

MRI-HISTOPATHOLOGY CONCORDANCE AND PATTERNS OF STAGING DISCREPANCY IN BLADDER CANCER: A SINGLE-CENTER STUDY

Syed Muhammad Faiq^{*1}, Bisma Ghulam Mustafa Arijio², Munazzah Ibrar³,
Syeda Nuzhat Zehra⁴, Afifa Mehmood⁵, Aiman Akif⁶

^{*1} FCPS, ^{2,3}PGT FCPS, ^{4,5,6}BSMT

^{*1}sfaig69@hotmail.com, ²bismamustafab12@gmail.com, ³drmunazzahibrar@gmail.com,
⁴nuzhatzehra994@gmail.com, ⁵afifamehmood@gmail.com, ⁶aimanakif635@gmail.com

Corresponding Author:

Syeda Nuzhat Zehra

DOI: <http://doi.org/10.5281/zenodo.23498894>

Received	Accepted	Published
25 January 2026	10 March 2026	24 March 2026

ABSTRACT

Background: MRI may identify muscularis propria invasion in bladder cancer, but accuracy at the muscle-invasion threshold does not necessarily imply agreement across individual pathological T stages. We assessed stage-specific concordance and the direction of MRI staging discrepancies.

Methods: We analyzed MRI and histopathological T-stage classifications recorded for 60 patients with bladder cancer evaluated at a single tertiary center in Karachi, Pakistan, between April and November 2019. Pelvic MRI was performed using a 1.5-T scanner. Exact T-stage agreement, understaging, overstaging and unweighted Cohen's kappa were determined using the recorded seven-category staging matrix. Sensitivity, specificity, predictive values and accuracy for muscle-invasive ($\geq T2$) versus non-muscle-invasive ($\leq T1$) disease were calculated with 95% Wilson confidence intervals.

Results: Of 60 patients, 52 (86.7%) were male; mean age was 47.21 ± 14.82 years. Exact MRI-histopathology agreement was observed in 46 patients (76.7%; 95% CI, 64.6–85.6%), whereas MRI understaged 10 (16.7%) and overstaged four (6.7%). Unweighted kappa was 0.723. For detecting muscle invasion, MRI had 92.5% sensitivity (49/53), 71.4% specificity (5/7), 96.1% positive predictive value (49/51), 55.6% negative predictive value (5/9) and 90.0% accuracy (54/60). Four of the 14 staging discrepancies crossed the T1–T2 boundary in the direction of understaging.

Conclusion: Conventional MRI achieved high overall accuracy at the muscle-invasion threshold but lower concordance for exact T-stage assignment. Understaging was more frequent than overstaging, highlighting the need to interpret MRI together with pathological and clinical findings. These findings reflect a small, predominantly muscle-invasive, single-center cohort.

Keywords: bladder cancer; magnetic resonance imaging; histopathology; T staging; staging concordance; muscle invasion; understaging

1. Introduction

Accurate assessment of local tumor extent is essential in the management of urinary bladder cancer. Distinguishing tumors confined to the mucosa or lamina propria from those invading the muscularis propria helps determine whether patients proceed along non-muscle-invasive or

muscle-invasive treatment pathways. Histopathological evaluation remains central to diagnosis and staging, while imaging provides complementary information about disease extent and adjacent structures. [1,2]

Magnetic resonance imaging (MRI) has high soft-tissue contrast resolution and can depict the

detrusor muscle and perivesical tissues. Previous diagnostic studies have reported useful performance for identifying muscle-invasive bladder cancer, but estimates vary with patient selection, imaging protocol and reference standard. In a systematic review of MRI staging studies, pooled sensitivity and specificity for identifying muscle invasion were 92% and 87%, respectively. [2,3]

The introduction of the Vesical Imaging-Reporting and Data System (VI-RADS) subsequently standardized acquisition and interpretation of multiparametric bladder MRI, incorporating T2-weighted, diffusion-weighted and dynamic contrast-enhanced imaging. Contemporary research has increasingly focused on the optimal application of VI-RADS, yet conventional MRI datasets remain informative when their diagnostic findings are interpreted according to the methods used at the time rather than retrospectively treated as VI-RADS studies. [4-7]

Most single-center reports emphasize sensitivity and specificity for the dichotomous presence of muscle invasion. A separate question is whether MRI assigns the same specific T stage as the pathological comparator, and, when disagreement occurs, whether the imaging assessment is lower or higher than the recorded reference stage. Understanding these patterns may identify clinically important boundaries at which additional assessment is warranted.

