

Correspondence of cystoscopy with pathology findings in the evaluation of bladder cancer: A systematic review and meta-analysis

Clara Blackburn^{1,2}, Sébastien Leclerc¹, Caio Vinícius Suartz², Typhaine Gris^{1,3}, Paul Toren^{1,3}

¹CHU de Québec-Université Laval, QC, Canada; ²Northern Ontario School of Medicine University, Thunder Bay, ON, Canada; ³Laboratoire d’Uro-Oncologie Expérimentale, Centre de recherche du CU de Québec-Université Laval, Axe Oncologie, QC, Canada

Cite as: Blackburn C, Leclerc S, Suartz CV, et al. Correspondence of cystoscopy with pathology findings in the evaluation of bladder cancer: A systematic review and meta-analysis. Can Urol Assoc J 2026 October 19; Epub ahead of print
<http://dx.doi.org/10.5489/cuaj.9675>

Published online October 19, 2026

Corresponding author: Dr. Paul Toren, CHU de Québec-Université Laval, QC, Canada; paul.toren@crchudequebec.ulaval.ca

ABSTRACT

Introduction: The diagnosis and treatment non-muscle-invasive bladder cancer relies on cystoscopy to detect bladder tumors. Few studies, even less in the modern era of cystoscopes, have assessed their diagnostic performance. We aimed to evaluate the literature on cystoscopy as a diagnostic test for pathology-confirmed bladder cancer. We further aimed to ascertain the utility of recorded images and video at cystoscopy to correctly predict the pathologic diagnosis among patients investigated for bladder cancer.

Methods: This systematic review followed the PRISMA guideline and was registered in PROSPERO. We reviewed MEDLINE/Pubmed, Cochrane, and EMBASE

KEY MESSAGES

- White-light cystoscopy demonstrates good overall diagnostic accuracy for pathology-confirmed bladder cancer, with pooled sensitivity of 83% and specificity of 79%.
- Diagnostic performance varies substantially across studies, particularly for flat lesions and across different clinical settings (primary diagnosis vs. surveillance).
- Recorded cystoscopy images and videos show diagnostic performance comparable to real-time direct observation.
- Systematic recording of cystoscopy procedures may support clinical decision-making, research standardization, education, and quality assurance.

databases from January 1995 to December 2024 for English- or French-language studies to address the research question: what is the utility of recorded images and video at cystoscopy to correctly predict the pathologic diagnosis among patients investigated for bladder cancer? Study selection was performed in duplicate, with consensus from a third reviewer to resolve differences, and the QUADAS-2 scale was used to assess the risk of bias in the study. After selecting the studies, the following variables were extracted: study characteristics, patient demographics, and diagnostic outcomes. The results were summarized in an Excel sheet and analyzed through a meta-analysis performed using RevMan software, version 5.4.

Results: Of 1165 extracted studies, we included 13 studies that reported on the correspondence of cystoscopy findings with pathology findings. Among studies with consecutive cystoscopy cohorts, six evaluated patients under investigation for bladder cancer, and five evaluated both patients with prior bladder cancer and those under investigation for bladder cancer. The sensitivity and specificity of white light cystoscopy from studies with per-lesion comparison to pathology are, respectively, 83.1% (95% confidence interval [CI] 71.9–90.5%) and 79.9% (95% CI 62.4–90.5%). Qualitatively, diagnostic results were lower for flat lesions compared to papillary lesions, with the accuracy of grade prediction variable among studies assessing this. Five studies reported data evaluating conserved photos or videos, with no significant differences in reported diagnostic test performance. Overall, most studies were at low risk of bias in the areas assessed. Nonetheless, limitations include the heterogeneity of clinical contexts and the detail of available information, with many studies indirectly related to our study question.

Conclusions: While relatively few studies directly address this topic, the current literature indicates that cystoscopy has high diagnostic value to detect pathology-confirmed bladder cancer; however, it may lead to both overdiagnosis or overtreatment of non-aggressive lesions and miss more aggressive ones. Lesion features may help indicate the level of diagnostic accuracy, while current data does not support the reliable diagnosis of grade at cystoscopy. Our findings support the use of recorded evaluations in clinical practice and for future research, which could complement real-time assessments.

INTRODUCTION

Bladder cancer is the seventh most common malignancy worldwide, with an estimated incidence of 9.5 per 100,000 men and 2.4 per 100,000 women. It is classified into two categories: non-muscle-invasive bladder cancer (NMIBC), which accounts for approximately 75% of diagnosed cases, and muscle-invasive bladder cancer (MIBC), which is associated with a higher risk of progression and metastasis. Because of its high

recurrence rate, NMIBC requires regular surveillance which renders bladder cancer one of the most expensive malignancies to manage over a lifetime.^{1,2}

Cystoscopy remains the gold standard for the diagnosis and surveillance of bladder cancer, especially for NMIBC.³ Despite its widespread use, it is an invasive procedure associated with discomfort and potential complications, including infection, bleeding, and in rare cases, bladder perforation. Moreover, data on its diagnostic performance remain limited.¹

In recent years, several technologies have been developed to assist urologists during cystoscopy, including fibre-optic and digital endoscopy, as well as enhanced techniques such as photodynamic diagnosis (PDD) and narrow-band imaging (NBI). More recently, the integration of high-resolution imaging, and integrated video recording capacity, and artificial intelligence (AI) has opened new possibilities for both clinical and educational purposes. These innovations support the creation of annotated datasets for research and machine learning^{4,5,6,7}.

