

Artificial Intelligence in the Diagnosis of Urological Diseases: A Narrative Review

Thaer Saleh Sabor Al-Omary¹ , Samer Lateef Salih^{2,*}, Husam Al-hraishawi^{3,*} 

¹Department of Surgery, College of Medicine, University of Misan, Misan, Iraq.

²Al-Hakem Teaching Hospital, Maysan Health Directorate, Misan, Iraq.

³Department of Physiology, College of Medicine, University of Misan, Misan, Iraq.

Corresponding Authors Email:

samir_pioneer@yahoo.com
hra10@scarletmail.rutgers.edu

Received: 19 June 2026,

Revised: 17 August 2026,

Accepted: 3 September 2026,

DOI: [10.57238/jbb.2026.7432.1183](https://doi.org/10.57238/jbb.2026.7432.1183)



Access this article online

Copyright: ©2026 The authors. This article is published by Nabea Al-Ajyal Foundation Press and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT

AI has made progress quicker in urology compared to other surgical fields due to prostate imaging, bladder cytology, stone classification, and urodynamic assessment producing the types of structured data that machine learning algorithms work best on. This narrative review takes stock of where diagnostic AI actually stands across four areas of urological practice prostate cancer, bladder cancer, urolithiasis, and functional or infectious disorders of the lower urinary tract and asks a question that is often glossed over in enthusiastic overviews of the field: how much of this evidence has actually been tested prospectively against clinicians, rather than only benchmarked against a retrospective dataset? Prostate cancer detection on MRI has the strongest claim to an answer, following an international confirmatory study (PI-CAI) that pitted AI systems directly against a large panel of radiologists. Bladder cancer has a comparable, if smaller, body of prospective evidence in AI-assisted urinary cytology. Stone disease and functional urology, by contrast, remain almost entirely at the retrospective, single- or two-center stage. We argue that the central obstacle to clinical adoption is no longer diagnostic accuracy, which is already respectable in most published models, but the slower and less glamorous work of external validation, interpretability, and regulatory clearance.

Keywords: *Artificial intelligence, Urological diagnosis, Prospective validation, Machine learning.*

1. Introduction

A urology clinic routinely combines several distinct diagnostic skills in a single day reading a prostate MRI, grading a biopsy, interpreting a urodynamic trace, deciding whether a urine culture is worth waiting for. Each task has historically relied on individual clinical judgment, and each is known to vary from one observer to the next [1]. With the increasing number of imaging databases, digitized pathology slide

databases, and structured electronic records in hospitals, ML and DL have emerged as the obvious choice to standardize such judgments. The convolutional neural networks (CNNs), especially, are particularly appropriate for the image-centric processes involved in urology diagnoses: segmentation of organs, localization of abnormalities and grading of tissues. Apart from being accurate, the utility of such processes in a busy urology department lies in the fact that they provide a quick second opinion, do not rely on the expertise of one individual, and deliver results promptly. What tends to be underrepresented in many overviews of the literature in this field is its uneven level of maturity. Only a select few applications, notably that of prostate MRI, have come to the point of having been studied through large prospective multicenter trials comparing their performance against clinicians.

All others, namely virtually all applications in functional urology as well as many in stone diseases, are still at the stage of being studied in single-center retrospective studies to prove the concept. Comparing all such applications in one continuum can be misleading about how far along urology has come in implementing AI.

Thus, the objectives of the current narrative review can be outlined as follows: to provide an overview of the current state of affairs regarding the development of AI-based tools for urological diseases, such as prostate cancer, bladder cancer, renal cell carcinoma, urolithiasis, and disorders related to the functionality or infections of the lower urinary tract; to point out the methods which are used repeatedly in different applications; and to specify clearly, on a case-by-case basis, what remains between the promising research results and their robust clinical application. The structure of these groups and technical approaches is summarized in Figure 1.