This study aimed to quantify (i) exact MRI-histopathology T-stage concordance, (ii) the frequency and direction of MRI understaging and overstaging, and (iii) diagnostic performance for differentiating muscle-invasive from non-muscle-invasive bladder cancer in a single-center cohort.

2. Materials and Methods

2.1 Study design and participants

This single-center observational diagnostic-accuracy study evaluated 60 adult patients with bladder cancer who underwent pelvic MRI in the Department of Radiology, Sindh Institute of Urology and Transplantation (SIUT), Karachi, Pakistan, during an eight-month period from April to November 2019. Study documentation described prospective case recruitment. The recorded age range was 20–80 years, and patients

had an established or incidentally detected bladder tumor.

Reported exclusion criteria included an additional malignancy, pregnancy or lactation, claustrophobia and incomplete follow-up. A further exclusion criterion concerning invasive diagnostic procedures was not operationally defined in the available study-method documentation; therefore, it is not used to infer exclusion of histopathological assessment. Complete screening and eligibility counts were not available for a participant flow diagram.

2.2 MRI acquisition and interpretation

Pelvic MRI was acquired with a Toshiba Vantage 1.5-T system and a body phased-array coil. The documented protocol comprised T1-weighted spin-echo imaging (repetition time [TR], 400–800 ms; echo time [TE], 15–30 ms), T2-weighted spin-echo imaging (TR, 1,600–2,500 ms; TE, 80–120 ms) and short-tau inversion recovery imaging (TR, 2,000 ms; TE, 60–75 ms; inversion time, 150–275 ms). Section thickness was 4–5 mm with a 1-mm interslice gap. Axial images were acquired in every case; coronal and sagittal imaging was obtained as indicated by lesion location.

Tumors were assigned MRI T-stage categories according to the extent of mural involvement, suspected perivesical extension and extension to adjacent structures. In patients with multiple tumors, the highest MRI T category represented the patient-level stage. No contemporaneous standardized VI-RADS scores were recorded in the analyzed dataset; the results should therefore be interpreted as conventional MRI staging rather than a VI-RADS validation study.

2.3 Pathological comparator and study outcomes

MRI staging was compared with the corresponding T-stage category recorded as histopathological in the study database. The available paired matrix contained the categories T1, T2a, T2b, T3a, T3b, T4a and T4b for all 60 patients. Exact-stage concordance was defined as identical MRI and recorded histopathological T categories. MRI understaging indicated a lower MRI category than the reference category; MRI overstaging indicated a higher MRI category.

For the binary analysis, T1 was grouped as non-muscle-invasive and T2a through T4b as muscle-invasive disease. No Ta or carcinoma in situ-only

cases were present in the staging matrix. The specific surgical or tissue-sampling procedure used to assign each pathological category, especially advanced T3–T4 substages, was not documented in the available tabulation. Accordingly, the comparison is presented as concordance with recorded histopathological categories, not independently re-reviewed pathologic staging.

2.4 Statistical analysis

Patient characteristics and T-stage distributions were summarized using counts, percentages and the recorded mean and standard deviation for age. We calculated the proportion of exact stage matches and directional disagreements from the 7×7 T-stage cross-tabulation. Unweighted Cohen's kappa was used to describe chance-corrected agreement across the seven categories. Diagnostic performance for muscle invasion was calculated from the corresponding 2×2 table as

sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy. Wilson score intervals were used for 95% binomial confidence intervals. No multivariable modeling or time-to-event analysis was performed because patient-level covariates and longitudinal outcomes were not part of the available staging dataset.

3. Results

3.1 Patient and tumor-stage characteristics

The cohort included 52 males (86.7%) and eight females (13.3%) with a mean age of 47.21 ± 14.82 years. MRI assigned T1 to nine patients, T2a to six, T2b to eight, T3a to five, T3b to 11, T4a to 13 and T4b to eight. The recorded histopathological distribution was T1 ($n = 7$), T2a ($n = 8$), T2b ($n = 8$), T3a ($n = 3$), T3b ($n = 11$), T4a ($n = 15$) and T4b ($n = 8$). The full patient-level paired-stage distribution is summarized in Table 1.

