Explainable artificial intelligence (XAI): closing the gap between image analysis and navigation in complex invasive diagnostic procedures

S. O'Sullivan[1] · M. Janssen[1] · Andreas Holzinger[2,3] · Nathalie Nevejans[4] · O. Eminaga[5,6] · C. P. Meyer[7] · Arkadiusz Miernik[8,9]

Abstract

Literature review Cystoscopy is the gold standard for initial macroscopic assessments of the human urinary bladder to rule out (or diagnose) bladder cancer (BCa). Despite having guidelines, cystoscopic findings are diverse and often challenging to classify. The extent of the false negatives and false positives in cystoscopic diagnosis is currently unknown. We suspect that there is a certain degree of under-diagnosis (like the failure to detect malignant tumours) and over-diagnosis (e.g. sending the patient for unnecessary transurethral resection of bladder tumors with anesthesia) that put the patient at risk.

Conclusions XAI robot-assisted cystoscopes would help to overcome the risks/flaws of conventional cystoscopy. Cystoscopy is considered a less life-threatening starting point for automation than open surgical procedures. Semi-autonomous cystoscopy requires standards and cystoscopy is a good procedure to establish a model that can then be exported/copied to other procedures of endoscopy and surgery. Standards also define the automation levels—an issue for medical product law. These cystoscopy skills do not give full autonomy to the machine, and represent a surgical parallel to 'Autonomous Driving' (where a standard requires a human supervisor to remain in the 'vehicle'). Here in robotic cystoscopy, a human supervisor remains bedside in the 'operating room' as a 'human-in-the-loop' in order to safeguard patients. The urologists will be able to delegate personal- and time-consuming cystoscopy to a specialised nurse. The result of automated diagnostic cystoscopy is a short video (with pre-processed photos from the video), which are then reviewed by the urologists at a more convenient time.

Keywords Cystoscopy · Bladder cancer detection · Bladder cancer diagnosis · Robotic endoscopy · Autonomous surgery · Autonomous endoscopy

Introduction

AI robotics has enormous potential to assist health professionals in many daily tasks [1]. Especially, AI methods from the domain of the so-called deep learning [2]. These are deep neural networks that are very successful and can practically exceed human performance [3]. Unfortunately, these most successful models have a serious drawback. They are high dimensional, very non-linear, and therefore, very complex that they are no longer interpretable by human experts

✉ S. O'Sullivan
sosullivan810@gmail.com

1 Department of Urology, University Hospital of Münster (UKM), Muenster, Germany

2 Human-Centered AI Lab, Institute for Medical Informatics/Statistics, Medical University of Graz, Graz, Austria

3 xAI Lab, Alberta Machine Intelligence Institute, University of Alberta, Edmonton, Canada

4 AI Responsible Chair, Research Center in Law, Ethics and Procedures, Faculty of Law of Douai, University of Artois, Arras, France

5 Department of Urology, Stanford University School of Medicine, Stanford, CA, USA

6 Center for Artificial Intelligence in Medicine and Imaging, Stanford University School of Medicine, Stanford, CA, USA

7 Urology Clinic, Ruhr-University of Bochum, Bochum, Germany

8 Department of Urology, Faculty of Medicine, University of Freiburg-Medical Centre, Freiburg, Germany

9 RaVeNNA 4Pi-Consortium of the German Federal Ministry of Education and Research (BMBF), Freiburg, Germany

[4]. However, in medicine, transparency, re-traceability and explainability are essential criteria and even mandatory, both for legal reasons [5] and for ethical reasons [6–10]. It is often that the question of why is even more important than the actual AI result. The area of computer science that addresses the implementation of transparency, re-traceability and comprehensibility of such so-called black-box approaches is called explainable artificial intelligence (XAI). Currently, this is becoming more and more important, because the quality requirements of medical AI solutions are increasing accordingly and the human expert must be able to explain and understand why a machine decision has been made, and must not rely on the usual parameters such as accuracy and precision alone [11].

Translated to urology, one field of action would be an invasive diagnostic procedure such as cystoscopy. Cystoscopy is the Gold Standard for initial macroscopic assessments of the human urinary bladder to rule out (or diagnose) bladder cancer (BCa). Around 7% of all newly diagnosed cancers are BCa in the United States of America (USA), making it the fourth most common malignancy and among the top ten lethal cancers in men [12]. Only 4.7% of all newly diagnosed cancers are BCa in Germany, accounting for 3.2% of all cancer-related deaths [13].