Although cystoscopy plays a central role in the diagnosis and follow-up of bladder cancer, it remains a technique that depends heavily on the operator's experience and is subject to variability between observers. This poses potential problems, such as encouraging repeat cystoscopies in transfers of patients between urologists; it also decreases validity in clinical trials relying on cystoscopic endpoints. In this context, the use of recorded images and video has gained attention.⁸ Video recordings can facilitate retrospective review, peer discussions, and contribute to medical training and quality assurance. They may also serve to adjudicate recurrences in clinical trials and monitor changes over time. As the use of image and video recording becomes more common in both clinical and research environments, evaluating their diagnostic contribution is increasingly relevant.^{8,9,10}

In this systematic review we evaluated the literature supporting cystoscopy as a diagnostic test for pathology-confirmed bladder cancer, including the utility of recorded images and video.

METHODS

Search strategy

This systematic review was registered in PROSPERO (CRD42024565274) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We reviewed MEDLINE/Pubmed, Cochrane and EMBASE databases from 1995-2024 for English or French-language studies to address the research question: What is the utility of live observations or recorded images and video at cystoscopy to correctly predict the pathologic diagnosis among patients investigated for bladder cancer? Important keywords for the research were cystoscopy, bladder cancer, diagnostic and image. The complete search strategy is presented in Supplementary Table 2.

Eligibility criteria

We included both prospective and retrospective studies, whether randomized or non-randomized. Eligible studies enrolled adult patients with known or suspected bladder cancer and specifically evaluated white-light cystoscopy as a diagnostic modality. Studies were required to assess cystoscopic detection of primary or recurrent bladder tumors using direct visualization, recorded images, or video documentation. Studies assessing visual diagnostic criteria or diagnostic scale used for adjudication of diagnosis of tumor recurrence were also included. Only studies in which cystoscopic findings were compared against histopathological confirmation obtained through surgical resection or biopsy were included. Case reports and small cohort series involving fewer than 20 patients were excluded, as were studies involving the use of intravenous agents for fluorescent imaging.

Data collection

Study selection was independently performed in duplicate, with discrepancies resolved by consensus or consultation with a third reviewer. Data extraction was conducted using Rayyan (Rayyan Systems Inc. ; web-based software, version 2024), and the following variables were collected and entered into a standardized Excel spreadsheet: patient demographics, clinical characteristics, and diagnostic accuracy outcomes, including true positives, true negatives, false positives, and false negatives.

Data analysis

For the Statistical Analysis, pooled sensitivity and specificity were calculated using a bivariate random-effects model. Results were summarized using Summary Receiver Operating Characteristic (SROC) curves and Forest plots. All analyses were performed using RevMan software, version 5.4.1.

Diagnostic accuracy data were extracted as 2×2 tables. Pooled sensitivity and specificity were estimated using a bivariate random-effects model. Between-study heterogeneity was assessed using the bivariate I^2 statistic. Publication bias was evaluated with Deeks' funnel plot asymmetry test, based on a regression of log diagnostic odds ratio against the inverse square root of the effective sample size, with $p < 0.10$ indicating significant asymmetry. Analyses were performed using SPSS 29.0 (IBM Corp., Armonk, NY, USA).

Risk of bias and applicability of studies

Risk of bias and applicability concerns were assessed using the QUADAS-2 tool, which includes four key domains as recommended by the Cochrane Collaboration: patient selection, index test, reference standard, and flow and timing¹². Two independent reviewers conducted the assessment, with discrepancies resolved together with a third

reviewer. Summary and graphical representations were generated using RevMan version 5.4.1. Each domain was rated as having a high, low, or unclear risk of bias.

RESULTS

Literature search

The study selection process was conducted in accordance with the PRISMA guidelines (Fig.1). A total of 1,234 articles were initially identified through database searches. After eliminating duplicates, 1,165 studies remained. Following a screening of titles and abstracts, 1,085 articles were excluded for not meeting the inclusion criteria. A further 46 studies were excluded during the full-text assessment phase: 37 for not reporting outcomes of interest, seven due to an inappropriate study design, and two for addressing a topic outside the scope of the review. 13 articles were ultimately included in the meta-analysis, and 27 were included for a qualitative analysis.

