BMJ Open

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: a systematic review and meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-089125
Article Type:	Protocol
Date Submitted by the Author:	22-May-2024
Complete List of Authors:	BAANA, MOHAMED; Northwick Park Hospital, Urology Arkwazi, Murtada; London North West University Healthcare NHS Trust Zhao, Yi; Imperial College London School of Medicine Ofagbor, Ojone; Norwich University Hospital NHS Trust Bhardwaj, Gaurika; Imperial College London NHS Trust Lami, Mariam; Imperial College London NHS Trust Bolton, Eva; Imperial College London NHS Trust Heer, Rakesh; Imperial College London NHS Trust
Keywords:	Artificial Intelligence, UROLOGY, Diagnostic Imaging

SCHOLARONE™
Manuscripts

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: a systematic review and meta-analysis

Mohamed Baana[1], Murtada Arkwazi[1], Yi Zhao[2], Ojone Ofagbor[3], Gaurika Bhardwaj[4], Mariam Lami[4], Eva Bolton[4], Rakesh Heer[4]

MB and MA are joint first authors

EB and RH are joint senior authors

[1]London North West University Healthcare NHS Trust, London, UK

[2]Imperial College London School of Medicine, London, UK

[3]Norfolk and Norwich University Hospital NHS Trust

[4]Imperial College London NHS Trust, London, UK

Correspondence to: YZ14017@ic.ac.uk

Imperial College London

Exhibition Rd, South Kensington, London SW7 2BU

YZ14017@ic.ac.uk

Ethics: No ethical considerations apply

Word count: 1754

ABSTRACT

Introduction: Cystoscopy has revolutionised the process of diagnosing bladder cancer, leading to better categorization of risk levels and more precise treatment plans. Nonetheless, concerns arise about the lack of uniformity among observers in predicting tumour stage and grade. To address these concerns, artificial intelligence (AI) is being incorporated into clinical settings to aid in the analysis of diagnostic and therapeutic images. The subsequent report outlines a systematic review and meta-analysis protocol aimed at evaluating the effectiveness of AI in predicting bladder cancer based on cystoscopic images.

Methods and Analysis: Our systematic search will utilise databases including Pubmed, MEDLINE, Embase, and Cochrane. The articles published between May 2015 and April 2024 will be eligible for inclusion. For articles to be considered, they must employ AI for analysis of cystoscopic images to identify bladder cancer, present original data and be written in English. The protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) 2015 checklist. Quality and bias risk across chosen studies will be evaluated using the QUADAS-2 score.

Ethics and dissemination: Ethical clearance won't be necessary for conducting this systematic review since results will be disseminated through peer-reviewed publications and presentations at both national and international conferences.

PROSPERO registration number: CRD42024528345

Strengths and limitations of this study

- This protocol adheres to PRISMA guidelines and will incorporate subgroup and sensitivity analyses to further explore the variability among the studies included.
- AI-specific metrics like the F1-score and precision-recall AUC will be utilised to address limitations inherent in traditional pooled analysis, such as the impact of imbalanced classes.
- Given the novelty of AI technology in cystoscopy, long-term data regarding its impact on clinical outcomes may be scarce.
- Limitations on language and the exclusive use of cystoscopy may result in a limited number of eligible studies for inclusion

BACKGROUND

Bladder cancer is the commonest cancer of the urinary tract, and ranks as the ninth most common cancer worldwide according to the World Health Organization's 2020 report(1,2). The gold standard for diagnosing bladder cancer remains cystoscopy, as advocated by both the National Institute for Health and Care Excellence (NICE) and the European Association of Urology (EAU) guidelines(3). While white light cystoscopy (WLC) is the conventional method widely employed, however it potentially overlooks up to 20% of lesions. Furthermore, a recent systematic review and meta-analysis demonstrated that conventional WLC exhibits low diagnostic sensitivity compared to alternative modalities like blue light cystoscopy (BLC)(4). Despite the superior detection rates of bladder cancer associated with these alternative methods recommended by EAU, their adoption remains limited, likely due to their higher initial costs and limited availability(4). Transurethral resection of bladder tumour (TURBT) with white light cystoscopy (WLC) remains fundamental for diagnosing and treating non-muscle invasive bladder cancer (NMIBC) which accounts for around 75% of bladder tumours at the time of diagnosis(5).

Artificial intelligence (AI) now boasts a remarkable ability to accurately recognise images. AI offers a promising solution to improve the diagnosis of bladder cancer during cystoscopy(6). Artificial intelligence broadly describes the modelling of intelligent behaviour by use of a computer model(7). Deep learning is a subset of artificial intelligence which more specifically positions AI within the context of medical imaging(8). Augmented cystoscopy employing deep learning holds promise in also enhancing tumour localization, intraoperative navigation, and surgical resection of bladder cancer during TURBT(9).

While Artificial intelligence shows promise in the diagnosis of bladder cancer using cystoscopy, several limitations to the deployment of this technology need to be addressed. Given the cystoscopy imaging data used, such as WLC and BLC, studies are strongly encouraged to follow the Checklist for Artificial Intelligence in Medical Imaging (CLAIM)(10). This is to encourage the reproducibility of AI-models in development and forwarding the collaboration of research groups in external validation of their models. The checklist also provide a focus on the use of radiomic features as well as computer-aided diagnosis of imaging data. In order to be better suited for clinical development, models should be explainable in their decision-making process which may be currently under reported(11,12).

Existing reviews on this topic have provided a robust summary to the feasibility of the application of AI in cystoscopy(13,14). But further investigation of the reported studies with a goal of clinical deployment should be conducted next. Therefore, this systematic review and meta-analysis seeks to outline the precision of artificial intelligence (AI) in forecasting bladder cancer based on cystoscopic images and evaluate its potential influence on patient clinical outcomes.

METHODS AND ANALYSIS:

This systematic review protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 checklist(15). This study has been prospectively registered with the PROSPERO review database (CRD42024528345), and all methodologies detailed herein have been established prior to implementation. The statistical analysis will focus on evaluating the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area-under-the-curve (AUC) associated with the application of artificial intelligence (AI) in detecting bladder cancer during cystoscopy, along with its impact on clinical outcomes. These parameters will be derived through comprehensive analysis and thematic synthesis of included studies. Pooled sensitivities and specificities across the studies will be determined before calculating PPV and NPV values. AI-specific matrices such as F1-scores and Precision-Recall AUC will also be investigated.

Search Methodology

A comprehensive search will be conducted across multiple databases including PubMed, MEDLINE, Embase, and the Cochrane Library. The search strategy will incorporate medical subject heading (MeSH) terms and free text combined with appropriate Boolean operators. Articles from May 2015 to April 2024 will be included to ensure a thorough retrieval of relevant evidence. The search will encompass the following key terms: "bladder", "cancer", "diagnosis", “cystoscopy”, “cystoscopic images”, "artificial intelligence", and "deep learning". The complete search strategy is outlined in Supplementary file 1. To streamline the initial screening phase, we will be utilising Rayyan, a semi-automated tool crafted to enhance the efficiency and precision of systematic review (16). All eligible articles identified in the initial search will be imported into Rayyan. Additionally, a manual examination of references cited in all included articles will be conducted to uncover any additional pertinent literature not captured by the initial search strategy. In instances where data is lacking or unclear, corresponding authors will be contacted for clarification.