Bladder cancer is the 10th most common form of cancer worldwide, with an estimated 549,000 new cases and 200,000 deaths in 2018 [14]. Bladder cancer is more common in men than in women, with respective incidence and mortality rates of 9.6 and 3.2 per 100,000 in men: about 4 times that of women globally. The disease ranks higher among men, in whom it is the 6th most common cancer and 9th leading cause of cancer death. For the Age-Standardised Rate per 100,000 population in 2018, Germany had a bladder cancer rate of 26.4 in men and a 6.3 in women [14].

Bladder cancer has two entities: non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive bladder cancer (MIBC) (see Fig. 1). Its notably high recurrence rates range from 50 to 70% at 20 years, and put patients at risk for MIBC [15]. Periodic cystoscopic examinations thus remain the cornerstone for patients' follow-ups to detect early recurrences and lower the risk of progression.

NMIBC represents 75% of primary diagnoses and is characterized by frequent recurrence [16]. Bladder cancer remains a heterogeneous disease with varying pathology, molecular background, diagnostic options, and subsequent therapies for the individual patient. The improvement of bladder cancer outcomes remains the goal of future research.

Although it is such an invasive procedure, cystoscopy of the urinary bladder plays a critical role in the diagnosis and follow-up of patients with bladder cancer. Non-invasive biomarkers like nuclear matrix protein 22 (NMP22) or human complement factor H-related protein (BTA TRAK) from urine (or others from blood samples) have failed to replace
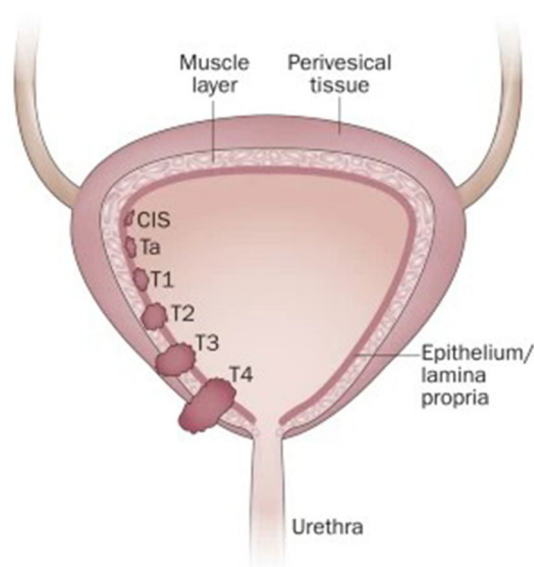

Fig. 1 Diagram of 'bladder cancer staging', starting with carcinoma in situ (CIS), then Ta, T1, T2, T3 and T4. Source: [16]

cystoscopy in early detection and follow-up [EAU guidelines on non-muscle-invasive bladder cancer (NMIBC)] [17]. In fact, they may only enhance the Gold Standard cystoscopy [18].

Molecular diagnostics are making rapid progress. Competition to AI could come in the form of non-invasive biomarkers such as Oncoprotein Keratin 17 (K17) [19]. K17 is highly sensitive and specific for detecting urogenital cancer recurrence in the urogenital tract during treatment follow up. However, new biomarkers like K17 still do not rule out cystoscopy, or rule in TURBT, overall.

However, conventional cystoscopy still has several shortcomings, as it is especially difficult to detect flat or small lesions like carcinomas in situ (Cis) via standard white light cystoscopy (WLC) [20]. The latest data suggest that early and higher recurrence rates after transurethral resection (TUR) for BCa may be due to lesions going undetected during initial cystoscopy [21, 22].

Several methods to enhance optical techniques were introduced in the last few decades, and most have proven to improve initial detection rates and operative outcomes after therapy [transurethral resection of bladder tumours (TUR-BT)] of BCa (see Fig. 2). For example, hexaminolevulinate (HAL) TUR-B [photodynamic diagnosis (PDD)], an adjunct to WLC which involves fluorescence to localise abnormal tissue. This method proved to achieve higher detection and better resection rates due to the enhanced visualisation of malignant lesions during TURBT than cystoscopy and white light resection alone [23].