Study characteristics

The studies varied in terms of population characteristics and clinical context. Most studies focused on patients with hematuria or those under surveillance for cancer recurrence. Study designs were predominantly prospective ($n = 11$), with a minority being retrospective ($n = 2$). The demographic characteristics of participants from the included studies are summarized in Table 1.

Quality assessment

The overall RoB assessment indicated a low risk of bias in the subdomains of “flow and timing” (85%), “reference standard” (100%), “index test” (100%), and “patient selection” (77%) (Figs. 2 and 3). However, the applicability assessments revealed limitations. 54% of studies were judged to have low concerns regarding the reference standard, 69% for the index test, and 46% for patient selection. The moderate to high concerns regarding applicability were primarily due to variations in the definition of what constituted a negative reference standard. Another contributing factor is the broad scope of our review question, which encompasses a wide range of patients investigated for bladder cancer.

Substantial between-study heterogeneity was observed in the analysis of the pool, with a bivariate I^2 statistic of 0.885. Publication bias assessment using Deeks’ funnel plot asymmetry test revealed significant asymmetry ($p < 0.001$), indicating the presence of small-study effects (Supplementary Figure 1).

Diagnostic performance of white light cystoscopy

A bivariate random-effects model was used to pool sensitivity and specificity results. For all 13 included studies, the pooled sensitivity was 83% (95% CI: 71.9-90.5%) and the specificity was 79% (95% CI: 62.4-90.5%). The Positive Likelihood Ratio was 4.1 (95% CI: 1.95-8.77%), and the Negative Likelihood Ratio was 0.2 (95% CI: 0.11-0.40) (Fig.4,

Fig. 5). This translates to an estimated 17% false negative rate and 21% false positive rate. These results include studies focusing on patients undergoing evaluation for either primary bladder cancer or suspected recurrence.

The study by Cheng et al. (2000)¹⁴, the earliest included, reported the lowest sensitivity and a notably low specificity. This may reflect the limited imaging resolution of cystoscopy available at the time. In addition, 78% of lesions assessed were flat, a morphology associated with reduced detection accuracy. This trial aimed to compare the performance of white light versus fluorescence cystoscopy following intravesical 5-aminolevulinic acid.

Similarly, Fradet et al. (2007), the second oldest study, involved only patients with carcinoma in situ (CIS) and evaluated fluorescence in comparison of white light cystoscopy. It reported both the lowest specificity and among the lowest sensitivity values. A similar decline in diagnostic performance was observed in three other studies focusing on flat lesions, which likewise reported the lowest overall results.^{17,23,24}

Factors influencing diagnostic performance: Clinical spectrum

Beyond differences in technology and lesion types, the variety of patient populations across the included studies appeared to be a factor contributing to the heterogeneity in the results. Our review includes a broad clinical spectrum, ranging from patients undergoing initial assessment for primary bladder cancer with first-time gross hematuria^{13,22}, to those being monitored for tumor recurrences¹⁸. Some studies also included mixed populations—combining both primary and recurrent cases, without clearly distinguishing diagnostic outcomes between these subgroups^{13,17}. The detection of a primary tumor can present a different diagnostic challenge than identifying a subtle recurrence in a previously treated bladder.

Cystoscopy performance for primary diagnosis

Given the heterogeneity of clinical settings across the included studies, subgroup analyses were performed. Three studies specifically evaluated cystoscopy performance in patients undergoing assessment for primary bladder cancer^{13,22,23}. This subgroup demonstrated higher pooled sensitivity and specificity; however, substantial between-study heterogeneity persisted ($I^2 = 0.89$ for sensitivity and 0.85 for specificity) (Fig 6, Fig 7). The study reporting the highest diagnostic performance, by Blick et al.¹³, included a cohort of patients referred to a hospital haematuria rapid diagnosis clinic and aimed to compare the diagnostic accuracy of computed tomography (CT) urography, flexible cystoscopy, and voided urine cytology for detecting bladder cancer.

Direct observation compared to video-based assessments

Five studies specifically evaluated the diagnostic performance of white light cystoscopy using photos or video recordings of the bladder, rather than real-time direct observation.^{18,20,26,27,28} The median sensitivity or specificity of these studies resembled the overall cohort (Figure 4).