Study selection and data extraction

Two researchers, M.B. and M.A., will independently conduct the screening process. They will carefully review the titles and abstracts of eligible studies, eliminating any irrelevant articles. Full-text versions of relevant articles will then be retrieved for further evaluation. In the event of any discrepancies between the researchers, a third reviewer (Y.Z.) will be consulted, and a consensus will be reached through discussion. The reasons for excluding articles will be meticulously documented and outlined in a PRISMA flow diagram. Prior to commencing the screening process, calibration exercises will be conducted to ensure consistency among the researchers, thereby minimising potential inter-reviewer bias.

Inclusion and exclusion criteria

This systematic review will include studies employing either fully automated or semi-automated artificial intelligence (AI) for analysing cystoscopic images to detect bladder cancer. Both prospective and retrospective studies will be considered. The main comparisons will focus on evaluating sensitivities, specificities, positive predictive values (PPVs), negative predictive values (NPVs), and area under the curve (AUC) values. Patient cohorts may include individuals with suspected or confirmed bladder cancer cases, with histological findings serving as reference standards.

Excluded from analysis will be correspondence papers, ongoing studies, case reports, and conference abstracts. Additionally, non-English language articles, studies not utilising cystoscopy as the primary diagnostic modality, and those involving patients with a history of previous bladder cancer treatment will be excluded.

Data extraction (table of collection)

The data outlined in Table 1 will be gathered from all selected studies. Each researcher will independently conduct data extraction, consolidating the obtained information into a comprehensive datasheet. Any discrepancies in data extraction will be reviewed by a third evaluator, with the aim of reaching consensus for resolution.

If accessible, pertinent figures such as true positives, true negatives, false positives, false negatives, and their derived calculations will be extracted accordingly. If these figures are not explicitly provided, efforts will be made to compute them from available data. In cases where computation is unattainable, the authors of the respective paper will be contacted to provide the required data.

Table 1. Data collection items

Item No.	**Data Title**	**Data Type**
1	Year of publication	Study characteristic
2	Study authors	Study characteristic
3	Patient population	Demographics
4	Study size	Demographics
5	Cystoscopic images	Methodology
6	Histopathology results	Methodology
7	AI models utilised	Methodology

8	Definition for significant clinical disease	Methodology

AI: artificial intelligence.

Endpoints

The main endpoint of analysis will be the statistically significant quantification of accuracy when employing AI in bladder cancer detection during cystoscopy, aiming to assess its potential impact on clinical outcomes. Additional outcomes will encompass various parameters examining patient demographics.

Meta-analysis

Should an ample number of suitable studies be accessible, we will proceed with a meta-analysis to amalgamate a quantitative measure of AI performance in identifying bladder cancer from cystoscopic images. Initially, sensitivity and specificity values will be retrieved from studies, or if not accessible, computed from clinical data or solicited from authors. If a notable fraction of studies employ alternative metrics like F1-score or precision-recall AUC, these metrics will be acquired and scrutinised independently. The distributions of untransformed, logit, and double-arcsine transformed proportions will be compared and evaluated for normality using density plots and Shapiro-Wilk tests. The set of ratios most resembling a normal distribution will be selected for further analysis.

Heterogeneity and inter-study variation will be quantified using I2, and if statistically significant, a random-effects model will be employed for estimating the summary estimate. Leave-one-out analysis (LOO) and accompanying diagnostic plots will be used to identify influential studies after the model fits all relevant studies. Summary

proportions will be re-estimated based on the remaining studies after each study is removed. Studies exerting a statistically significant influence on the model will be identified as outliers and excluded. After excluding these studies, the model will then be refitted with a summary estimate comprising the remaining studies will be calculated to estimate the accuracy using AI in cystoscopy for bladder cancer detection. All data analysis and visualisation will be conducted using the R statistical environment with the "mvmeta" and "meta" packages.

Risk of bias in individual studies : This study will use the QUADAS-2 tool to assess the quality and the risk of bias within each of the included studies (17). The scoring system is split into four main sections: patient selection, index test, reference standard, and flow and timing. Within each section, signalling questions assessed the quality of the research methodology and results at the study level. As included studies utilises imaging data, the CLAIM-AI checklist will be used to assess imaging specific considerations such as the classification, image reconstruction, text analysis and workflow optimisation (10). Two reviewers (M.B. and M.A.) will independently engage in this procedure, with any discrepancies resolved through consensus. The bias assessment will analyse the appropriateness and dependability of the data utilised. This analysis will contribute to evidence synthesis and enhance transparency, potentially leading to the exclusion of articles of low quality or indicative of high bias levels. If included, relevant commentaries will be integrated into the discussion.

Ethics and Dissemination

The potential of AI to assist in diagnosing bladder cancer during cystoscopy remains to be thoroughly assessed. Our aim is to consolidate available information to elucidate AI's role in cystoscopy for bladder cancer diagnosis and outline its impact on clinical practice.

However, it is crucial to acknowledge several limitations. Firstly, given the novelty of AI in this domain, there may be a limited number of studies available for evaluation. Additionally, issues related to the diversity of reporting protocols due to regional variations may arise. Furthermore, the presence of different neural networks could potentially act as confounders due to variations in training datasets, leading to discrepancies in reported outcomes.

We intend to disseminate our findings through publication in a peer-reviewed journal. The synthesis of this data will contribute to a deeper understanding of current AI methodologies in bladder cancer diagnosis. Ultimately, this knowledge may facilitate the enhancement of existing systems, both supervised and unsupervised, to refine reporting protocols and improve bladder cancer diagnosis in the future.

Patient and public involvement

There will be no patient or public involvement in this study.

Trial status

- **-** Preliminary searches: Started
- - Piloting of the study selection process: Not started
- - Formal screening: Not started
- - Data extraction: Not started

- - Risk of bias assessment: Not started
- - Data analysis: Not started

Author Affiliations

[1]London North West University Healthcare NHS Trust, London, UK

[2]Imperial College London School of Medicine, London, UK

[3]Norfolk and Norwich University Hospital NHS Trust

[4]Imperial College London NHS Trust, London, UK

Email: Mohamed Baana: m.baana1@nhs.net; Murtada Arkwazi: murtada.arkwazi@nhs.net; Yi Zhao: yi.zhao18@imperial.ac.uk, Ojone Ofagbor: ojone.lawrence@nhs.net, Gaurika Bhardwaj: gaurika.bhardwaj@nhs.net, Mariam Lami: mariam.lami1@nhs.net, Eva Bolton: eva.bolton@nhs.net, Rakesh Heer: r.heer@imperial.ac.uk

Contributors The authors' contribution includes, but is not limited to, the following: MB and MA wrote the manuscript. YZ created the study concept. LM, EB and RH provided supervision and guidance, GB and OO checked the manuscript in its current form. YZ and MB are the guarantors of this work.