Table 1. Cross-classification of MRI-assigned and histopathological T stage ($n = 60$)

MRI stage	T1	T2a	T2b	T3a	T3b	T4a	T4b	Total
T1	5	4	–	–	–	–	–	9
T2a	2	4	–	–	–	–	–	6
T2b	–	–	8	–	–	–	–	8
T3a	–	–	–	2	3	–	–	5
T3b	–	–	–	1	8	2	–	11
T4a	–	–	–	–	–	12	1	13
T4b	–	–	–	–	–	1	7	8
Total	7	8	8	3	11	15	8	60

Rows represent MRI classification; columns represent the recorded histopathological category. A dash denotes zero cases. The main diagonal represents exact-stage agreement.

3.2 Exact-stage agreement and staging discrepancies

Exact MRI–histopathology agreement was found in 46/60 patients (76.7%; 95% CI, 64.6–85.6%). Fourteen patients (23.3%) had a discordant T stage. MRI assigned a lower stage in 10 patients (16.7%; 95% CI, 9.3–28.0%) and a higher stage in four patients (6.7%; 95% CI, 2.6–15.9%). Unweighted Cohen's kappa for seven-category agreement was 0.723, consistent

with substantial chance-corrected agreement under conventional descriptive interpretations. All discordant classifications involved neighboring categories in the seven-category staging matrix.

Four MRI T1 assignments corresponded to histopathological T2a, indicating missed muscle invasion. Other understaged cases involved MRI T3a versus reference T3b ($n = 3$), MRI T3b versus T4a ($n = 2$), and MRI T4a versus T4b ($n = 1$). Conversely, MRI overstaging included T2a versus reference T1 ($n = 2$), T3b versus T3a ($n = 1$), and T4b versus T4a ($n = 1$). These transitions are detailed in Table 2.

Table 2. Direction and location of MRI staging discrepancies

MRI category	Recorded pathological category	Direction	Cases
T1	T2a	Understaging	4
T3a	T3b	Understaging	3
T3b	T4a	Understaging	2
T4a	T4b	Understaging	1
T2a	T1	Overstaging	2
T3b	T3a	Overstaging	1
T4b	T4a	Overstaging	1
Total discordances	—	—	14

3.3 Performance at the muscle-invasion threshold

When muscle-invasive disease ($\geq T2$) was regarded as the positive condition, the contingency table comprised 49 true-positive, two false-positive, five true-negative and four

false-negative cases. This yielded a sensitivity of 92.5%, specificity of 71.4%, PPV of 96.1%, NPV of 55.6% and overall accuracy of 90.0% (Table 3). The reference-positive group comprised 53/60 patients (88.3%), leaving only seven reference-negative cases.

Table 3. Diagnostic performance of MRI for identifying muscle-invasive bladder cancer

Measure	Numerator denominator	/	Estimate	95% CI
Sensitivity	49 / 53		92.5%	82.1–97.0%
Specificity	5 / 7		71.4%	35.9–91.8%
Positive predictive value	49 / 51		96.1%	86.8–98.9%
Negative predictive value	5 / 9		55.6%	26.7–81.1%
Binary accuracy	54 / 60		90.0%	79.9–95.3%
Exact seven-stage concordance	46 / 60		76.7%	64.6–85.6%

CI = confidence interval; Wilson score intervals were used. The binary threshold compares T1 (non-muscle-invasive) with T2a–T4b (muscle-invasive).