Cystoscopy performance for predicting cancer grade

The diagnostic performance for predicting tumor stage or grade is present in a subset of studies. During et al²⁵ specifically assessed newly diagnosed patients and reported a high ability of cystoscopy to differentiate between MIBC and NMIBC, with a sensitivity of 88.9% and specificity of 91.0% for MIBC's detection. The highest results were observed for low-grade non-muscle-invasive tumors (pTa), with over 90% correctly identified. However, results decreased for more advanced non-muscle-invasive (pT1) and muscle-invasive (pT2+) lesions. Moreover, when urologists attempted to predict both tumor grade and stage, the concordance rate with histology was lower (70.3%). In contrast, Herr et al²⁹ focused on the unique context of recurrent papillary bladder tumors and demonstrated a particularly high accuracy for visual prediction of lower-risk lesions. This study reported that 93% of non-invasive and 94% of low-grade neoplasms were correctly identified by cystoscopy. The authors also provided detailed visual cues, noting that low-grade, non-invasive papillary tumors typically presented as small (<0.5 cm) lesions with discrete papillary fronds, whereas more invasive lesions appeared papillary-nodular or solid. In four studies with concomitant NBI imaging and per-lesion data on CIS, pooled data demonstrate that white light cystoscopy had a sensitivity for CIS of 61.2% (95% CI 54.8% – 67.6%) and specificity of 77.5% (68.4% – 84.5%)^{15,16,23,24}. Two studies reported outcomes for papillary tumors. Klein et al³⁰ demonstrated a sensitivity of 85% with relatively low specificity (22%). Inoue et al³¹ reported sensitivity ranging from 75.8% to 84.4% across two arms, potentially reflecting differences in group composition. These findings suggest that though the general ability to distinguish between invasive and non-invasive lesions is high, its performance varies according to tumor stage, grade, and type of lesions.

DISCUSSION

To our knowledge, this is the first systematic review to evaluate the diagnostic performance of white light cystoscopy across both primary diagnosis and surveillance settings, as well as the first to examine the performance of recorded images and videos. Cystoscopy showed good diagnostic performance in our review, but the results notably varied widely: reported sensitivities ranged from 31 % to 98 %, and specificities from 10 % to 100 % across the included studies. By comparison, a previous systematic review focused on adults presenting with haematuria found narrower ranges, with sensitivities between 87 % and 100 %, and specificities between 64 % and 100 %. The broader range

observed in our findings may be explained by the diverse patient populations included, such as individuals under surveillance for recurrent bladder cancer or those investigated following abnormal urine cytology or incidental imaging findings, and not specifically haematuria.³² It is also may be related to the publication bias we detected, with many smaller studies included in our review. Overall, the observed variability in results underscores the necessity for clinicians to interpret cystoscopy findings within the appropriate clinical framework.

While limited in number, our review did not find recorded images or videos were inferior to direct observation. Interestingly, similar approaches have been explored in gastrointestinal endoscopy. One study using the validated GiECAT tool demonstrated strong agreement between live and video-based evaluations of colonoscopy skills, with excellent intra and inter-rater reliability.³³ Prospective studies are needed to establish whether cystoscopy recordings could prove useful for educational or research purposes. Our review has limitations. The included studies showed a considerable clinical and methodological heterogeneity, particularly regarding patient populations and lesion types. The broad scope of the review, encompassing primary diagnosis and surveillance settings, as well as different clinical indications limits the generalizability of pooled estimates to specific clinical contexts. Finally, alternative cystoscopy modalities (e.g. narrow-band imaging or fluorescence cystoscopy) were excluded, narrowing the review's focus but potentially missing comparative diagnostic advances.

CONCLUSIONS

This systematic review demonstrates that white light cystoscopy maintains a high diagnostic value for detecting pathology-confirmed bladder cancer across both primary diagnosis and surveillance settings. Assessments based on recorded images or videos perform comparably to real-time evaluations, supporting their integration into clinical practice. Routine systematic recording of cystoscopy procedures offers promising opportunities for enhancing research, training, and quality assurance, by providing a valuable resource for retrospective analyses and standardization.