Funding Imperial College London Open Access Fund

Competing Interests None

Patient consent for publication Not required

Provenance and peer review Not commissioned; externally peer-reviewed

ORCID IDs

- Mohamed Baana: https://orcid.org/0009-0006-0384-9423
- Murtada Arkwazi: https://orcid.org/0009-0001-1344-9849
- Yi Zhao: https://orcid.org/0000-0002-4563-4344
- Ojone Ofagbor: https://orcid.org/0009-0004-5033-9050
- Gaurika Bhardwaj: https://orcid.org/0009-0008-7502-9920
- Mariam Lami: https://orcid.org/0000-0002-5485-8315
- Eva Bolton: https://orcid.org/0000-0002-1191-0849

References

(1) Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209–49.

(2) Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Int J Cancer. 2021 Apr 5;

(3) Witjes JA, Bruins HM, Cathomas R, Compérat EM, Cowan NC, Gakis G, et al. European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. European Urology. 2021 Jan 1;79(1):82–104.

(4) Russo GI, Sholklapper TN, Cocci A, Broggi G, Caltabiano R, Smith AB, et al. Performance of Narrow Band Imaging (NBI) and Photodynamic Diagnosis (PDD) Fluorescence Imaging Compared to White Light Cystoscopy (WLC) in Detecting Non-Muscle Invasive Bladder Cancer: A Systematic Review and Lesion-Level Diagnostic Meta-Analysis. Cancers. 2021 Jan;13(17):4378.

(5) Bochenek K, Aebisher D, Międzybrodzka A, Cieślar G, Kawczyk-Krupka A. Methods for bladder cancer diagnosis – The role of autofluorescence and photodynamic diagnosis. Photodiagnosis and Photodynamic Therapy. 2019 Sep 1;27:141–8.

(6) Chan EOT, Pradere B, Teoh JYC. The use of artificial intelligence for the diagnosis of bladder cancer: a review and perspectives. Current Opinion in Urology. 2021 Jul;31(4):397.

(7) Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017 Apr;69S:S36–40.

(8) Howard J. Artificial intelligence: Implications for the future of work. Am J Ind Med. 2019 Nov;62(11):917–26.

(9) Shkolyar E, Jia X, Chang TC, Trivedi D, Mach KE, Meng MQH, et al. Augmented Bladder Tumor Detection Using Deep Learning. European Urology. 2019 Dec 1;76(6):714–8.

(10) Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. Radiol Artif Intell. 2020 Mar;2(2):e200029.

(11) de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. npj Digit Med. 2022 Jan 10;5(1):1–13.

(12) Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health. 2021 Nov 1;3(11):e745–50.

(13) Ferro M, Falagario UG, Barone B, Maggi M, Crocetto F, Busetto GM, et al. Artificial Intelligence in the Advanced Diagnosis of Bladder Cancer-Comprehensive Literature Review and Future Advancement. Diagnostics (Basel). 2023 Jul 7;13(13):2308.

(14) Borhani S, Borhani R, Kajdacsy-Balla A. Artificial intelligence: A promising frontier in bladder cancer diagnosis and outcome prediction. Critical Reviews in Oncology/Hematology. 2022 Mar 1;171:103601.

(15) Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews. 2015 Jan 1;4(1):1.

(16) Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016 Dec 5;5(1):210.

(17) Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011 Oct 18;155(8):529–36.

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: a systematic review and meta-analysis

The search terms for this systematic review are as follows:

((artificial intelligence OR machine intelligence OR computational intelligence OR neural network OR deep learning OR machine learning OR reinforcement learning OR convolutional neural network OR artificial neural network OR recurrent neural network OR deep neural network OR intelligent model OR AI OR NN OR DL OR ML OR CNN OR ANN OR RNN OR DNN) AND ((Bladder AND (cancer OR malignant OR transitional cell or urothelial OR squamous cell OR adenocarcinoma OR papillary OR carcinoma in situ OR non muscle invasive OR lesions OR disease)) AND (Cystoscopy OR cystoscopic images OR flexible cystoscopy OR cystourethroscopy)) AND diagnosis).

Searches will be conducted on the following databases:

PubMed, MEDLINE, Embase, and Cochrane databases

Reporting checklist for protocol of a systematic review and meta analysis.

Based on the PRISMA-P guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the PRISMA-Preporting guidelines, and cite them as:

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

		Reporting Item	Page Number
Title			
Identification	#1a	Identify the report as a protocol of a systematic review	1
Update	#1b	If the protocol is for an update of a previous systematic review, identify as such	n/a
Registration			
	#2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors			
Contact	#3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1, 12
Contribution	#3b	Describe contributions of protocol authors and identify the guarantor of the review	12

Amendments

	#4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	n/a

Support

Sources	#5a	Indicate sources of financial or other support for the review	n/a
Sponsor	#5b	Provide name for the review funder and / or sponsor	n/a
Role of sponsor or funder	#5c	Describe roles of funder(s), sponsor(s), and / or institution(s), if any, in developing the protocol	n/a

Introduction

Rationale	#6	Describe the rationale for the review in the context of what is already known	2
Objectives	#7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	2

Methods

Eligibility criteria	#8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	2
Information sources	#9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	5-6
Search strategy	#10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Supplementary file
Study records - data management	#11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	7-8

Study records - selection process	#11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	7-8
Study records - data collection process	#11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	6-8
Data items	#12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6-8
Outcomes and prioritization	#13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	9-10
Risk of bias in individual studies	#14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	10
Data synthesis	#15a	Describe criteria under which study data will be quantitatively synthesised	8-10
Data synthesis	#15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I2, Kendall's τ)	8-10
Data synthesis	#15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	8-10
Data synthesis	#15d	If quantitative synthesis is not appropriate, describe the type of summary planned	8-10
Meta-bias(es)	#16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	8-10
Confidence in cumulative evidence	#17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	8-10

 This checklist was completed on 20. March 2023 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: A protocol for a systematic review and meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-089125.R1
Article Type:	Protocol
Date Submitted by the Author:	14-Aug-2024
Complete List of Authors:	BAANA, MOHAMED; Northwick Park Hospital, Urology Arkwazi, Murtada; London North West University Healthcare NHS Trust Zhao, Yi; Imperial College London School of Medicine Ofagbor, Ojone; Norwich University Hospital NHS Trust Bhardwaj, Gaurika; Imperial College London NHS Trust Lami, Mariam; Imperial College London NHS Trust Bolton, Eva; Imperial College London NHS Trust Heer, Rakesh; Imperial College London NHS Trust
Primary Subject Heading:	Urology
Secondary Subject Heading:	Surgery
Keywords:	Artificial Intelligence, UROLOGY, Diagnostic Imaging

SCHOLARONE™
Manuscripts

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: A Protocol for a systematic review and meta-analysis

Mohamed Baana[1], Murtada Arkwazi[1], Yi Zhao[2], Ojone Ofagbor[3], Gaurika Bhardwaj[4], Mariam Lami[4], Eva Bolton[4], Rakesh Heer[4]

MB and MA are joint first authors

EB and RH are joint senior authors

[1]London North West University Healthcare NHS Trust, London, UK

[2]Imperial College London School of Medicine, London, UK

[3]Norfolk and Norwich University Hospital NHS Trust

[4]Imperial College London NHS Trust, London, UK

Correspondence to: YZ14017@ic.ac.uk

Imperial College London

Exhibition Rd, South Kensington, London SW7 2BU

YZ14017@ic.ac.uk

Ethics: No ethical considerations apply

Word count: 1754

ABSTRACT

Introduction: Cystoscopy has revolutionised the process of diagnosing bladder cancer, leading to better categorization of risk levels and more precise treatment plans. Nonetheless, concerns arise about the lack of uniformity among observers in predicting tumour stage and grade. To address these concerns, artificial intelligence (AI) is being incorporated into clinical settings to aid in the analysis of diagnostic and therapeutic images. The subsequent report outlines a systematic review and meta-analysis protocol aimed at evaluating the effectiveness of AI in predicting bladder cancer based on cystoscopic images.