In some countries, the vast majority of cystoscopies (e.g. over 90% in Germany) are performed with white light
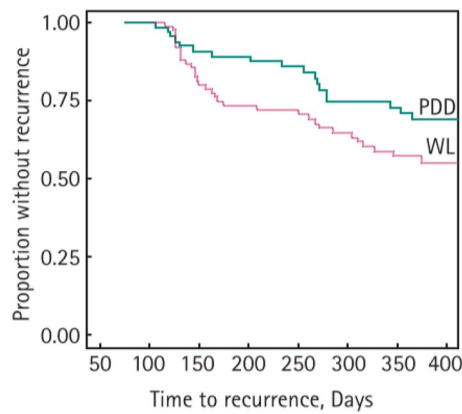
Fig. 2 Kaplan–Meier estimates of 12-month recurrence-free survival stratified by treatment. WL TUR-B, $n=77$, or HAL TUR-B (PDD), $n=68$; $P=0.02$. Source: [21]

cystoscopy (WLC) which can make it difficult to view certain lesions. Different technical improvements also include blue-light cystoscopy (BLC) and narrow-band imaging (NBI) [24, 25] (see Fig. 3).

The aforementioned technical improvements require a human professional to do the navigating and interpreting of what the cystoscopy reveals. The interpretation of such optical information would be one field for AI methods that could rely on convolutional neural networks (CNNs), machine learning (ML), or deep learning (DL) [26]. At present, the RaVeNNA-4p1 consortium is developing a digital platform with 4PI Endo-imaging for 3D-reconstruction, semantic segmentation and visualisation of the bladder for BCa.

The consortium focusses on the multi-class segmentation of cystoscopic images using DL. This enables the precise segmentation of classes of interest such as bladder tumours [26]. In addition, the RaVeNNA-4p1 consortium trained CNNs with confirmed images of bladder cancers vs. healthy urothelium. Their output revealed up to 93.0% sensitivity and up to 83.7% specificity differentiating

between tumourous/malignant and healthy urothelium; the neural network was pre-trained by feeding it 1.2 million images from the ImageNet Dataset [27] (Imagenet: a large-scale hierarchical image database). Several CNN models were evaluated by Eminaga et al. who found that the Xception-based model achieved the highest accuracy score (99.52%) [28].

Besides optical and imaged-based advancements, all these introduced methods assume that the cystoscopy itself was a thorough scan. They also assume that every millimeter of urothelium in the bladder was recorded. At present, cystoscopic navigation is limited, especially for flexible cystoscopy. We are convinced that better navigation will lead to automation, and that automation will ensure more reproducibility—essential during follow-up for non-muscle-invasive bladder cancer patients.

Methods

The literature search was performed on the following databases: PubMed/MEDLINE, Scopus, dblp, SSRN, Academia.edu, ResearchGate, and eur-lex.europa.eu. The review focused on articles that were published between the years of 2007 and 2021. All available publications were analyzed and summarized herein after an interdisciplinary collaborative review process. The search used the following keywords: cystoscopy, blue-light cystoscopy, narrow-band imaging, endoscopic documentation, endo-imaging, semantic segmentation, instance segmentation, visualization, non-invasive biomarkers, bladder cancer detection, bladder cancer diagnosis, explainable artificial intelligence, black box, robotic endoscopy, autonomous endoscopy, autonomous surgery, human rights, high-risk AI systems, fundamental rights, provider, user, conformity, compliance, accountability, responsibility, liability, human control.
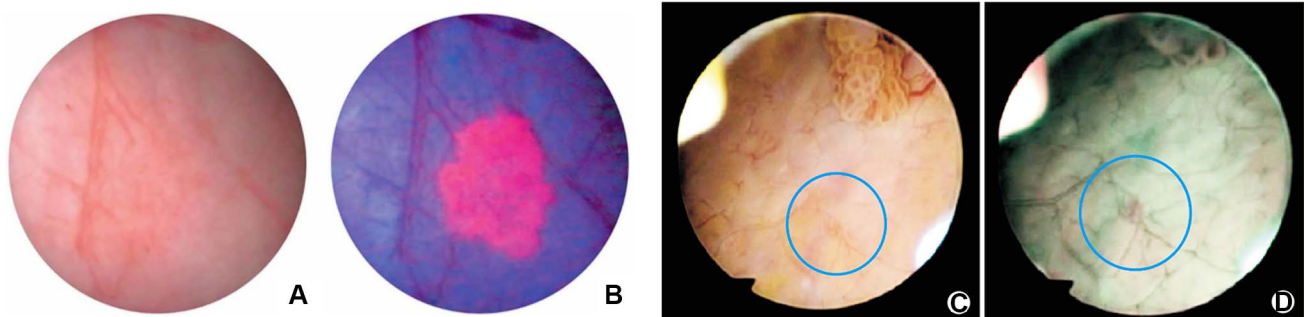