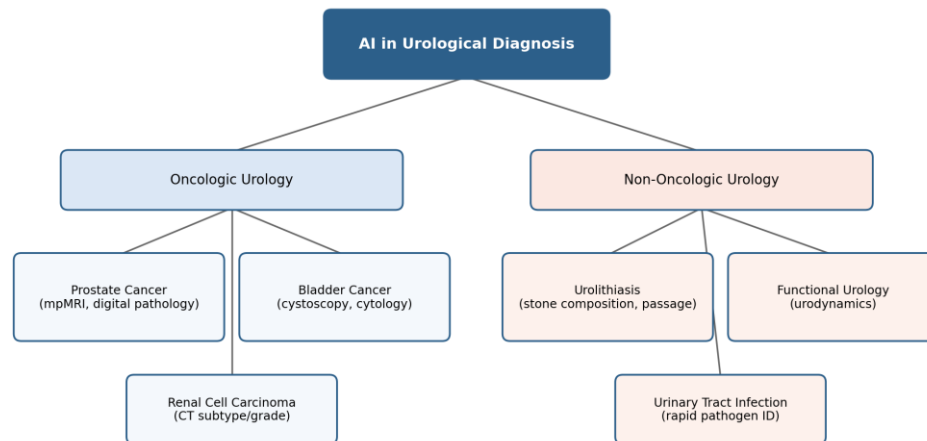


Figure 1. Overview of AI application areas across urological diagnostic practice.

Figure 1. Overview of AI application areas across urological diagnostic practice.

2. AI in Urologic Oncology

2.1. Prostate Cancer

Of all the conditions that affect urology, prostate cancer has received the most research using AI, primarily due to the nature of multiparametric MRI (mpMRI) and digital pathology being amenable to image recognition techniques. The prostate gland is now often segmented using CNN-based applications, and lesions identified by such techniques are used to highlight areas relevant to PI-RADS.

Perhaps the most compelling example in support of this assertion is the PI-CAI study, which was an international, paired, non-inferiority confirmatory trial published in *The Lancet Oncology* in 2024. The initiative recruited over 200 AI development teams and 62 radiologists from 20 different countries to compare their accuracy in detecting clinically significant prostate cancers on MRI using a reference standard, where the patient follow-up period was about five years.

The primary result was that the AI ensemble detected clinically significant tumors that were seven percent higher than the average of the radiologists' results, while they also pointed out suspicious but benign areas almost half as much - a combination that if proven reliable in further research may significantly reduce unnecessary biopsy procedures. It should be noted that the editorial accompanying the trial was cautious in interpreting the results of the PI-CAI study as evidence that the technology is "ready for prospective evaluation" but not unsupervised use because the trial itself was only a retrospective reader study [1,3].

The radiomic features that have been generated from the same images for texture, shape, and intensity descriptors which cannot be seen by the naked eye train separate classifiers for differentiating between significant and insignificant cancer, especially in the PI-RADS 3 category where there are conflicts among the radiologists.

On the pathological side, whole slide imaging is becoming a tool for detecting any malignant glands as well as assigning Gleason grades where there have been a number of studies showing that their accuracy is comparable to that of urologists, in some cases even higher, in small-volume and borderline cases [4]. The prostate cancer remains the only area in the field of urology where the chain of evidence starts from pilot study, systematic review to an international confirmatory trial.

2.2. Bladder Cancer

Though, cystoscopy is the method used to detect bladder cancer but it highly depends on experience of the operator and has limitations in diagnosing flat tumors such as carcinoma in situ. AI assisted cystoscopy involves analysis of the video obtained during the endoscopy process in order to identify suspicious regions which would normally go undetected.

It has been found to be particularly effective in identifying small papillary tumors, although much of this research is still unpublished. Nevertheless, the prospective evidence which shows higher levels of efficacy in cases of bladder cancer is not the use of cystoscopy but the artificial intelligence-assisted urinary cytology.