Methods and Analysis: Our systematic search will utilise databases including Pubmed, MEDLINE, Embase, and Cochrane. The articles published between May 2015 and April 2024 will be eligible for inclusion. For articles to be considered, they must employ AI for analysis of cystoscopic images to identify bladder cancer, present original data and be written in English. The protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) 2015 checklist. Quality and bias risk across chosen studies will be evaluated using the QUADAS-2 score.

Ethics and dissemination: Ethical clearance won't be necessary for conducting this systematic review since results will be disseminated through peer-reviewed publications and presentations at both national and international conferences.

PROSPERO registration number: CRD42024528345

Strengths and limitations of this study

- This protocol adheres to PRISMA guidelines and will incorporate subgroup and sensitivity analyses to further explore the variability among the studies included.
- AI-specific metrics like the F1-score and precision-recall AUC will be utilised to address limitations inherent in traditional pooled analysis, such as the impact of imbalanced classes.
- Given the novelty of AI technology in cystoscopy, long-term data regarding its impact on clinical outcomes may be scarce.
- Limitations on language and the exclusive use of cystoscopy may result in a limited number of eligible studies for inclusion

BACKGROUND

Bladder cancer is the commonest cancer of the urinary tract, and ranks as the ninth most common cancer worldwide according to the World Health Organization's 2020 report(1,2). The gold standard for diagnosing bladder cancer remains cystoscopy, as advocated by both the National Institute for Health and Care Excellence (NICE) and the European Association of Urology (EAU) guidelines(3). While white light cystoscopy (WLC) is the conventional method widely employed, it potentially overlooks up to 20% of lesions. Furthermore, a recent systematic review and meta-analysis demonstrated that conventional WLC exhibits low diagnostic sensitivity compared to alternative modalities like blue light cystoscopy (BLC)(4). Despite the superior detection rates of bladder cancer associated with these alternative methods recommended by EAU, their adoption remains limited, likely due to their higher initial costs and limited availability(4). Transurethral resection of bladder tumour (TURBT) with WLC remains fundamental for diagnosing and treating non-muscle invasive bladder cancer (NMIBC) which accounts for around 75% of bladder tumours at the time of diagnosis(5).

Artificial intelligence (AI) now boasts a remarkable ability to accurately recognise images. AI offers a promising solution to improve the diagnosis of bladder cancer during cystoscopy(6). Artificial intelligence broadly describes the modelling of intelligent behaviour by use of a computer model(7). Deep learning is a subset of artificial intelligence which more specifically positions AI within the context of medical imaging(8). Augmented cystoscopy employing deep learning holds promise in also enhancing tumour localization, intraoperative navigation, and surgical resection of bladder cancer during TURBT(9).

While Artificial intelligence shows promise in the diagnosis of bladder cancer using cystoscopy, several limitations to the deployment of this technology need to be addressed. Given the cystoscopy imaging data used, such as WLC and BLC, studies are strongly encouraged to follow the Checklist for Artificial Intelligence in Medical Imaging (CLAIM)(10). This is to encourage the reproducibility of AI models in development and forwarding the collaboration of research groups in external validation of their models. The checklist also provides a focus on the use of radiomic features as well as computer-aided diagnosis of imaging data. In order to be better suited for clinical development, models should be explainable in their decision-making process which may be currently under reported(11,12).

Existing reviews on this topic have provided a robust summary to the feasibility of the application of AI in cystoscopy(13,14). But further investigation of the reported studies with a goal of clinical deployment should be conducted next. Therefore, this systematic review and meta-analysis seeks to outline the precision of artificial intelligence (AI) in forecasting bladder cancer based on cystoscopic images and evaluate its potential influence on patient clinical outcomes.

METHODS AND ANALYSIS:

This systematic review protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 checklist(15). This study has been prospectively registered with the PROSPERO review database (CRD42024528345), and all methodologies detailed herein have been established prior to implementation. The statistical analysis will focus on evaluating the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area-under-the-curve (AUC) associated with the application of artificial intelligence (AI) in detecting bladder cancer during cystoscopy, along with its impact on clinical outcomes. These parameters will be derived through comprehensive analysis and thematic synthesis of included studies. Pooled sensitivities and specificities across the studies will be determined before calculating PPV and NPV values. AI-specific matrices such as F1-scores and Precision-Recall AUC will also be investigated.

Search Methodology

A comprehensive search will be conducted across multiple databases including PubMed, MEDLINE, Embase, and the Cochrane Library. The search strategy will incorporate medical subject heading (MeSH) terms and free text combined with appropriate Boolean operators. Articles from May 2015 to April 2024 will be included to ensure a thorough retrieval of relevant evidence. The search will encompass the following key terms: "bladder", "cancer", "diagnosis", “cystoscopy”, “cystoscopic images”, "artificial intelligence", and "deep learning". The complete search strategy is outlined in Supplementary file 1. To streamline the initial screening phase, we will be utilising Rayyan, a semi-automated tool crafted to enhance the efficiency and precision of systematic review (16). All eligible articles identified in the initial search will be imported into Rayyan. Additionally, a manual examination of references cited in all included articles will be conducted to uncover any additional pertinent literature not captured by the initial search strategy. In instances where data is lacking or unclear, corresponding authors will be contacted for clarification.

Study selection and data extraction

Two researchers, M.B. and M.A., will independently conduct the screening process. They will carefully review the titles and abstracts of eligible studies, eliminating any irrelevant articles. Full-text versions of relevant articles will then be retrieved for further evaluation. In the event of any discrepancies between the researchers, a third reviewer (Y.Z.) will be consulted, and a consensus will be reached through discussion. The reasons for excluding articles will be meticulously documented and outlined in a PRISMA flow diagram. Prior to commencing the screening process, calibration exercises will be conducted to ensure consistency among the researchers, thereby minimising potential inter-reviewer bias.

Inclusion and exclusion criteria

This systematic review will include studies employing either fully automated or semi-automated artificial intelligence (AI) for analysing cystoscopic images to detect bladder cancer. Both prospective and retrospective studies will be considered. The main comparisons will focus on evaluating sensitivities, specificities, positive predictive values (PPVs), negative predictive values (NPVs), and area under the curve (AUC) values. Patient cohorts may include individuals with suspected or confirmed bladder cancer cases, with histological findings serving as reference standards.