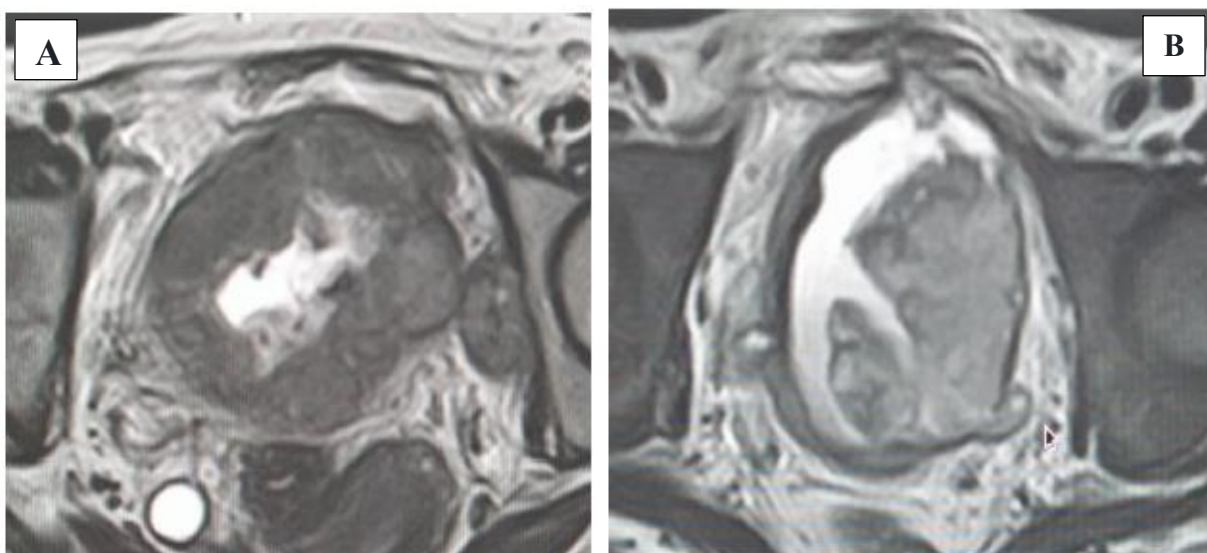


Figure 1. Axial magnetic resonance imaging (MRI) of the urinary bladder. (A) Large, heterogeneous soft-tissue mass occupying a substantial portion of the urinary bladder lumen, resulting in marked distortion of the normal intravesical architecture. (B) Corresponding axial MR image demonstrating the

lobulated morphology and heterogeneous internal signal characteristics of the lesion, with residual hyperintense fluid outlining portions of the mass.

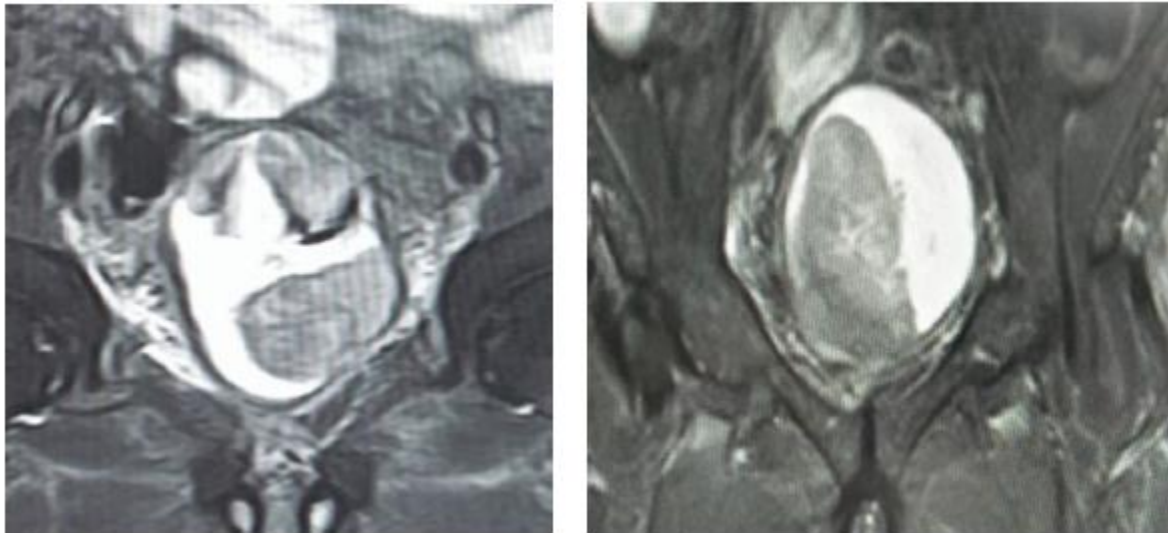


Figure 2. Coronal short tau inversion recovery (STIR) magnetic resonance imaging (MRI) of the urinary bladder demonstrating intravesical soft-tissue masses. (A) Heterogeneous, polypoid soft-tissue lesion involving the superior and left lateral walls of the urinary bladder, protruding into the bladder lumen and distorting its normal contour. (B) Large, irregular, lobulated soft-tissue mass arising from the anterior bladder wall, demonstrating substantial intraluminal extension and apparent extravescical involvement toward the anterior abdominal wall and linea alba, with reportedly associated regional lymphadenopathy.