REFERENCES

1. Leslie SW, Soon-Sutton TL, Aeddula NR. Bladder cancer. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK536923/>. Accessed July 14, 2025.
2. Sanchez A, Agarwal N. Quantifying the costs of care among patients with high-risk non-muscle-invasive bladder cancer treated in the Veterans Health Administration. *JAMA Netw Open* 2021;4:e213816. <https://doi.org/10.1001/jamanetworkopen.2021.3816>
3. Holzbeierlein J, Bixler BR, Buckley DI, et al. Diagnosis and treatment of non-muscle invasive bladder cancer: AUA/SUO guideline: 2024 amendment. *J Urol* 2024. <https://doi.org/10.1097/JU.0000000000004452>
4. Shkolyar E, Zhou SR, Carlson CJ, et al. Optimizing cystoscopy and TURBT: Enhanced imaging and artificial intelligence. *Nat Rev Urol* 2025;22:46-54. <https://doi.org/10.1038/s41585-024-00904-9>
5. Samplaski MK, Jones JS. Two centuries of cystoscopy: The development of imaging, instrumentation and synergistic technologies. *BJU Int* 2009;103:154-8. <https://doi.org/10.1111/j.1464-410X.2008.08244.x>
6. Herr HW. Max Nitze, the cystoscope and urology. *J Urol* 2006;176:1313-6. <https://doi.org/10.1016/j.juro.2006.06.085>
7. Hwang WK, Jo SB, Han DE, et al. Artificial intelligence-based classification and segmentation of bladder cancer in cystoscopy images. *Cancers (Basel)* 2024;17:57. <https://doi.org/10.3390/cancers17010057>
8. Basile G, Pradere B, Bellmunt J, et al. A plea for the use of video-recorded cystoscopy for blinded independent central review of tumor response in open-label non-muscle invasive bladder cancer trials. *Eur Urol Oncol* 2023;6:453-5. <https://doi.org/10.1016/j.euo.2023.07.005>
9. McCroskey Z, Pambuccian SE, Kleithers, et al. Accuracy and interobserver variability of the cytologic diagnosis of low-grade urothelial carcinoma in instrumented urinary tract cytology specimens. *Am J Clin Pathol* 2015;144:902-8. <https://doi.org/10.1309/AJCPE1O9YKMRSQKG>
10. Dagnaes-Hansen J, Mahmood O, Bube S, et al. Direct observation vs. video-based assessment in flexible cystoscopy. *J Surg Educ* 2018;75:671-7. <https://doi.org/10.1016/j.jsurg.2017.10.005>
11. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
12. Basak M, Ozkurt H, Tanriverdi O, et al. Sixteen-slice multidetector computed tomographic virtual cystoscopy in the evaluation of a patient with suspected bladder tumor and history of bladder carcinoma operation. *J Comput Assist Tomogr* 2009;33:867-71. <https://doi.org/10.1097/RCT.0b013e318199d6da>
13. Blick CGT, Nazir SA, Mallett S, et al. Evaluation of diagnostic strategies for bladder cancer using computed tomography (CT) urography, flexible cystoscopy and voided urine cytology: Results for 778 patients from a hospital haematuria clinic. *BJU Int* 2012;110:84-94. <https://doi.org/10.1111/j.1464-410X.2011.10664.x>

14. Cheng CW, Lau WK, Tan PH, et al. Cystoscopic diagnosis of bladder cancer by intravesical instillation of 5-aminolevulinic acid induced porphyrin fluorescence--the Singapore experience. *Ann Acad Med Singap* 2000;29:153-8. <https://doi.org/10.47102/annals-acadmedsg.V29N2p153>
15. Fradet Y, Grossman HB, Gomella L, et al. A comparison of hexaminolevulinate fluorescence cystoscopy and white light cystoscopy for the detection of carcinoma in situ in patients with bladder cancer: A phase III, multicenter study. *J Urol* 2007;178:68-73. <https://doi.org/10.1016/j.juro.2007.03.028>
16. Geavlete B, Muțescu R, Georgescu D, et al. Hexvix blue light fluorescence cystoscopy - a promising approach in superficial bladder tumors diagnosis. *J Med Life* 2008;1:355-62.
17. Giulianelli R, Gentile BC, Albanes L, et al. Narrow band imaging (NBI) cystoscopy and assisted bipolar TURBT: A preliminary experience in a single centre. *Arch Ital Urol Androl* 2017;89:232-5. <https://doi.org/10.4081/aiua.2017.3.232>
18. Silberg Guldhammer C, Vásquez JL, Møllegaard Kristensen V, et al. Cystoscopy accuracy in detecting bladder tumors: A prospective video-confirmed study. *Cancers* 2023;16:160. <https://doi.org/10.3390/cancers16010160>
19. Herr HW, Donat SM. A comparison of white-light cystoscopy and narrow-band imaging cystoscopy to detect bladder tumour recurrences. *BJU Int* 2008;102:1111-4. <https://doi.org/10.1111/j.1464-410X.2008.07846.x>
20. Mulawkar P, Sharma G, Garge P. Evaluation of spectra A and B modes in diagnosis of suspicious bladder lesions. *J Endourol* 2021;35. <https://doi.org/10.1089/end.2020.0291>
21. Sadow CA, Silverman SG, O'Leary MP, et al. Bladder cancer detection with CT urography in an academic medical center. *Radiology* 2008;249:195-202. <https://doi.org/10.1148/radiol.2491071860>
22. Schmitz-Dräger B, Tirsar LA, Schmitz-Dräger C, et al. Immunocytology in the assessment of patients with painless gross haematuria. *BJU Int* 2008;101:455-8. <https://doi.org/10.1111/j.1464-410X.2007.07287.x>
23. Shen YJ, Zhu YP, Ye DW, et al. Narrow-band imaging flexible cystoscopy in the detection of primary non-muscle invasive bladder cancer: A "second look" matters? *Int Urol Nephrol* 2012;44:451-7. <https://doi.org/10.1007/s11255-011-0036-5>
24. Tatsugami K, Kuroiwa K, Kamoto T, et al. Evaluation of narrow-band imaging as a complementary method for the detection of bladder cancer. *J Endourol* 2010;24:1807-11. <https://doi.org/10.1089/end.2010.0055>
25. During VA, Sole GM, Jha AK, et al. Prediction of histological stage based on cystoscopic appearances of newly diagnosed bladder tumours. *Ann R Coll Surg Engl* 2016;98:547-51. <https://doi.org/10.1308/rcsann.2016.0246>
26. Kufukihara R, Kikuchi E, Shigeta K, et al. Diagnostic performance of the vesical imaging-reporting and data system for detecting muscle-invasive bladder cancer in real clinical settings: Comparison with diagnostic cystoscopy. *Urol Oncol* 2022;40:61.e1-8. <https://doi.org/10.1016/j.urolonc.2021.07.005>