Excluded from analysis will be correspondence papers, ongoing studies, case reports, and conference abstracts. Additionally, non-English language articles, studies not utilising cystoscopy as the primary diagnostic modality, and those involving patients with a history of previous bladder cancer treatment will be excluded.

Data extraction (table of collection)

The data outlined in Table 1 will be gathered from all selected studies. Each researcher will independently conduct data extraction, consolidating the obtained information into a comprehensive datasheet. Any discrepancies in data extraction will be reviewed by a third evaluator, with the aim of reaching consensus for resolution.

If accessible, pertinent figures such as true positives, true negatives, false positives, false negatives, and their derived calculations will be extracted accordingly. If these figures are not explicitly provided, efforts will be made to compute them from available data. In cases where computation is unattainable, the authors of the respective paper will be contacted to provide the required data.

Table 1. Data collection items

Item No.	**Data Title**	**Data Type**
1	Year of publication	Study characteristic
2	Study authors	Study characteristic
3	Patient population	Demographics
4	Study size	Demographics
5	Cystoscopic images	Methodology
6	Histopathology results	Methodology
7	AI models utilised	Methodology

8	Definition for significant clinical disease	Methodology

AI: artificial intelligence.

Endpoints

The main endpoint of analysis will be the statistically significant quantification of accuracy when employing AI in bladder cancer detection during cystoscopy, aiming to assess its potential impact on clinical outcomes. Additional outcomes will encompass various parameters examining patient demographics.

Meta-analysis

Should an ample number of suitable studies be accessible, we will proceed with a meta-analysis to amalgamate a quantitative measure of AI performance in identifying bladder cancer from cystoscopic images. Initially, sensitivity and specificity values will be retrieved from studies, or if not accessible, computed from clinical data or solicited from authors. If a notable fraction of studies employ alternative metrics like F1-score or precision-recall AUC, these metrics will be acquired and scrutinised independently. The distributions of untransformed, logit, and double-arcsine transformed proportions will be compared and evaluated for normality using density plots and Shapiro-Wilk tests. The set of ratios most resembling a normal distribution will be selected for further analysis.

Heterogeneity and inter-study variation will be quantified using I2, and if statistically significant, a random-effects model will be employed for estimating the summary estimate. Leave-one-out analysis (LOO) and accompanying diagnostic plots will be used to identify influential studies after the model fits all relevant studies. Summary

proportions will be re-estimated based on the remaining studies after each study is removed. Studies exerting a statistically significant influence on the model will be identified as outliers and excluded. After excluding these studies, the model will then be refitted with a summary estimate comprising the remaining studies will be calculated to estimate the accuracy using AI in cystoscopy for bladder cancer detection. All data analysis and visualisation will be conducted using the R statistical environment with the "mvmeta" and "meta" packages.

Risk of bias in individual studies : This study will use the QUADAS-2 tool to assess the quality and the risk of bias within each of the included studies (17). The scoring system is split into four main sections: patient selection, index test, reference standard, and flow and timing. Within each section, signalling questions assessed the quality of the research methodology and results at the study level. As included studies utilises imaging data, the CLAIM-AI checklist will be used to assess imaging specific considerations such as the classification, image reconstruction, text analysis and workflow optimisation (10). Two reviewers (M.B. and M.A.) will independently engage in this procedure, with any discrepancies resolved through consensus. The bias assessment will analyse the appropriateness and dependability of the data utilised. This analysis will contribute to evidence synthesis and enhance transparency, potentially leading to the exclusion of articles of low quality or indicative of high bias levels. If included, relevant commentaries will be integrated into the discussion.

Ethics and Dissemination

The potential of AI to assist in diagnosing bladder cancer during cystoscopy remains to be thoroughly assessed. Our aim is to consolidate available information to elucidate AI's role in cystoscopy for bladder cancer diagnosis and outline its impact on clinical practice.

However, it is crucial to acknowledge several limitations. Firstly, given the novelty of AI in this domain, there may be a limited number of studies available for evaluation. Additionally, issues related to the diversity of reporting protocols due to regional variations may arise. Furthermore, the presence of different neural networks could potentially act as confounders due to variations in training datasets, leading to discrepancies in reported outcomes.

We intend to disseminate our findings through publication in a peer-reviewed journal. The synthesis of this data will contribute to a deeper understanding of current AI methodologies in bladder cancer diagnosis. Ultimately, this knowledge may facilitate the enhancement of existing systems, both supervised and unsupervised, to refine reporting protocols and improve bladder cancer diagnosis in the future.

Patient and public involvement

There will be no patient or public involvement in this study.

Trial status

- **-** Preliminary searches: Started
- - Piloting of the study selection process: Not started
- - Formal screening: Not started
- - Data extraction: Not started

- - Risk of bias assessment: Not started
- - Data analysis: Not started

Author Affiliations

[1]London North West University Healthcare NHS Trust, London, UK

[2]Imperial College London School of Medicine, London, UK

[3]Norfolk and Norwich University Hospital NHS Trust

[4]Imperial College London NHS Trust, London, UK

Email: Mohamed Baana: m.baana1@nhs.net; Murtada Arkwazi: murtada.arkwazi@nhs.net; Yi Zhao: yi.zhao18@imperial.ac.uk, Ojone Ofagbor: ojone.lawrence@nhs.net, Gaurika Bhardwaj: gaurika.bhardwaj@nhs.net, Mariam Lami: mariam.lami1@nhs.net, Eva Bolton: eva.bolton@nhs.net, Rakesh Heer: r.heer@imperial.ac.uk

Contributors The authors' contribution includes, but is not limited to, the following: MB and MA wrote the manuscript. YZ created the study concept. LM, EB and RH provided supervision and guidance, GB and OO checked the manuscript in its current form. YZ and MB are the guarantors of this work.

Funding Imperial College London Open Access Fund

Competing Interests None

Patient consent for publication Not required

Provenance and peer review Not commissioned; externally peer-reviewed

ORCID IDs

- Mohamed Baana: https://orcid.org/0009-0006-0384-9423
- Murtada Arkwazi: https://orcid.org/0009-0001-1344-9849
- Yi Zhao: https://orcid.org/0000-0002-4563-4344
- Ojone Ofagbor: https://orcid.org/0009-0004-5033-9050
- Gaurika Bhardwaj: https://orcid.org/0009-0008-7502-9920
- Mariam Lami: https://orcid.org/0000-0002-5485-8315
- Eva Bolton: https://orcid.org/0000-0002-1191-0849
- Rakesh Heer: https://orcid.org/0000-0003-1952-7462

References

(1) Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209–49.

(2) Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Int J Cancer. 2021 Apr 5;

(3) Witjes JA, Bruins HM, Cathomas R, Compérat EM, Cowan NC, Gakis G, et al. European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. European Urology. 2021 Jan 1;79(1):82–104.

(4) Russo GI, Sholklapper TN, Cocci A, Broggi G, Caltabiano R, Smith AB, et al. Performance of Narrow Band Imaging (NBI) and Photodynamic Diagnosis (PDD) Fluorescence Imaging Compared to White Light Cystoscopy (WLC) in Detecting Non-Muscle Invasive Bladder Cancer: A Systematic Review and Lesion-Level Diagnostic Meta-Analysis. Cancers. 2021 Jan;13(17):4378.