Indeed, the VISIOCYT1 trial in France, a prospective multicenter study, examined an artificial intelligence-based digital cytology method against standard reading of cytology and was blind to histopathology. According to the results of the trial, the artificial intelligence method provided significant improvements in sensitivity in detecting the urothelial carcinoma, compared to conventional cytology, especially in detecting low grade tumors which are usually missed by the cytology, at the price of decreased specificity a price which was clearly recognized by the investigators.

As far as staging is concerned, the possibility of using radiomics analysis on CT/MRI is under research, alongside the independent research on urinary biomarkers as a non-invasive alternative to cystoscopy.

2.3. Renal Cell Carcinoma

Renal tumors have a more limited application of AI, with the emphasis being on the differentiation between malignant and benign tumors and prediction of the tumor's histology or grade without biopsy. Deep learning classifiers based on CT images, including deep learning models that incorporate an explicit measure of uncertainty, have proven to be effective in differentiating subtypes of renal cell carcinomas and, therefore, can be used in preoperative planning rather than as a replacement of pathology.

Most research studies in the field of renal tumor detection have so far been limited to single-center data sets, and the reproducibility issue is the biggest limitation in this respect. Renal oncology has not so far provided prospective multicenter trials similar to prostate and bladder cancers [6,7].

3. AI in Non-Oncologic Urology

3.1. Urolithiasis

The imaging standard for kidney stone disease is non-contrast CT, where AI has been applied to address the three clinical questions: composition of the stone, probability of spontaneous passage, and postoperative risk of complications including sepsis. There is a difference between predicting composition of the stone using two different approaches. First, one can predict stone composition using CT attenuation values or endoscopic images taken during surgery.

Second, one can use electronic medical record data such as 24-hour urine chemistry and demographic information without any imaging at all. When applying second strategy, models usually predict stone composition rather well in the binary case of discriminating between calcium and non-calcium stones, while performance falls significantly when trying to predict the specific type of stone. This difference should be taken into account, although sometimes the summary is too simplified by stating “AI predicts stone composition” [8,9].

Other predictive algorithms have focused on calculating the probability of spontaneous passage of stones and identifying patients who are at high risk for developing infections after undergoing endoscopic procedures; more recent developments have explored the application of AI in prognosis following surgery, such as predicting recurrence based on 24-hour urine analysis.

3.2. Functional Urology and Urinary Tract Infection

The interpretation of urodynamic testing relies on the analysis of complicated pressure-flow relationships that may differ even among experienced analysts. The use of ML-based classifiers to detect features including detrusor overactivity, decreased bladder compliance, and bladder outlet obstruction is evaluated as one possible method for standardizing this process, as several retrospective analyses have shown performance similar to specialists in urology.

Similar to the case with stones, these are promising retrospective data and not yet confirmatory evidence; there have not been any achievements in the field of functional urology like those achieved in PI-CAI and VISIOCYT1 [10,11].

A parallel line of work targets urinary tract infection: because standard culture takes one to three days to return a result, models combining urinalysis parameters, microscopy, and clinical data are being developed to flag a likely infection and in some cases a likely resistance pattern well before culture confirmation. This is particularly valuable in catheter-associated and recurrent infections, where early, targeted treatment matters most.

4. Barriers to Clinical Translation

Figure 2 summarizes the current evidence tier reached by each application discussed above, from single-center retrospective work through to prospective multicenter confirmatory trials.

