indicate that the effects of L1 distance and body habitus are approximately additive rather than synergistic, and that the right-triangle analogy, while heuristically useful, does not capture additional variance beyond the two main effects.

The nested model comparison provides the clearest evidence for the clinical value of the L1–pubic symphysis distance. A model containing only conventional anthropometric parameters—height, BMI, age, sex, and side—explained 60.6% of the variance in UAS insertion depth (LOOCV $R^2 = 0.606$). Addition of the single L1–pubic symphysis measurement increased the explained variance to 64.9% ($\Delta R^2 = +0.043$), a statistically significant improvement ($F(1, 117) = 14.19$, $p < .001$) that also reduced RMSE by 0.10 cm and increased the proportion of predictions within ± 2 cm by 4.0 percentage points. This pattern underscores that the L1–pubic symphysis distance is not merely a proxy for body size but captures independent anatomical variation in truncal proportions that is not reflected in standard anthropometric measurements. Given that this measurement can be obtained from routinely acquired preoperative

imaging at no additional cost or radiation, its incremental predictive value justifies its inclusion in preoperative planning for UAS placement.

From a machine learning perspective, the relatively modest sample size (124 patients) favored regularized linear models over flexible nonlinear algorithms. The random forest achieved a test-set R^2 comparable to Ridge (0.652) but showed clear overfitting, with a training R^2 of 0.797. SVR and the MLP overfit more severely, with the MLP being particularly unstable (test R^2 = 0.028). Ridge regression, by imposing an L2 penalty that shrinks coefficients toward zero, produced the most stable and interpretable model, a finding consistent with the broader prediction modeling literature showing that regularized linear models often match or outperform complex algorithms in small clinical datasets [15-17]. The choice of Ridge also yields a transparent prediction formula that can be implemented without specialized software, facilitating clinical adoption.

The clinical implications are practical. The model achieved a mean absolute error of approximately 1.4 cm (optimism-corrected MAE = 1.39 cm; test-set MAE = 1.24 cm), which is well below the a priori defined 2 cm threshold for clinically acceptable prediction. In practical terms, an average error of 1.4 cm means that for most patients, the predicted depth will fall within a range that requires at most one minor fluoroscopic adjustment, rather than repeated repositioning. This level of accuracy is particularly meaningful given that the total UAS insertion depth ranges from 26.5 to 40.0 cm (mean 33.7 cm), so a 1.4 cm error corresponds to approximately 4% of the mean depth. Had the MAE been in the range of 3–4 cm, the clinical utility would be substantially more limited, and the model would warrant characterization as having statistical predictive ability but uncertain clinical value. At the observed accuracy level, a preoperative estimate within ± 2 cm in nearly 80% of cases provides a useful starting point for sheath placement, potentially reducing the number of fluoroscopic adjustments needed to achieve optimal tip position. This is particularly relevant for trainees and surgeons early in their learning curve, for whom depth estimation is most uncertain. The online calculator makes the model accessible on any device with a web browser. We do not propose that the model replace intraoperative judgment or fluoroscopic confirmation; rather, it serves as a planning aid that narrows the initial insertion range, after which fine adjustment proceeds under direct visualization as usual.

This study has several limitations. First, the sample size is modest, and the model has undergone only internal validation through split-sample and leave-one-out procedures. External validation in independent cohorts, ideally from multiple centers with diverse patient populations and different sheath brands, is essential before the model can be recommended for broad clinical use. Second, the cohort was predominantly male (72.6%), reflecting the demographic of patients undergoing fURS at our institution; the model may be less well calibrated in female-predominant populations, and the sex-specific coefficient should be interpreted with this imbalance in mind. Third, we did not record the surgeon's preoperative estimate or the number of fluoroscopic adjustments, so we could not directly compare the model's accuracy with current clinical practice or quantify the reduction in radiation exposure; this should be addressed in a prospective impact study. Finally, the model does not incorporate ureteral diameter,

tortuosity, or prior stone passage, which may influence both the achievable depth and the risk of injury; these variables could be added in future iterations as larger datasets become available.