(5) Bochenek K, Aebisher D, Międzybrodzka A, Cieślar G, Kawczyk-Krupka A. Methods for bladder cancer diagnosis – The role of autofluorescence and photodynamic diagnosis. Photodiagnosis and Photodynamic Therapy. 2019 Sep 1;27:141–8.

(6) Chan EOT, Pradere B, Teoh JYC. The use of artificial intelligence for the diagnosis of bladder cancer: a review and perspectives. Current Opinion in Urology. 2021 Jul;31(4):397.

(7) Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017 Apr;69S:S36–40.

(8) Howard J. Artificial intelligence: Implications for the future of work. Am J Ind Med. 2019 Nov;62(11):917–26.

(9) Shkolyar E, Jia X, Chang TC, Trivedi D, Mach KE, Meng MQH, et al. Augmented Bladder Tumor Detection Using Deep Learning. European Urology. 2019 Dec 1;76(6):714–8.

(10) Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. Radiol Artif Intell. 2020 Mar;2(2):e200029.

(11) de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. npj Digit Med. 2022 Jan 10;5(1):1–13.

(12) Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health. 2021 Nov 1;3(11):e745–50.

(13) Ferro M, Falagario UG, Barone B, Maggi M, Crocetto F, Busetto GM, et al. Artificial Intelligence in the Advanced Diagnosis of Bladder Cancer-Comprehensive Literature Review and Future Advancement. Diagnostics (Basel). 2023 Jul 7;13(13):2308.

(14) Borhani S, Borhani R, Kajdacsy-Balla A. Artificial intelligence: A promising frontier in bladder cancer diagnosis and outcome prediction. Critical Reviews in Oncology/Hematology. 2022 Mar 1;171:103601.

(15) Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews. 2015 Jan 1;4(1):1.

(16) Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016 Dec 5;5(1):210.

(17) Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011 Oct 18;155(8):529–36.

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: A Protocol systematic review and meta-analysis

The search terms for this systematic review are as follows:

((artificial intelligence OR machine intelligence OR computational intelligence OR neural network OR deep learning OR machine learning OR reinforcement learning OR convolutional neural network OR artificial neural network OR recurrent neural network OR deep neural network OR intelligent model OR predictive model OR computational model OR algorithm OR AI OR NN OR DL OR ML OR CNN OR ANN OR RNN OR DNN) AND (random forests OR support vector machines OR nearest neighbours OR logistic regression) AND ((Bladder AND (cancer OR malignant OR transitional cell or urothelial OR squamous cell OR adenocarcinoma OR papillary OR carcinoma in situ OR non muscle invasive OR lesions OR disease)) AND (cystoscop* OR cystoscop* images OR flexible cystoscopy OR cystourethroscopy)) AND (diagnos* OR detect*)).

Searches will be conducted on the following databases:

PubMed, MEDLINE, Embase, and Cochrane databases

BMJ Open

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: A protocol for a systematic review and meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-089125.R2
Article Type:	Protocol
Date Submitted by the Author:	30-Aug-2024
Complete List of Authors:	BAANA, MOHAMED; Northwick Park Hospital, Urology Arkwazi, Murtada; London North West University Healthcare NHS Trust Zhao, Yi; Imperial College London School of Medicine Ofagbor, Ojone; Norwich University Hospital NHS Trust Bhardwaj, Gaurika; Imperial College London NHS Trust Lami, Mariam; Imperial College London NHS Trust Bolton, Eva; Imperial College London NHS Trust Heer, Rakesh; Imperial College London NHS Trust
Primary Subject Heading:	Urology
Secondary Subject Heading:	Surgery
Keywords:	Artificial Intelligence, UROLOGY, Diagnostic Imaging

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: A Protocol for a systematic review and meta-analysis

Mohamed Baana[1], Murtada Arkwazi[1], Yi Zhao[2], Ojone Ofagbor[3], Gaurika Bhardwaj[4], Mariam Lami[4], Eva Bolton[4], Rakesh Heer[4]

MB and MA are joint first authors

EB and RH are joint senior authors

[1]London North West University Healthcare NHS Trust, London, UK

[2]Imperial College London School of Medicine, London, UK

[3]Norfolk and Norwich University Hospital NHS Trust

[4]Imperial College London NHS Trust, London, UK

Correspondence to: YZ14017@ic.ac.uk

Imperial College London

Exhibition Rd, South Kensington, London SW7 2BU

YZ14017@ic.ac.uk

Ethics: No ethical considerations apply

Word count: 1754

ABSTRACT

Introduction: Cystoscopy has revolutionised the process of diagnosing bladder cancer, leading to better categorization of risk levels and more precise treatment plans. Nonetheless, concerns arise about the lack of uniformity among observers in predicting tumour stage and grade. To address these concerns, artificial intelligence (AI) is being incorporated into clinical settings to aid in the analysis of diagnostic and therapeutic images. The subsequent report outlines a systematic review and meta-analysis protocol aimed at evaluating the effectiveness of AI in predicting bladder cancer based on cystoscopic images.

Methods and Analysis: Our systematic search will utilise databases including Pubmed, MEDLINE, Embase, and Cochrane. The articles published between May 2015 and April 2024 will be eligible for inclusion. For articles to be considered, they must employ AI for analysis of cystoscopic images to identify bladder cancer, present original data and be written in English. The protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) 2015 checklist. Quality and bias risk across chosen studies will be evaluated using the QUADAS-2 score.

Ethics and dissemination: Ethical clearance won't be necessary for conducting this systematic review since results will be disseminated through peer-reviewed publications and presentations at both national and international conferences.

PROSPERO registration number: CRD42024528345

Strengths and limitations of this study

- This protocol adheres to PRISMA guidelines and will incorporate subgroup and sensitivity analyses to further explore the variability among the studies included.
- AI-specific metrics like the F1-score and precision-recall AUC will be utilised to address limitations inherent in traditional pooled analysis, such as the impact of imbalanced classes.
- Given the novelty of AI technology in cystoscopy, long-term data regarding its impact on clinical outcomes may be scarce.
- Limitations on language and the exclusive use of cystoscopy may result in a limited number of eligible studies for inclusion

BACKGROUND

Bladder cancer is the commonest cancer of the urinary tract, and ranks as the ninth most common cancer worldwide according to the World Health Organization's 2020 report(1,2). The gold standard for diagnosing bladder cancer remains cystoscopy, as advocated by both the National Institute for Health and Care Excellence (NICE) and the European Association of Urology (EAU) guidelines(3). While white light cystoscopy (WLC) is the conventional method widely employed, it potentially overlooks up to 20% of lesions. Furthermore, a recent systematic review and meta-analysis demonstrated that conventional WLC exhibits low diagnostic sensitivity compared to alternative modalities like blue light cystoscopy (BLC)(4). Despite the superior detection rates of bladder cancer associated with these alternative methods recommended by EAU, their adoption remains limited, likely due to their higher initial costs and limited availability(4). Transurethral resection of bladder tumour (TURBT) with WLC remains fundamental for diagnosing and treating non-muscle invasive bladder cancer (NMIBC) which accounts for around 75% of bladder tumours at the time of diagnosis(5).