Fig. 3 Comparison of WLC with BLC and NBI. **A** Bladder cancer tumour viewed with WLC. **B** Same bladder cancer tumour as in A, but viewed with BLC (Cysview™), which reveals a lesion in pink against the BL (Source: [24]). **C** Ta cancer growth viewed with WLC. **D** Same Ta cancer growth as in C, but viewed with NBI (Source: [25])

Urethrocystoscopy for diagnosis and follow-up of bladder diseases

Since the German urologist Maximilian Nitze introduced cystoscopy to urology, it has improved continuously, driven by the need to enhance patient comfort and increase diagnostic accuracy (see Fig. 4). Urethrocystoscopy is one of the cornerstones for diagnosis and the follow-up of patients with urinary bladder disorders [15, 17]. Cytoscopic findings significantly help us explain the cause of symptoms and draw correlations with different bladder illnesses or systematic diseases. These patients may have visited the urologist after they noticed haematuria (i.e., blood in the urine).

Being so essential to detecting and identifying bladder tumours, guidelines have been made available outlining how a cystoscopy is performed and documented [15, 17]. These guidelines help avoid the severe consequences of failing to identify malignant lesions. Indeed, in follow-up, patients must undergo frequent cystoscopies (e.g. every 3 months for a 2-year period in high-risk tumour cases). About a million cystoscopies are performed every year in the USA alone [29].

In the surveillance setting, the enhanced flexible cystoscope [Federal Drug Administration (FDA) approved in 2018] revealed improved detection rates. However, it remains an invasive and sometimes painful procedure [30]. Several distraction methods have been studied (e.g. hand-holding, playing music, speaking, etc.) in an attempt to minimise patient discomfort and ultimately increase acceptance rates, especially for follow-up examinations of non-muscle invasive bladder cancer (NMIBC) [31]. This also illustrates the need for a more comfortable procedure, as this procedure must be repeated frequently with regard to NMIBC, to detect recurrence and progression [32].

Cystoscopy's technical situation

Despite having guidelines [15, 17], cystoscopic findings are diverse and often challenging to classify [28]. Arriving at an accurate diagnosis is still largely dependent on the examiner's experience, and the inter-observer variability is therefore wide [33]. Often in training, some spots on the bladder wall go undetected because of improper manipulation, especially during flexible cystoscopy. Since cystoscopy began, the difficulty of training junior urologists on how to use a cystoscope is a major point, as visually identifying a tumour within the bladder is so challenging and often imprecise because of various patient–surgeon- or cancer-specific factors [34]. Cystoscopy is usually performed as an outpatient procedure with the patient awake during the procedure. Patients often exhibit low tolerance for this procedure, and their tolerance correlates with the urologist's conduct during the cystoscopy, be the cystoscope flexible or rigid. The cystoscope is rotated slowly to inspect the bladder sidewalls and bladder dome. The urologist can also apply manual pressure to the lower abdominal wall, just superior to the pubic bone, to visualize the anterior bladder neck and dome.