Conclusion

A multivariable Ridge regression model incorporating the L1–pubic symphysis distance and anthropometric parameters can predict UAS insertion depth with clinically acceptable accuracy. The novel L1–pubic symphysis radiographic measurement, obtainable from routine preoperative imaging without additional examination, contributes independently to depth prediction.

Declarations

Ethics approval and consent to participate: This study was approved by the Institutional Review Board of Shanghai Jinshan District Central Hospital (Approval Number:LXKY202511). Due to the retrospective nature of the study, the requirement for informed consent was waived by the ethics committee.

Consent for publication: Not applicable.

Availability of data and materials: The datasets and codes used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Authors' contributions: Hengda Hu (first and corresponding author): Conceptualization, Methodology, Data curation and supervision, Formal analysis (machine learning modeling), Software (online calculator development), Writing – original draft, Writing – review & editing, Visualization. Weixiong Ma: Data curation, Investigation, Writing – original draft. Wenpu Chen: Data curation, Formal analysis, Writing – review & editing. Guofeng Yu (co-corresponding author): Supervision, Writing – review & editing, Final approval of the manuscript.

All authors have read and agreed to the published version of the manuscript.

Funding: The authors declare that this research was funded by Health Commission of Jinshan District, Shanghai City(number:JSKJ-KTQN-2024-09)

References

1. De S, Sarkissian C, Torricelli FC, Brown R, Monga M: **New ureteral access sheaths: a double standard.** *Urology* 2015, **85**(4):757-763.
2. Breda A, Emiliani E, Millán F, Scoffone CM, Knoll T, Osther PJ, Liatsikos E: **The new concept of ureteral access sheath with guidewire disengagement: One wire does it all.** *World journal of urology* 2016, **34**(4):603-606.

3. Traxer O, Thomas A: **Prospective evaluation and classification of ureteral wall injuries resulting from insertion of a ureteral access sheath during retrograde intrarenal surgery.** *The Journal of urology* 2013, **189**(2):580-584.
4. Loftus CJ, Ganesan V, Traxer O, Schold JD, Noble M, Sivalingam S, Muruve N, Monga M: **Ureteral Wall Injury with Ureteral Access Sheaths: A Randomized Prospective Trial.** *Journal of endourology* 2020, **34**(9):932-936.
5. Fulla J, Prasanchaimontri P, Rizk A, Loftus C, Remer EM, Monga M: **Ureteral Diameter as Predictor of Ureteral Injury during Ureteral Access Sheath Placement.** *The Journal of urology* 2021, **205**(1):159-164.
6. Tapiero S, Kaler KS, Jiang P, Lu S, Cottone C, Patel RM, Okhunov Z, Klopfer MJ, Landman J, Clayman RV: **Determining the Safety Threshold for the Passage of a Ureteral Access Sheath in Clinical Practice Using a Purpose-Built Force Sensor.** *The Journal of urology* 2021, **206**(2):364-372.
7. Wang L, Luo Z, Huang W, Jia Q: **Efficacy and safety of the flexible negative-pressure ureteral sheath in retrograde intrarenal surgery: a systematic review and meta-analysis.** *Frontiers in surgery* 2025, **12**:1649574.
8. Gauhar V, Traxer O, Castellani D, Sietz C, Chew BH, Fong KY, Hamri SB, Gökce MI, Gadzhiev N, Galosi AB et al: **Could Use of a Flexible and Navigable Suction Ureteral Access Sheath Be a Potential Game-changer in Retrograde Intrarenal Surgery? Outcomes at 30 Days from a Large, Prospective, Multicenter, Real-world Study by the European Association of Urology Urolithiasis Section.** *European urology focus* 2024, **10**(6):975-982.
9. Hruby GW, Ames CD, Yan Y, Monga M, Landman J: **Correlation of ureteric length with anthropometric variables of surface body habitus.** *BJU international* 2007, **99**(5):1119-1122.
10. Kawahara T, Sakamaki K, Ito H, Kuroda S, Terao H, Makiyama K, Uemura H, Yao M, Miyamoto H, Matsuzaki J: **Developing a preoperative predictive model for ureteral length for ureteral stent insertion.** *BMC urology* 2016, **16**(1):70.
11. Hsu JT, Tseng JS, Chen M, Sun FJ, Chen CW, Lin WR, Chiang PK, Chiu AW: **Preoperative estimate of natural ureteral length based on computed tomography and/or plain radiography.** *Sci Rep* 2021, **11**(1):12202.
12. Collins GS, Reitsma JB, Altman DG, Moons KG: **Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement.** *Annals of internal medicine* 2015, **162**(1):55-63.
13. Paick SH, Park HK, Byun SS, Oh SJ, Kim HH: **Direct ureteric length measurement from intravenous pyelography: does height represent ureteric length?** *Urological research* 2005, **33**(3):199-202.
14. Pilcher JM, Patel U: **Choosing the correct length of ureteric stent: a formula based on the patient's height compared with direct ureteric measurement.** *Clinical radiology* 2002, **57**(1):59-62.
15. Hui Y, Hu H, Xiang J, Du X: **Comparative analysis of machine learning algorithms for predicting tibial intramedullary nail length from patient characteristics.** *Journal of orthopaedics and traumatology : official journal of the Italian Society of Orthopaedics and Traumatology* 2025, **26**(1):56.

16. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B: **A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models.** *Journal of clinical epidemiology* 2019, **110**:12-22.
17. van der Ploeg T, Austin PC, Steyerberg EW: **Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints.** *BMC medical research methodology* 2014, **14**:137.

Figures

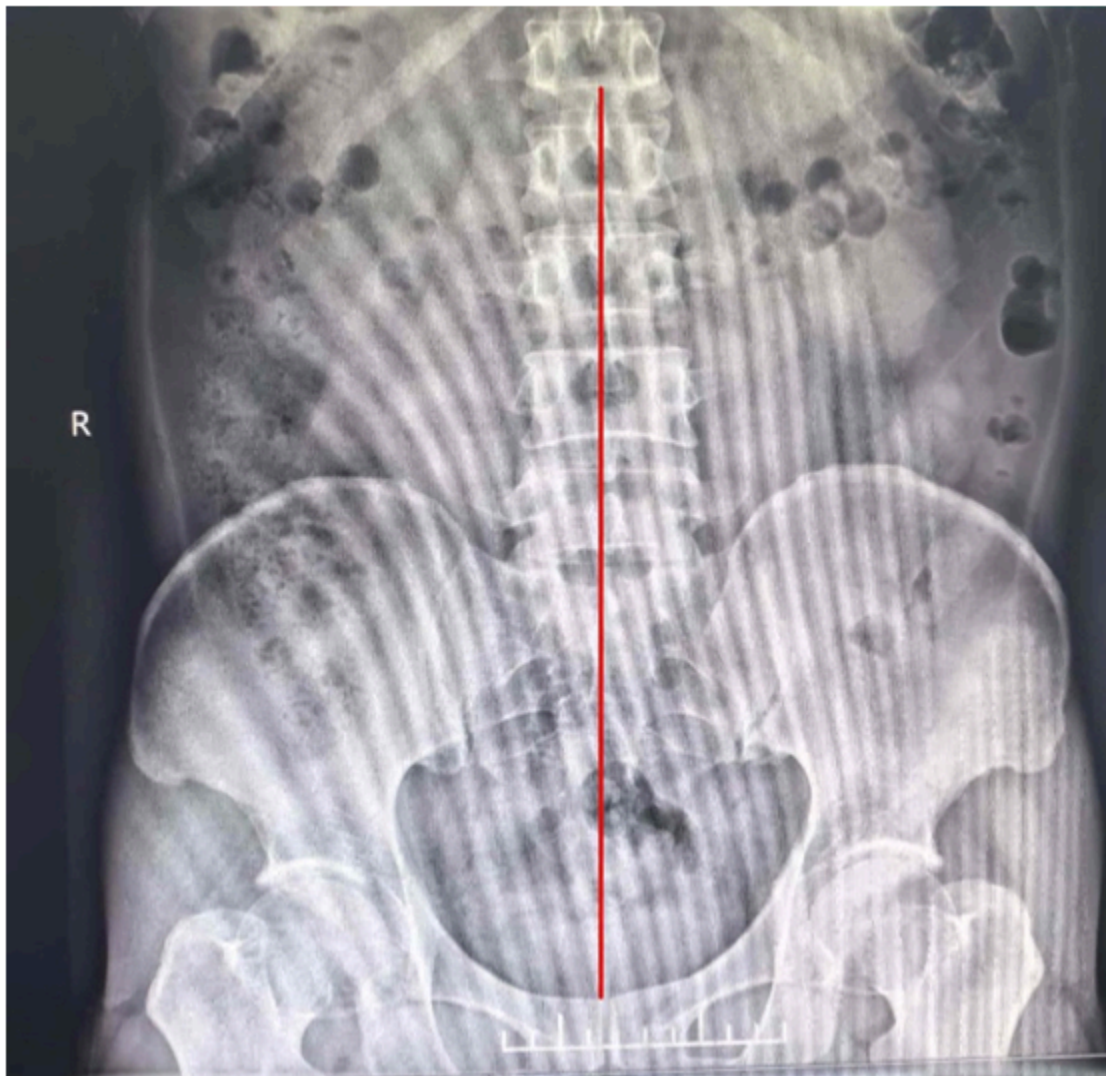


Figure 1

Measurement of the L1–pubic symphysis distance on a preoperative anteroposterior plain radiograph image. The distance is measured from