Open access Protocol

BMJ Open

Using 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]

To cite: Baana M, Arkwazi M, Zhao Y, *et al*. Using artificial intelligence for bladder cancer detection during cystoscopy and its impact on clinical outcomes: a protocol for a systematic review and meta-analysis. *BMJ Open* 2024;**14**:e089125. doi:10.1136/bmjopen-2024-089125

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2024-089125).

OO and GB contributed equally.

MB and MA are joint first authors.
EB and RH are joint senior authors.

Received 22 May 2024
Accepted 26 September 2024

For numbered affiliations see end of article.

Correspondence to
Dr Mohamed Baana; m.baana1@nhs.net

ABSTRACT

Introduction Cystoscopy has revolutionised the process of diagnosing bladder cancer leading to better categorisation 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 use 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 2015 checklist. Quality and bias risk across chosen studies will be evaluated using the Quality Assessment of Diagnostic Accuracy Studies-2 score.

Ethics and dissemination Ethical clearance will not 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 Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines and will incorporate subgroup and sensitivity analyses to further explore the variability among the studies included.
- ⇒ Artificial intelligence (AI)-specific metrics like the F1-score and precision-recall area under the curve will be used 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 most common cancer of the urinary tract and ranks as the ninth most common cancer worldwide according to the WHO'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 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 with 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 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] AI broadly describes the modelling of intelligent behaviour by use of a computer model.[7] Deep learning is a subset of AI 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 AI 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 to forward the collaboration of research groups in the 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 of the feasibility of the application of AI in cystoscopy.[13 14] However, 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 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 and all methodologies detailed here 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 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 1 October 2024 and end on the 30 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 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 online supplemental file 1. To streamline the initial screening phase, we will be using 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, MB and MA, 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 (YZ) 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 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, PPVs, NPVs and 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 using 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 a consensus for resolution.

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 used	Methodology
8	Definition for significant clinical disease	Methodology

AI, artificial intelligence.

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.

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 interstudy 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 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 Quality Assessment of Diagnostic Accuracy Studies-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 use 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 (MB and MA) will independently engage in this procedure with any discrepancies resolved through consensus. The bias assessment will analyse the appropriateness and dependability of the data used. 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. First, 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 data sets 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, Harrow, UK
[2]Imperial College London School of Medicine, London, UK
[3]Norwich University Hospital NHS Trust, Norfolk, UK
[4]Imperial College London NHS Trust, London, UK

Contributors The authors' contribution includes but is not limited to the following: MB and MA wrote the manuscript. YZ created the study concept. ML, 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 Funding for the Open Access publication charges for this article was provided by the Imperial Open Access Fund.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iD
Mohamed Baana http://orcid.org/0009-0006-0384-9423

REFERENCES

1 Sung H, Ferlay J, Siegel RL, *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J Clinicians* 2021;71:209–49.
2 Ferlay J, Colombet M, Soerjomataram I, *et al*. Cancer statistics for the year 2020: An overview. *Intl Journal of Cancer* 2021;149:778–89.
3 Witjes JA, Bruins HM, Cathomas R, *et al*. European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. *Eur Urol* 2021;79:82–104.
4 Russo GI, Sholklapper TN, Cocci A, *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 (Basel)* 2021;13:4378.
5 Bochenek K, Aebisher D, Międzybrodzka A, *et al*. Methods for bladder cancer diagnosis – The role of autofluorescence and photodynamic diagnosis. *Photodiagnosis Photodyn Ther* 2019;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. *Curr Opin Urol* 2021;31:397–403.
7 Hamet P, Tremblay J. Artificial intelligence in medicine. *Metab Clin Exp* 2017;69:S36–40.
8 Howard J. Artificial intelligence: Implications for the future of work. *American J Industrial Med* 2019;62:917–26.
9 Shkolyar E, Jia X, Chang TC, *et al*. Augmented Bladder Tumor Detection Using Deep Learning. *Eur Urol* 2019;76:714–8.
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;2:e200029.
11 de Hond AAH, Leeuwenberg AM, Hooft L, *et al*. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ Digit Med* 2022;5:2.
12 Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Dig Health* 2021;3:e745–50.
13 Ferro M, Falagario UG, Barone B, *et al*. Artificial Intelligence in the Advanced Diagnosis of Bladder Cancer-Comprehensive. *Lit Rev Future Advanc Diagn (Basel)* 2023;13:2308.
14 Borhani S, Borhani R, Kajdacsy-Balla A. Artificial intelligence: A promising frontier in bladder cancer diagnosis and outcome prediction. *Crit Rev Oncol Hematol* 2022;171:103601.
15 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
16 Ouzzani M, Hammady H, Fedorowicz Z, *et al*. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016;5:210.
17 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;155:529.