After lubrication, a rigid or flexible endoscope of 16–25 Fr is inserted via the female or male urethra into the bladder. Rigid telescopes with different angles (from 0° to 30° to 70°) can be used. Rigid cystoscopy is an often painful procedure, especially for male patients (as they have a longer urethra and/or enlarged prostate gland), which is why flexible
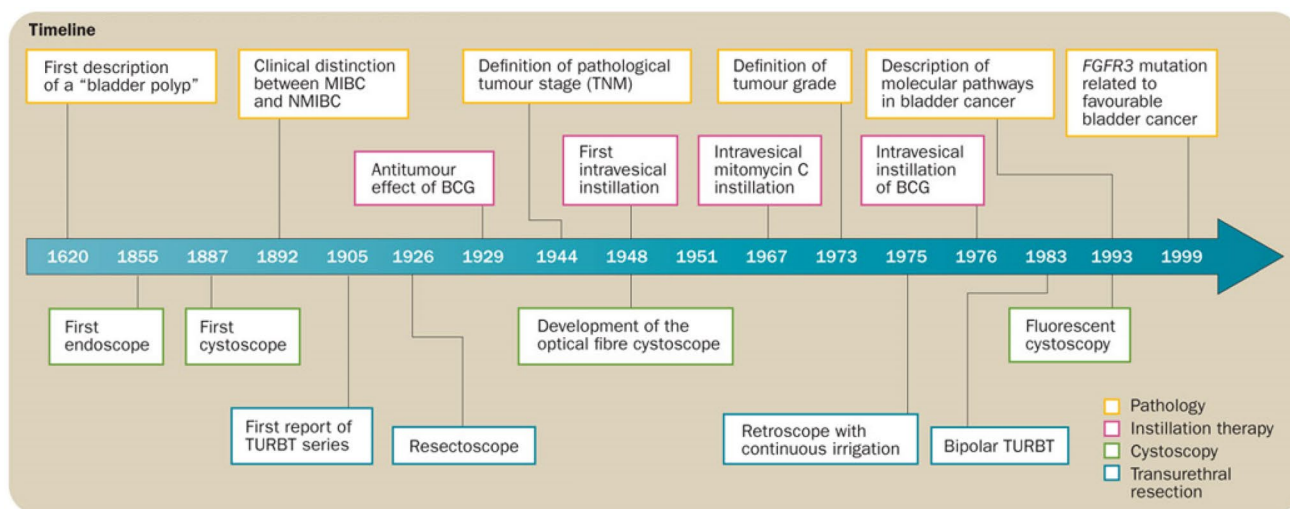

Fig. 4 Timeline for the R&D of MIBC and NMIBC. Source: [16]. Diagram abbreviations: TNM stands for Tumour, Node, and Metastasis. BCG stands for Bacillus Calmette–Guèrin (a type of immuno-therapy drug). FGFR3 stands for fibroblast growth factor receptor 3 (a protein coding gene). TUR-BT stands for transurethral resection of bladder tumours

cystoscopy (first introduced by Tsuchida and Sugaware in 1973) [35] is recommended for males [36].

Since the introduction of flexible cystoscopy, comparable to flexible scopes in endoscopy, the complexity of handling this instrument has increased. The flexible cystoscope enables the urologist to navigate throughout the entire bladder. Its orientation inside the bladder during the procedure is made harder and more complex because of the limited optical information provided. That in turn means the procedure lasts longer, leading to more discomfort for the patient. The American Cancer Society states that the average cystoscopic procedure itself takes 5–10 min; however, it can often take as long as 10–20 min depending on the patient's history [37]. In addition to being complex and time-consuming, the procedure can be more mental and physical workload for the urologist. It is therefore important to feel confident with the instrument in order to minimise the diagnostic procedure's duration without risking the failure to detect bladder lesions requiring surgical treatment.

The learning curve of flexible cystoscopy is quite long (i.e. an average of 122 flexible cystoscopies must have been done to attain an acceptable expertise level for a novice urologist) [38]. Flexible cystoscopies done by an inadequately trained trainee are usually considerably longer than rigid cystoscopies. First, the technical handling of a flexible cystoscope must be learned. Second, the urologist's orientation, once inside the bladder, is somewhat difficult to maintain in some situations. Losing one's orientation means pulling the cystoscope back and having to re-insert it, making this one of the most painful moments, especially for male patients; retraction can also lead to bleeding in the bladder neck, thus impeding an accurate visualisation of the bladder mucosa.

Finally, this procedure is usually carried out in an outpatient setting with the patient fully awake. Nevertheless, this intervention is crucial in the follow-up and primary diagnosis of bladder cancer, especially in the early stages, making it vital for the patient to receive a thorough, accurate diagnosis. Even with the recent advantages in local anaesthesia and enhanced visualisation using visual chips instead of fiber optics, the procedure is still uncomfortable for many patients—a major concern in light of patients' reluctance, especially in cancer follow-up investigations [35, 39].