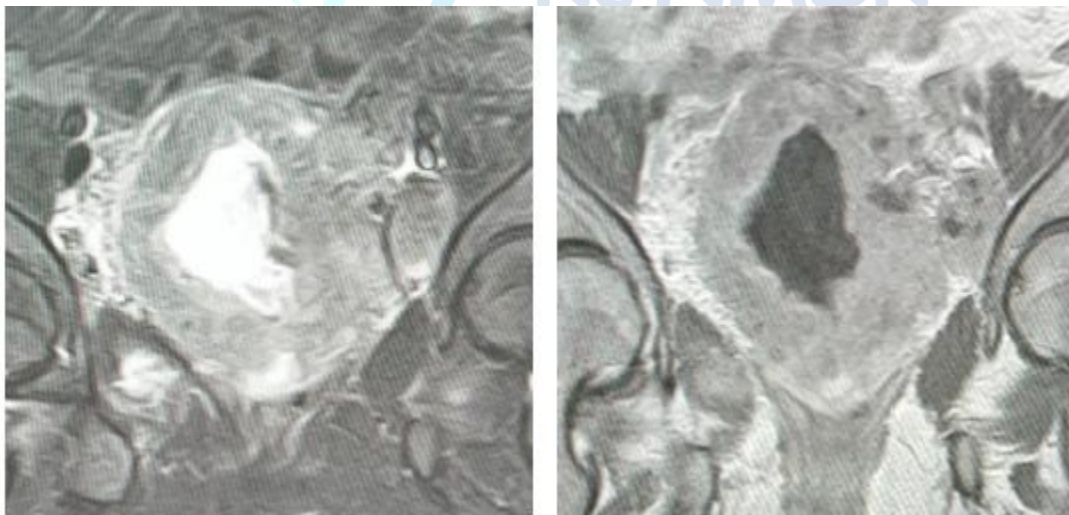


Figure 3. Coronal magnetic resonance imaging (MRI) of the pelvis demonstrating an extensive heterogeneous urinary bladder mass. (A) Large, irregular soft-tissue lesion involving the urinary bladder, with heterogeneous signal characteristics, marked distortion of the normal bladder contour, and residual hyperintense intraluminal fluid. (B) Corresponding coronal MR image demonstrating pronounced irregular bladder wall thickening and a heterogeneous soft-tissue component surrounding the residual bladder lumen, with substantial disruption of the normal vesical architecture.

4. Discussion

This study identifies an important difference between the performance of MRI at the clinically decisive muscle-invasion threshold and

its agreement with the exact T-stage category. MRI correctly classified 54 of 60 patients as muscle-invasive or non-muscle-invasive, yet agreement on the specific seven-category stage

was lower (46/60). Reporting only the 90% binary accuracy would therefore conceal the 14 individual-stage disagreements identified by the full cross-classification.

The dominant pattern was understaging: 10 patients received an MRI category below their recorded histopathological category, compared with four patients who were overstaged. Four understaged cases were assigned MRI T1 despite a recorded comparator of T2a. This boundary is clinically salient because detrusor muscle invasion often alters the nature and intensity of subsequent treatment. The findings indicate potential for under-recognition of muscle invasion on conventional MRI in a minority of cases; they do not demonstrate actual treatment delays, changes in management, recurrence or survival because such outcomes were not assessed.

The overall sensitivity of 92.5% in the present cohort is broadly compatible with the pooled sensitivity of 92% reported by Woo and colleagues for MRI detection of muscle invasion. [3] However, the 71.4% specificity observed here is imprecise because only seven comparator-negative patients were available, producing a wide confidence interval. Furthermore, the unusually high prevalence of muscle-invasive disease (88.3%) increases the observed PPV and reduces the generalizability of predictive values to populations with a lower proportion of advanced tumors. These factors argue against treating the single-center estimates as universal measures of MRI performance.

Recent practice has evolved beyond conventional anatomical MRI. VI-RADS integrates multiple sequences to communicate the likelihood of muscularis propria invasion using a standardized scoring framework. [4] Contemporary clinical reviews and comparative studies explore ways to improve reproducibility and diagnostic performance using multiparametric or biparametric protocols. [5–7] By contrast, our documented acquisition relied on conventional T1-weighted, T2-weighted and STIR sequences, and VI-RADS scores were not available. Direct numerical comparison with contemporary VI-RADS studies would therefore be inappropriate without considering technical differences, case selection and pathology procedures.