27. Chang TC, Liu JJ, Hsiao ST, et al. Interobserver agreement of confocal laser endomicroscopy for bladder cancer. *J Endourol* 2013;27:598-603. <https://doi.org/10.1089/end.2012.0549>
28. Lee JY, Lee YS, Tae JH, et al. Selection of convolutional neural network model for bladder tumor classification of cystoscopy images and comparison with humans. *J Endourol* 2024;38:1036-43. <https://doi.org/10.1089/end.2024.0250>
29. Herr HW, Donat SM, Dalbagni G. Correlation of cystoscopy with histology of recurrent papillary tumors of the bladder. *J Urol* 2002;168:978-80. [https://doi.org/10.1016/S0022-5347\(05\)64555-7](https://doi.org/10.1016/S0022-5347(05)64555-7)
30. Klein C, Brunelle S, Illy M, et al. Multiparametric magnetic resonance imaging in the follow-up of non-muscle-invasive bladder tumors after intravesical instillations: A promising tool. *World J Urol* 2024;42:178. <https://doi.org/10.1007/s00345-024-04868-8>
31. Inoue K, Anai S, Fujimoto K, et al. Oral 5-aminolevulinic acid mediated photodynamic diagnosis using fluorescence cystoscopy for non-muscle-invasive bladder cancer: A randomized, double-blind, multicentre phase II/III study. *Photodiagnosis Photodyn Ther* 2015;12:193-200. <https://doi.org/10.1016/j.pdpdt.2015.03.008>
32. Devlies W, de Jong JJ, Hofmann F, et al. The diagnostic accuracy of cystoscopy for detecting bladder cancer in adults presenting with haematuria: A systematic review from the European Association of Urology Guidelines Office. *Eur Urol Focus* 2024;10:115-22. <https://doi.org/10.1016/j.euf.2023.08.002>
33. Scaffidi MA, Grover SC, Carnahan H, et al. A prospective comparison of live and video-based assessments of colonoscopy performance. *Gastrointest Endosc* 2018;87:766-75. <https://doi.org/10.1016/j.gie.2017.08.020>

FIGURES AND TABLES

Figure 1. PRISMA flow diagram of included studies.