Simulators for urological procedures have been introduced to overcome the limited teaching/training time and practical operating room experience of residents/fellows [40], as well as increasing healthcare demands, fiscal constraints, and medico-legal considerations. Several types of simulators have been assessed. Their capability to enhance the training of residents/fellows has been widely successful [41, 42]. However, many of these simulators are still not employed during training. Some complain that they are too expensive/pricey. Also, novices have reported that they had trouble successfully relating the computer display to actual conditions encountered during cystoscopy—a factor raising the risk of error and longer learning curves [43].

Solution: dedicated diagnostic robots in Urology

A combination of robotic cystoscopy and XAI-enhanced imaging would overcome the manoeuvring and visual diagnostic issues encountered with flexible cystoscopy. Here the visualization component is important because it is here that precise human-AI interaction is sought, and for this, the international research community in the domain of visualization has developed a number of methods to improve the interactive processes, especially for tight interaction [44]. Such a combination would raise the standards of bladder cancer diagnosis and increase patient comfort. It would not require much experience to operate (i.e., a urologist is not required to operate it), and it could shorten the duration (e.g. 2–5 min instead of 10–20 min) of the procedure and improve diagnostic accuracy.

However, some patients have expressed worry about an invasive robotic procedure. Since the introduction of the DaVinci® tele-manipulator to urology, the term "robot" is commonly used in the literature and urology practice [45]. In everyday practice, it is sometimes hard to make the patient understand that it is not an autonomous robot alone performing the surgical procedure, but instead actually the urologist who steers the DaVinci® tele-manipulator. Most patients are relieved to learn that the "robot" is actually steered by a human and that it cannot make a single movement on its own. Acceptance and confidence in such a system must be established during the consultation before the patient gives informed-consent to this procedure.

In summary, an XAI robot-assisted flexible cystoscope could be inserted via the urethra into a patient's bladder like a soft transurethral catheter (an 8–12 Fr flexible scope instead of a 16–25 Fr rigid scope). A trained nurse could insert it following the prescribed standard procedure for transurethral catheterisation. The urethra is anaesthetised by a local lubricant gel, which is already standard care. The nurse could stay bedside during the entire robotic procedure, and take out the cystoscope when the scan is complete.

Once inserted, a flexible XAI robot-assisted cystoscope would be able to *semi-autonomously* perform a quicker, more comfortable and standardised diagnostic scan of the entire urinary bladder. The major anatomic landmarks can be identified. The cystoscope can fill (to a specific amount) and empty the bladder to maintain a perfectly clear view. If the patients feel any discomfort, they could also interrupt the procedure by pressing an emergency button (patient-in-the-loop) or simply tell the nurse to stop the procedure. After the entire bladder is scanned in a standardised mode, the nurse retracts/pulls out the cystoscope, and the patient is discharged. The collected data (scans) can then be processed

by AI and presented in a highly standardised mode to the urologist for the definitive diagnosis.

Urologists working in an outpatient setting would benefit from a quick and thorough performed diagnostic scan. Automated cystoscopy has the potential to increase the interobserver and intra-observational comparability and lead to more efficient diagnostics, which would ultimately improve patient safety. An XAI system that provides interpretable-explanations would also subsequently increase trust among urologists, nurses, and patients. Cystoscopy is considered a less life-threatening starting point for automation than open surgical procedures. Moreover, by not having to perform a diagnostic cystoscopy, urologists would reduce their mental and physical workload, thereby lowering the risk of adverse events/incidents when carrying out actual life-threatening surgeries. Insurance providers are also interested in new AI solutions that lower costs (i.e., the cost of having a nurse/technician attend a procedure is lower than paying a urologist to attend). The expected outcome would also be enabled by the AI-enhanced cystoscopy's professional second "opinion" that more accurately predicts a bladder cancer's recurrence, and would thus avoid the unnecessary transurethral resection of bladder tumors (TUR-BT) as well as increase patient safety [46–48]. AI should be at least as accurate as an experienced urologist in recognizing and differentiating bladder lesions, thereby improving the reproducibility of cystoscopy. AI can also play an important role in education for residency training.