Artificial intelligence (AI) now boasts a remarkable ability to accurately recognise images. AI offers a promising solution to improve the diagnosis of bladder cancer during cystoscopy(6). Artificial intelligence broadly describes the modelling of intelligent behaviour by use of a computer model(7). Deep learning is a subset of artificial intelligence which more specifically positions AI within the context of medical imaging(8). Augmented cystoscopy employing deep learning holds promise in also enhancing tumour localisation, intraoperative navigation, and surgical resection of bladder cancer during TURBT(9).

While Artificial intelligence shows promise in the diagnosis of bladder cancer using cystoscopy, several limitations to the deployment of this technology need to be addressed. Given the cystoscopy imaging data used, such as WLC and BLC, studies are strongly encouraged to follow the Checklist for Artificial Intelligence in Medical Imaging (CLAIM)(10). This is to encourage the reproducibility of AI models in development and forwarding the collaboration of research groups in external validation of their models. The checklist also provides a focus on the use of radiomic features as well as computer-aided diagnosis of imaging data. In order to be better suited for clinical development, models should be explainable in their decision-making process which may be currently under reported(11,12).

Existing reviews on this topic have provided a robust summary to the feasibility of the application of AI in cystoscopy(13,14). But further investigation of the reported studies with a goal of clinical deployment should be conducted next. Therefore, this systematic review and meta-analysis seeks to outline the precision of artificial intelligence (AI) in forecasting bladder cancer based on cystoscopic images and evaluate its potential influence on patient clinical outcomes.

METHODS AND ANALYSIS:

This systematic review protocol adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 checklist(15). This study has been prospectively registered with the PROSPERO review database (CRD42024528345), and all methodologies detailed herein have been established prior to implementation. The statistical analysis will focus on evaluating the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area-under-the-curve (AUC) associated with the application of artificial intelligence (AI) in detecting bladder cancer during cystoscopy, along with its impact on clinical outcomes. These parameters will be derived through comprehensive analysis and thematic synthesis of included studies. Pooled sensitivities and specificities across the studies will be determined before calculating PPV and NPV values. AI-specific matrices such as F1-scores and Precision-Recall AUC will also be investigated. This review intends to start on the 1st October 2024 and end on the 30th April 2025.

Search Methodology

A comprehensive search will be conducted across multiple databases including PubMed, MEDLINE, Embase, and the Cochrane Library. The search strategy will incorporate medical subject heading (MeSH) terms and free text combined with appropriate Boolean operators. Articles from May 2015 to April 2024 will be included to ensure a thorough retrieval of relevant evidence. The search will encompass the following key terms: "bladder", "cancer", "diagnosis", “cystoscopy”, “cystoscopic images”, "artificial intelligence", and "deep learning". The complete search strategy is outlined in Supplementary file 1. To streamline the initial screening phase, we will be utilising Rayyan, a semi-automated tool crafted to enhance the efficiency and precision of systematic review (16). All eligible articles identified in the initial search will be imported into Rayyan. Additionally, a manual examination of references cited in all included articles will be conducted to uncover any additional pertinent literature not captured by the initial search strategy. In instances where data is lacking or unclear, corresponding authors will be contacted for clarification.

Study selection and data extraction

Two researchers, M.B. and M.A., will independently conduct the screening process. They will carefully review the titles and abstracts of eligible studies, eliminating any irrelevant articles. Full-text versions of relevant articles will then be retrieved for further evaluation. In the event of any discrepancies between the researchers, a third reviewer (Y.Z.) will be consulted, and a consensus will be reached through discussion. The reasons for excluding articles will be meticulously documented and outlined in a PRISMA flow diagram. Prior to commencing the screening process, calibration exercises will be conducted to ensure consistency among the researchers, thereby minimising potential inter-reviewer bias.

Inclusion and exclusion criteria

This systematic review will include studies employing either fully automated or semi-automated artificial intelligence (AI) for analysing cystoscopic images to detect bladder cancer. Both prospective and retrospective studies will be considered. The main comparisons will focus on evaluating sensitivities, specificities, positive predictive values (PPVs), negative predictive values (NPVs), and area under the curve (AUC) values. Patient cohorts may include individuals with suspected or confirmed bladder cancer cases, with histological findings serving as reference standards.

Excluded from analysis will be correspondence papers, ongoing studies, case reports, and conference abstracts. Additionally, non-English language articles, studies not utilising cystoscopy as the primary diagnostic modality, and those involving patients with a history of previous bladder cancer treatment will be excluded.

Data extraction (table of collection)

The data outlined in Table 1 will be gathered from all selected studies. Each researcher will independently conduct data extraction, consolidating the obtained information into a comprehensive datasheet. Any discrepancies in data extraction will be reviewed by a third evaluator, with the aim of reaching consensus for resolution.

If accessible, pertinent figures such as true positives, true negatives, false positives, false negatives, and their derived calculations will be extracted accordingly. If these figures are not explicitly provided, efforts will be made to compute them from available data. In cases where computation is unattainable, the authors of the respective paper will be contacted to provide the required data.

Table 1. Data collection items

Item No.	Data Title	Data Type
1	Year of publication	Study characteristic
2	Study authors	Study characteristic
3	Patient population	Demographics
4	Study size	Demographics
5	Cystoscopic images	Methodology
6	Histopathology results	Methodology
7	AI models utilised	Methodology

8	Definition for significant clinical disease	Methodology

AI: artificial intelligence.

Endpoints

The main endpoint of analysis will be the statistically significant quantification of accuracy when employing AI in bladder cancer detection during cystoscopy, aiming to assess its potential impact on clinical outcomes. Additional outcomes will encompass various parameters examining patient demographics.

Meta-analysis

Should an ample number of suitable studies be accessible, we will proceed with a meta-analysis to amalgamate a quantitative measure of AI performance in identifying bladder cancer from cystoscopic images. Initially, sensitivity and specificity values will be retrieved from studies, or if not accessible, computed from clinical data or solicited from authors. If a notable fraction of studies employ alternative metrics like F1-score or precision-recall AUC, these metrics will be acquired and scrutinised independently. The distributions of untransformed, logit, and double-arcsine transformed proportions will be compared and evaluated for normality using density plots and Shapiro-Wilk tests. The set of ratios most resembling a normal distribution will be selected for further analysis.

Heterogeneity and inter-study variation will be quantified using I2, and if statistically significant, a random-effects model will be employed for estimating the summary estimate. Leave-one-out analysis (LOO) and accompanying diagnostic plots will be used to identify influential studies after the model fits all relevant studies. Summary

proportions will be re-estimated based on the remaining studies after each study is removed. Studies exerting a statistically significant influence on the model will be identified as outliers and excluded. After excluding these studies, the model will then be refitted with a summary estimate comprising the remaining studies will be calculated to estimate the accuracy using AI in cystoscopy for bladder cancer detection. All data analysis and visualisation will be conducted using the R statistical environment with the "mvmeta" and "meta" packages.