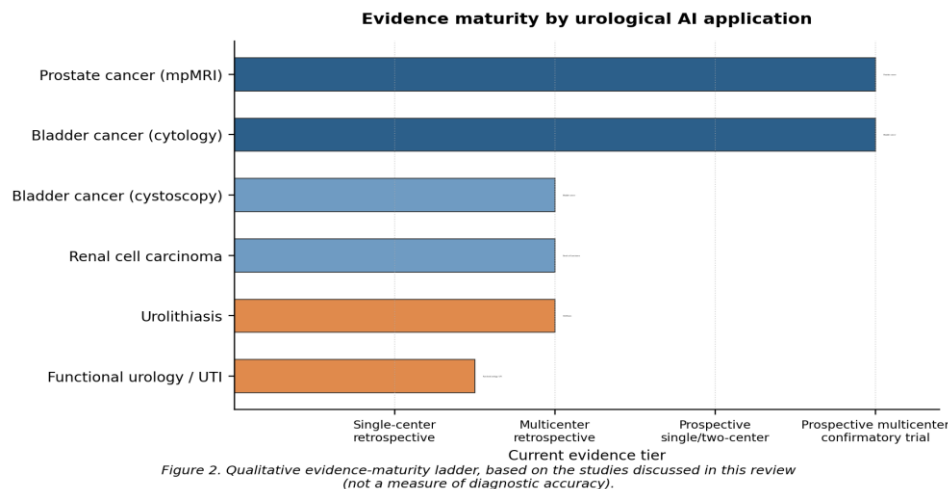


Figure 2. Qualitative evidence-maturity ladder, based on the studies discussed in this review (not a measure of diagnostic accuracy).

Three obstacles recur across all of the applications reviewed above. The first is interpretability: deep-learning models typically reach a conclusion through a process that is not readily inspectable by the treating clinician, which slows both trust and regulatory approval. The second is data quality most published models are trained and validated on single-center, retrospective cohorts, and multicenter prospective validation remains the exception rather than the norm [2].

As this review has tried to show, that exception is currently limited to prostate MRI and, to a lesser extent, bladder cytology; stone disease, renal masses, and functional urology remain almost entirely in the retrospective, single- or two-center evidence tier. The third and least-discussed obstacle is regulatory translation. The great majority of AI-enabled medical devices authorized by regulators to date sit in general radiology chest, brain, and breast imaging rather than in urology specifically [12,13].

That gap is striking given how much has been published on prostate MRI, bladder cytology, and stone characterization over the same period, and it suggests that the bottleneck for urological AI is now less about model performance than about the absence of a well-trodden regulatory pathway, standardized reporting of subgroup performance, and post-market monitoring plans that urology device developers have, so far, rarely pursued. Data privacy, bias in algorithms due to an unrepresentative set of subjects used for training, and lack of common regulatory frameworks constitute additional challenges and future development must consider fairness and transparency as core elements [14].

5. Conclusion

While artificial intelligence is gradually being developed into a clinically useful application for urological diagnosis, the extent of its maturity varies widely depending on the particular sub-specialty. Prostate cancer imaging represents the most advanced case, based on an actual international confirmatory trial; bladder cancer imaging boasts of a similar prospective trial for cytology, along with cystoscopy applications that are still in early stages of development; while stone disease, renal masses, and functional urology rely solely on retrospective, single center evidence at the moment.

In all of these domains, AI can be regarded as a decision support application rather than a replacement of human judgment, and in the future, development of this technology will be less about increasing its accuracy, which is sufficiently described in the literature, and more about addressing interpretation, prospective validation, and regulation problems.

Conflict of interest statement: The authors declare no conflict of interest.

The Authors Involved in the Research: The Thaer Saleh Sabor Al-Omary , Samer Lateef Salih and Husam Al-hraishawi, contributed to the research design to analyze the results and write the manuscript, and the authors approved the final version for submission.

Ethical Consideration: The study was approved by University of Misan, Misan, Iraq.

Funding: No Funding.