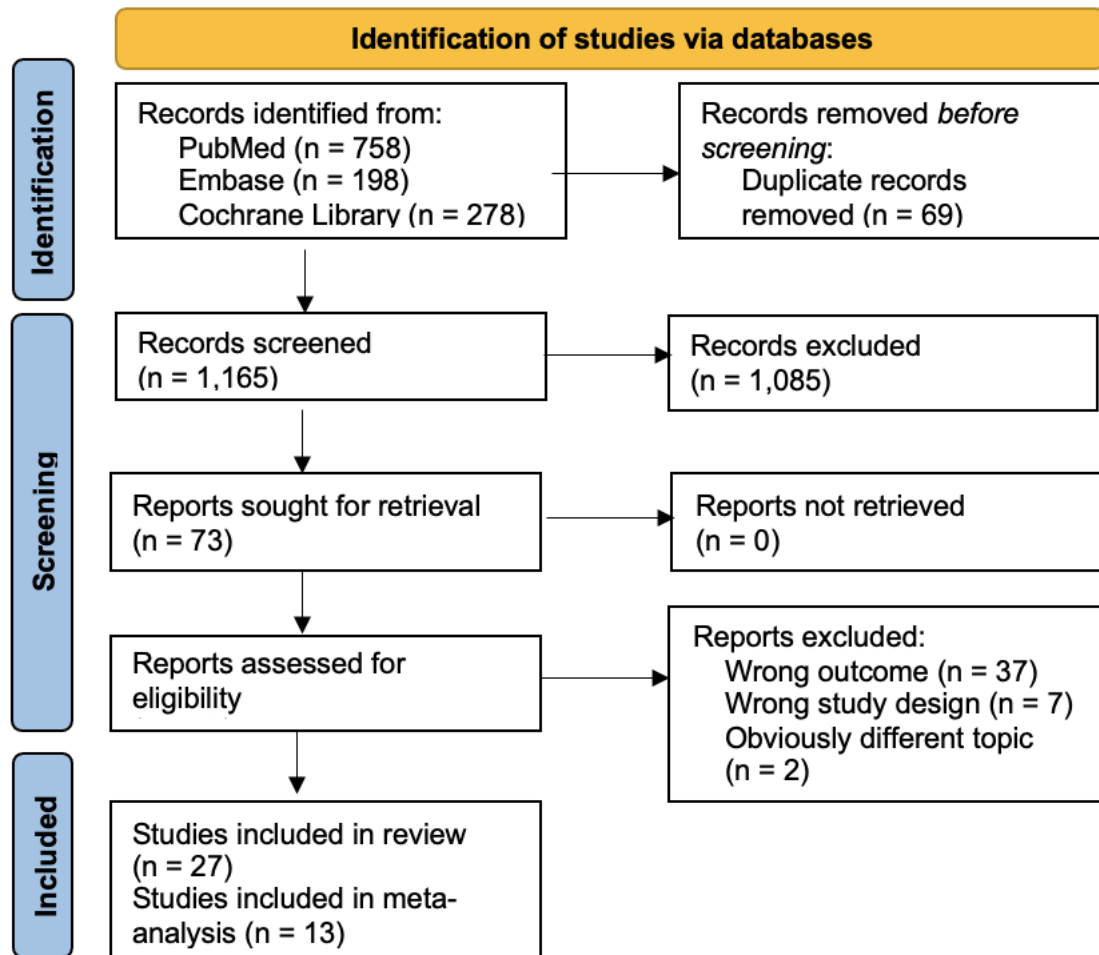


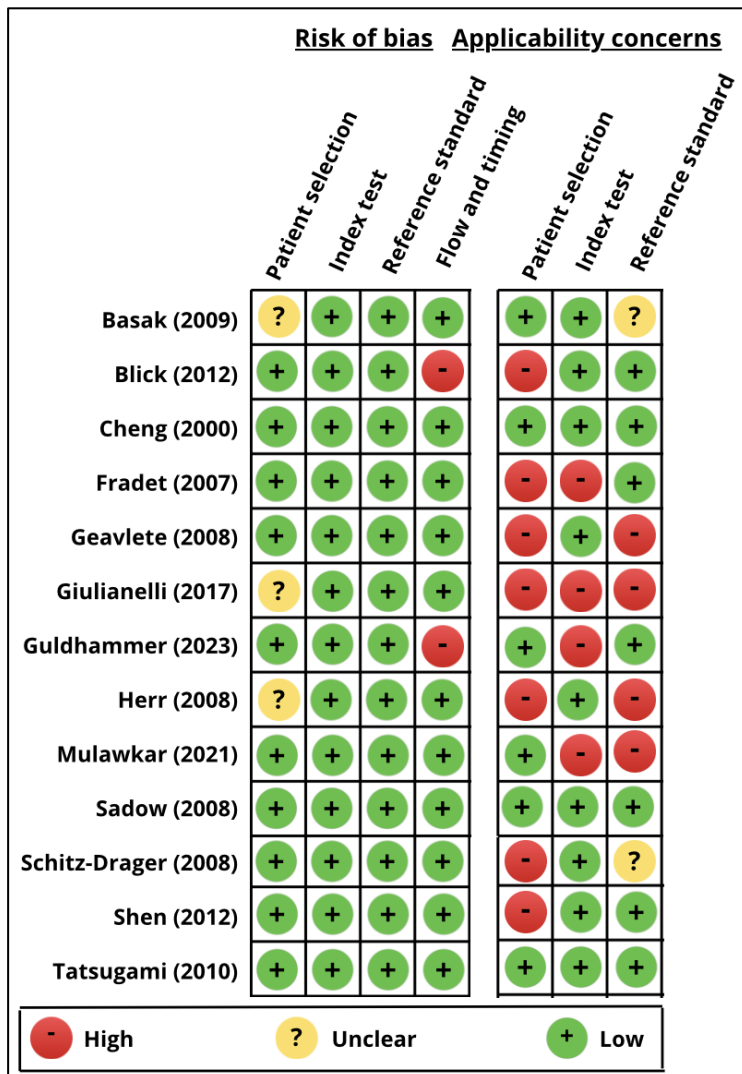
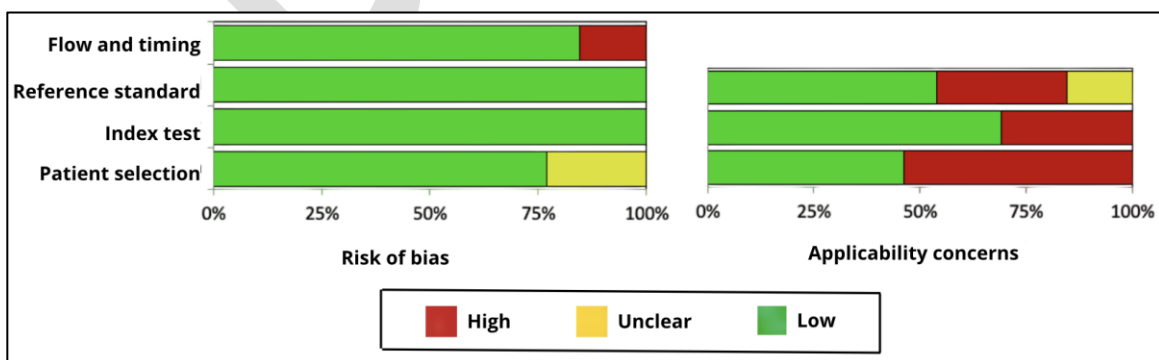
Figure 2. Risk of bias assessment of the included studies.**Figure 3.** Risk of bias and applicability concerns summary plot.