Risk of bias in individual studies: This study will use the QUADAS-2 tool to assess the quality and the risk of bias within each of the included studies (17). The scoring system is split into four main sections: patient selection, index test, reference standard, and flow and timing. Within each section, signalling questions assessed the quality of the research methodology and results at the study level. As included studies utilises imaging data, the CLAIM-AI checklist will be used to assess imaging specific considerations such as the classification, image reconstruction, text analysis and workflow optimisation (10). Two reviewers (M.B. and M.A.) will independently engage in this procedure, with any discrepancies resolved through consensus. The bias assessment will analyse the appropriateness and dependability of the data utilised. This analysis will contribute to evidence synthesis and enhance transparency, potentially leading to the exclusion of articles of low quality or indicative of high bias levels. If included, relevant commentaries will be integrated into the discussion.

Ethics and Dissemination

The potential of AI to assist in diagnosing bladder cancer during cystoscopy remains to be thoroughly assessed. Our aim is to consolidate available information to elucidate AI's role in cystoscopy for bladder cancer diagnosis and outline its impact on clinical practice.

However, it is crucial to acknowledge several limitations. Firstly, given the novelty of AI in this domain, there may be a limited number of studies available for evaluation. Additionally, issues related to the diversity of reporting protocols due to regional variations may arise. Furthermore, the presence of different neural networks could potentially act as confounders due to variations in training datasets, leading to discrepancies in reported outcomes.

We intend to disseminate our findings through publication in a peer-reviewed journal. The synthesis of this data will contribute to a deeper understanding of current AI methodologies in bladder cancer diagnosis. Ultimately, this knowledge may facilitate the enhancement of existing systems, both supervised and unsupervised, to refine reporting protocols and improve bladder cancer diagnosis in the future.

Patient and public involvement

There will be no patient or public involvement in this study.

Trial status

- **-** Preliminary searches: Started
- - Piloting of the study selection process: Not started
- - Formal screening: Not started
- - Data extraction: Not started

- - Risk of bias assessment: Not started
- - Data analysis: Not started

Author Affiliations

[1]London North West University Healthcare NHS Trust, London, UK

[2]Imperial College London School of Medicine, London, UK

[3]Norfolk and Norwich University Hospital NHS Trust

[4]Imperial College London NHS Trust, London, UK

Email: Mohamed Baana: m.baana1@nhs.net; Murtada Arkwazi: murtada.arkwazi@nhs.net; Yi Zhao: yi.zhao18@imperial.ac.uk, Ojone Ofagbor: ojone.lawrence@nhs.net, Gaurika Bhardwaj: gaurika.bhardwaj@nhs.net, Mariam Lami: mariam.lami1@nhs.net, Eva Bolton: eva.bolton@nhs.net, Rakesh Heer: r.heer@imperial.ac.uk

Contributors The authors' contribution includes, but is not limited to, the following: MB and MA wrote the manuscript. YZ created the study concept. LM, EB and RH provided supervision and guidance, GB and OO checked the manuscript in its current form. YZ and MB are the guarantors of this work.

Funding Imperial College London Open Access Fund

Competing Interests None

Patient consent for publication Not required

Provenance and peer review Not commissioned; externally peer-reviewed

ORCID IDs

- Mohamed Baana: https://orcid.org/0009-0006-0384-9423
- Murtada Arkwazi: https://orcid.org/0009-0001-1344-9849
- Yi Zhao: https://orcid.org/0000-0002-4563-4344
- Ojone Ofagbor: https://orcid.org/0009-0004-5033-9050
- Gaurika Bhardwaj: https://orcid.org/0009-0008-7502-9920
- Mariam Lami: https://orcid.org/0000-0002-5485-8315
- Eva Bolton: https://orcid.org/0000-0002-1191-0849
- Rakesh Heer: https://orcid.org/0000-0003-1952-7462

References

(1) Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209–49.

(2) Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Int J Cancer. 2021 Apr 5;

(3) Witjes JA, Bruins HM, Cathomas R, Compérat EM, Cowan NC, Gakis G, et al. European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. European Urology. 2021 Jan 1;79(1):82–104.

(4) Russo GI, Sholklapper TN, Cocci A, Broggi G, Caltabiano R, Smith AB, et al. Performance of Narrow Band Imaging (NBI) and Photodynamic Diagnosis (PDD) Fluorescence Imaging Compared to White Light Cystoscopy (WLC) in Detecting Non-Muscle Invasive Bladder Cancer: A Systematic Review and Lesion-Level Diagnostic Meta-Analysis. Cancers. 2021 Jan;13(17):4378.

(5) Bochenek K, Aebisher D, Międzybrodzka A, Cieślar G, Kawczyk-Krupka A. Methods for bladder cancer diagnosis – The role of autofluorescence and photodynamic diagnosis. Photodiagnosis and Photodynamic Therapy. 2019 Sep 1;27:141–8.

(6) Chan EOT, Pradere B, Teoh JYC. The use of artificial intelligence for the diagnosis of bladder cancer: a review and perspectives. Current Opinion in Urology. 2021 Jul;31(4):397.

(7) Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017 Apr;69S:S36–40.

(8) Howard J. Artificial intelligence: Implications for the future of work. Am J Ind Med. 2019 Nov;62(11):917–26.

(9) Shkolyar E, Jia X, Chang TC, Trivedi D, Mach KE, Meng MQH, et al. Augmented Bladder Tumor Detection Using Deep Learning. European Urology. 2019 Dec 1;76(6):714–8.

(10) Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. Radiol Artif Intell. 2020 Mar;2(2):e200029.

(11) de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. npj Digit Med. 2022 Jan 10;5(1):1–13.

(12) Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health. 2021 Nov 1;3(11):e745–50.

(13) Ferro M, Falagario UG, Barone B, Maggi M, Crocetto F, Busetto GM, et al. Artificial Intelligence in the Advanced Diagnosis of Bladder Cancer-Comprehensive Literature Review and Future Advancement. Diagnostics (Basel). 2023 Jul 7;13(13):2308.

(14) Borhani S, Borhani R, Kajdacsy-Balla A. Artificial intelligence: A promising frontier in bladder cancer diagnosis and outcome prediction. Critical Reviews in Oncology/Hematology. 2022 Mar 1;171:103601.

(15) Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews. 2015 Jan 1;4(1):1.

(16) Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016 Dec 5;5(1):210.

(17) Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011 Oct 18;155(8):529–36.

Utilising artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: A Protocol systematic review and meta-analysis

The search terms for this systematic review are as follows:

((artificial intelligence OR machine intelligence OR computational intelligence OR neural network OR deep learning OR machine learning OR reinforcement learning OR convolutional neural network OR artificial neural network OR recurrent neural network OR deep neural network OR intelligent model OR predictive model OR computational model OR algorithm OR AI OR NN OR DL OR ML OR CNN OR ANN OR RNN OR DNN) AND (random forests OR support vector machines OR nearest neighbours OR logistic regression) AND ((Bladder AND (cancer OR malignant OR transitional cell or urothelial OR squamous cell OR adenocarcinoma OR papillary OR carcinoma in situ OR non muscle invasive OR lesions OR disease)) AND (cystoscop* OR cystoscop* images OR flexible cystoscopy OR cystourethroscopy)) AND (diagnos* OR detect*)).

Searches will be conducted on the following databases:

PubMed, MEDLINE, Embase, and Cochrane databases