References

1. Saha A, Bosma JS, Twilt JJ, van Ginneken B, Bjartell A, Padhani AR, et al. Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. *Lancet Oncol.* 2024;25(7):879-87. [https://doi.org/10.1016/S1470-2045\(24\)00220-1](https://doi.org/10.1016/S1470-2045(24)00220-1)
2. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* 2019;17:195. <https://doi.org/10.1186/s12916-019-1426-2>
3. Eklund M. Artificial intelligence for scoring prostate MRI: ready for prospective evaluation. *Lancet Oncol.* 2024;25(7):827-8. [https://doi.org/10.1016/S1470-2045\(24\)00284-5](https://doi.org/10.1016/S1470-2045(24)00284-5)
4. Bulten W, Kartasalo K, Chen PHC, Ström P, Pinckaers H, Nagpal K, et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med.* 2022;28(1):154-63. <https://doi.org/10.1038/s41591-021-01620-2>
5. Lebre T, Paoletti X, Pignot G, Roumiguié M, Colombel M, Savareux L, et al. Artificial intelligence to improve cytology performance in urothelial carcinoma diagnosis: results from validation phase of the French, multicenter, prospective VISIOCYT1 trial. *World J Urol.* 2023;41(9):2381-8. <https://doi.org/10.1007/s00345-023-04519-4>
6. Wang Z, Zhang X, Wang X, Li J, Zhang Y, Zhang T, et al. Deep learning techniques for imaging diagnosis of renal cell carcinoma: current and emerging trends. *Front Oncol.* 2023;13:1152622. <https://doi.org/10.3389/fonc.2023.1152622>
7. Chandramohan D, Garapati HN, Nangia U, Simhadri PK, Lapsiwala B, Jena NK, et al. Diagnostic accuracy of deep learning in detection and prognostication of renal cell carcinoma: a systematic review and meta-analysis. *Front Med (Lausanne).* 2024;11:1447057. <https://doi.org/10.3389/fmed.2024.1447057>
8. Abraham A, Kavoussi NL, Sui W, Bejan C, Capra JA, Hsi R. Machine learning prediction of kidney stone composition using electronic health record-derived features. *J Endourol.* 2022;36(2):243-50. <https://doi.org/10.1089/end.2021.0211>
9. Chmiel JA, Stuivenberg GA, Wong JFW, Nott L, Burton JP, Razvi H, et al. Predictive modeling of urinary stone composition using machine learning and clinical data: implications for treatment strategies and pathophysiological insights. *J Endourol.* 2024;38(8):778-87. <https://doi.org/10.1089/end.2023.0446>
10. Mei H, Wang Z, Zheng Q, Jiao P, Lv S, Liu X, et al. Deep learning and numerical analysis for bladder outflow obstruction and detrusor underactivity diagnosis in men: a novel urodynamic evaluation scheme. *Neurourol Urodyn.* 2025;44(2):512-9. <https://doi.org/10.1002/nau.25665>

11. Liu X, Zhong P, Gao Y, Liao L. Applications of machine learning in urodynamics: a narrative review. *Neurourol Urodyn*. 2024;43(7):1617-25. <https://doi.org/10.1002/nau.25490>
12. Almarie B, Gonzalez-Gonzalez LF, Dos Santos Barbosa LA, Lutz A, Grosse U, Fregni F. Machine learning-enabled medical devices authorized by the US Food and Drug Administration in 2024: regulatory characteristics, predicate lineage, and transparency reporting. *Biomedicines*. 2025;13(12):3005. <https://doi.org/10.3390/biomedicines13123005>
13. U.S. Food and Drug Administration. Artificial intelligence-enabled medical devices [Internet]. Silver Spring (MD): FDA; 2026 [cited 2026 Sep 28]. Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
14. Char DS, Shah NH, Magnus D. Implementing machine learning in health care: addressing ethical challenges. *N Engl J Med*. 2018;378(11):981-3. <https://doi.org/10.1056/NEJMp1714229>

How to cite this article
Al-Omary TSS, Salih SL, Al-hraishawi H. Artificial Intelligence in the Diagnosis of Urological Diseases: A Narrative Review. <i>Journal of Biomedicine and Biochemistry</i> . 2026;5(3):74-79. doi: 10.57238/jbb.2026.7432.1184