An additional interpretative challenge concerns the attempted discrimination of detailed T3 and T4 substages. Although gross perivesical extension and invasion of adjacent structures may be visible on MRI, current European Association of Urology guidance cautions that imaging cannot reliably resolve microscopic perivesical fat invasion, which underlies distinction between T2 and pathological T3a. [1] Exact-stage agreement at this level should therefore be interpreted more cautiously than the clinically important distinction between non-muscle-invasive and muscle-invasive disease. Variation in bladder distension, partial-volume effects, inflammatory change and tumor location may further influence local-staging interpretation. [2,5]

A practical implication of these findings is that MRI reports should communicate not just the assigned T stage but also the uncertainty around muscle invasion and extension into adjacent tissues. Discordance between MRI and pathological assessment merits multidisciplinary review when it affects management. Nevertheless, the present study does not establish that any individual patient benefited from or was harmed by a particular MRI stage assignment. Its contribution is the systematic description of the direction of disagreement within the evaluated cohort.

5. Study Limitations

The study was restricted to one center and 60 patients, with a strong predominance of advanced-stage cases. Precision of specificity and NPV was limited by the small non-muscle-invasive comparison group. Individual clinical covariates, tumor size and location, treatment before MRI, and details of observer experience or blinding were not available for analysis. Although the paired-stage table recorded histopathological categories, the procedure establishing each advanced pathological substage and the timing of MRI relative to tissue sampling were not fully documented. These limitations may affect stage-specific validity and should be considered when interpreting the agreement measures.

The work also predates routine availability of standardized VI-RADS scores in the recorded dataset. No imaging re-read, interobserver agreement analysis, calibration exercise,

recurrence analysis or survival assessment was performed. Reporting followed the principal diagnostic accuracy elements recommended by STARD where information was available, but some protocol and participant-flow details could not be reconstructed. [8]

6. Conclusion

In this single-center cohort, conventional pelvic MRI demonstrated 90.0% accuracy for distinguishing muscle-invasive from non-muscle-invasive bladder cancer, whereas exact agreement with the recorded histopathological T stage was 76.7%. Understaging occurred more frequently than overstaging, and four cases crossed the T1-to-T2 boundary. These results underscore the importance of describing stage-specific discrepancies and integrating MRI findings with clinical and pathological assessment. Larger prospective studies employing standardized contemporary imaging protocols and clearly defined pathology reference standards are warranted.

References

1. European Association of Urology. EAU Guidelines on Muscle-invasive and Metastatic Bladder Cancer. 2026 edition. Diagnosis: Imaging for staging. Available at: <https://uroweb.org/guidelines/muscle-invasive-and-metastatic-bladder-cancer/chapter/diagnostic-evaluation> (accessed 16 October 2026).
2. Verma S, Rajesh A, Prasad SR, et al. Urinary bladder cancer: role of MR imaging. *RadioGraphics*. 2012;32(2):371–387. doi:10.1148/rg.322115125.
3. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic performance of MRI for prediction of muscle-invasiveness of bladder cancer: a systematic review and meta-analysis. *Eur J Radiol*. 2017;95:46–55. doi:10.1016/j.ejrad.2017.07.021.
4. Panebianco V, Narumi Y, Altun E, et al. Multiparametric magnetic resonance imaging for bladder cancer: development of VI-RADS (Vesical Imaging-Reporting And Data System). *Eur Urol*. 2018;74(3):294–306. doi:10.1016/j.eururo.2018.04.029.
5. Pecoraro M, Cipollari S, Messina E, et al. Multiparametric MRI for bladder cancer: a practical approach to the clinical application of VI-RADS. *Radiology*. 2025;314(3):e233459. doi:10.1148/radiol.233459.
6. Liu P, Cai L, Jiang L, et al. Comparative diagnostic performance of VI-RADS based on biparametric and multiparametric MRI in predicting muscle invasion in bladder cancer. *BMC Med Imaging*. 2025;25:60. doi:10.1186/s12880-025-01595-3.
7. Al-Qudimat AR, Sabir D, Elamin M, et al. Implementing VIRADS score for image-guided assessment of muscle invasiveness in bladder cancer pre-TURBT: an updated meta-analysis. *Arab J Urol*. 2025;23(2):97–108. doi:10.1080/20905998.2024.2442256.
8. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi:10.1136/bmj.h5527.