Figure 4. Receiver operating characteristic curve of all included studies.

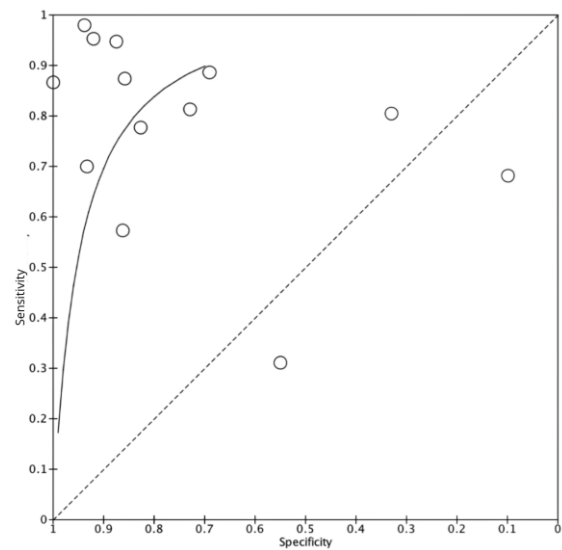


Figure 5. Forest plots of all included studies. CI: confidence interval.

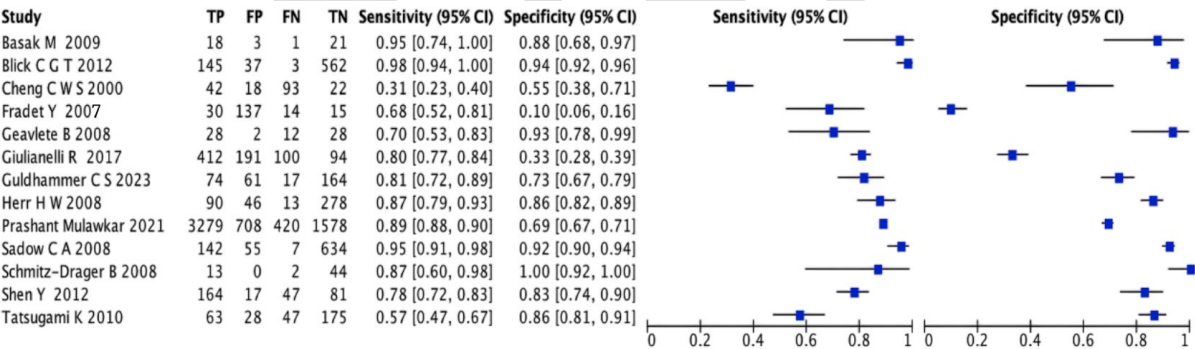


Figure 6. Receiver operating characteristic curve of the studies assessing cystoscopy at the initial diagnosis of bladder cancer.

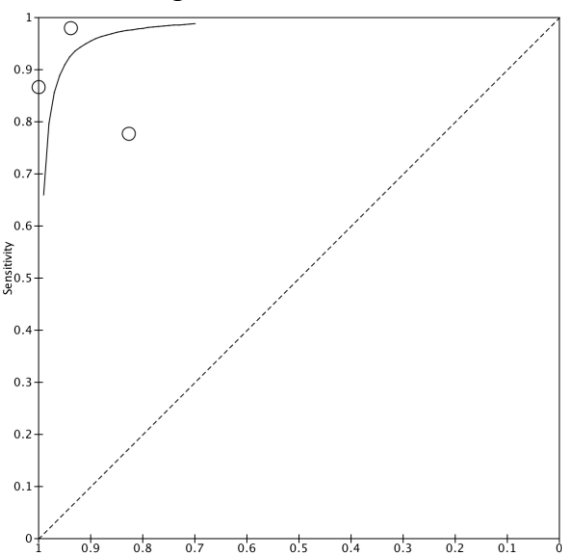


Figure 7. Forest plots of the studies assessing primary diagnoses. CI: confidence interval.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Blick C G T 2012	145	37	3	562	0.98 [0.94, 1.00]	0.94 [0.92, 0.96]		
Schmitz-Drager B 2008	13	0	2	44	0.87 [0.60, 0.98]	1.00 [0.92, 1.00]		
Shen Y 2012	164	17	47	81	0.78 [0.72, 0.83]	0.83 [0.74, 0.90]		