the inferior endplate of the L1 vertebral body to the superior margin of the pubic symphysis

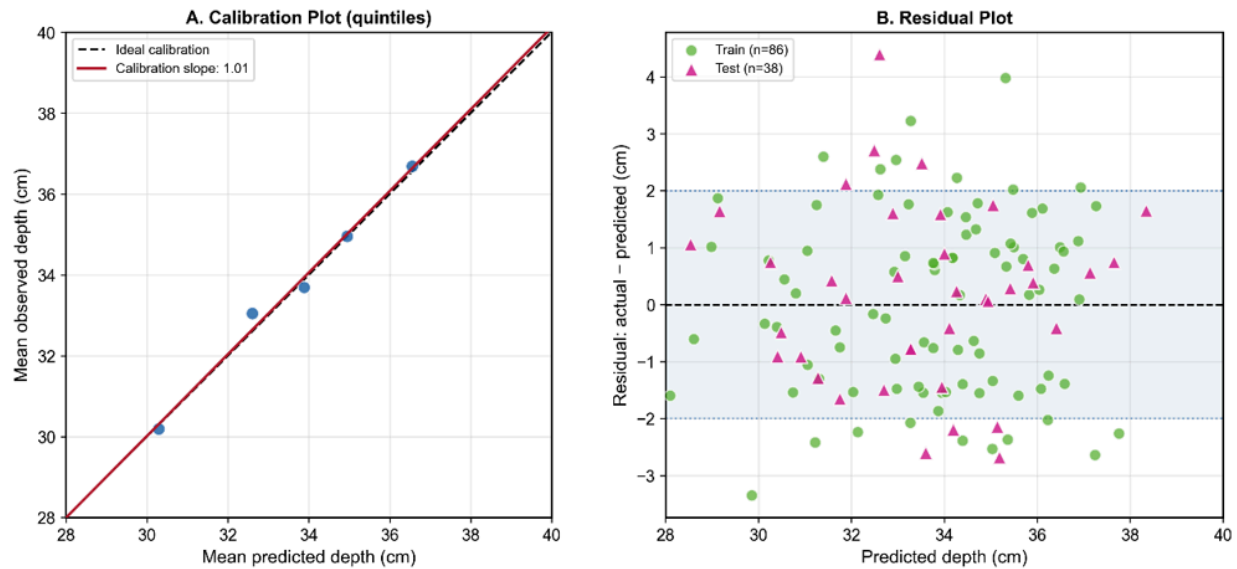


Figure 2

Calibration and residual plots. (A) Calibration plot of mean observed versus mean predicted depth by quintile; the dashed line represents ideal calibration. (B) Residual plot (actual minus predicted) against predicted values for the training (circles) and test (triangles) sets; dotted lines indicate ± 2 cm.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [UASInsertionDepthPredictor.html](#)