One central problem in cystoscopy is the lack of a standardized documentation [49] and descriptions for the lesions found inside the bladder (e.g. EAU Guidelines [15, 17] are limited). Waldbillig et al. found that in surveys of German and international urologists, only 1 in 5 rate their cystoscopy as complete (missing about 10% of bladder tissue during cystoscopy in outpatient setting) and a significant gap in information transfer from outpatient to clinic due to documentation that is only analog and descriptive, i.e., no standardized video or picture-based documentation in 52% of the cases [50]. For the first time, AI will offer a structured reporting system and solution for risk profiling of BCa.

Considerations for the practical use of AI solutions as medical products

AI medical devices are challenging from a regulatory perspective. We are experiencing a situation where algorithms that are still learning in the treatment process are not legally permitted. Before approval, the methodology must, therefore, be presented as a design freeze.

There are fundamental differences between medical technology markets in the USA and Europe. The FDA has already approved a significant number of AI applications in the medical field. As there were no products appropriate

for the simplified 510 (k) approval process, many manufacturers submitted applications for a de novo procedure [51]. This resulted in most products being assigned to the second-highest Risk Class II. The subsequent path went through the so-called "Breakthrough Device" status. All these approvals were made with already frozen algorithms.

In the EU Regulation 2017/745, most such applications in Europe are assigned to the third-highest Risk Class IIa. For the manufacturer, this means that in addition to risk management according to ISO 14971, software-specific standards IEC 62304 and 82304-1 and their associated standards must also be complied with. In this respect, AI-based medical devices or medical devices with AI software elements are currently requiring significant efforts for granting approval.

The FDA has signalled new directions for future regulatory-logical processes in a discussion paper from 2019 [52]. Here, approvals for continuously learning solutions were held out in prospect. They would only be associated with a comprehensive quality management system. The problem here, however, is that the responsibility for the performance of such products is transferred to the manufacturers, even in light of the future's unpredictability. This in turn can lead to scepticism and caution on industry's part. For this reason, manufacturers should not only ensure quality management on an ongoing basis, they must also ensure continuous validation steps based on robust clinical parameters. Similar political movements are not yet foreseeable in Europe. However, we expect them in the near future. In general, AI-based medical devices are a key technological contribution to improving patient care, and they are certain to eventually become an integral part of everyday clinical practice.

The importance of XAI for robotic cystoscopy and its legal and ethical considerations

The European Union is very active in the field of ethics and in protecting human rights in AI. Its strategy is to ensure citizens' trust in AI. This strategy places the principle of explainability at the center of its concerns, along with other essential values (auditability, traceability, transparency, accountability, etc.). In the 20 October 2020 [53] resolution on a framework of ethical aspects of artificial intelligence, robotics and related technologies, the European Parliament aims to enshrine in law the common ethical principles of AI issues to make them more effective (point Y). This resolution mentions explainability in paragraph 23. First, it asserts that the adoption of these principles helps to ensure citizens' trust in those technologies. Second, it recognizes, however, that the degree of explainability is relative to the complexity of the concerned technologies and admits that it is not always possible to explain why a model has led to a particular result or decision, in particular for black box algorithms.

Third, it concludes that abiding by these principles is an essential precondition for guaranteeing accountability. However, a European resolution has no binding force on Member States.

The European Union has gone further by developing its first legal framework in its 21 April 2021 [54] proposal for a regulation on artificial intelligence (AI Act). Contrary to a resolution, a European regulation, once adopted, shall apply in all Member States. This future regulation adopts a risk-based approach to AI systems, which is also adopted in the resolution on a framework of ethical aspects [53]. The AI Act is intended to regulate the professional uses of AI. When placing AI systems on the market and using them, it imposes obligations that are proportional to the level of risk. High-risk AI systems for natural persons' health and safety or European fundamental rights are the subject of reinforced supervision. The AI Act proposes two classification rules for high-risk AI systems (Article 6). Either the AI system is listed in Annex III of the AI Act, establishing a list of automatically high-risk AI, but it does not include AI systems in health, or the AI system meets the conditions of Article 6, paragraph 1. In this case, it shall be considered high risk if this AI system constitutes a safety component of a product or a product covered by Annex II and is required to undergo a third-party conformity assessment with a view to its placing on the market.

As EU Regulation 2017/745 on medical devices is listed in Annex II and an AI robot-assisted cystoscopy requires third-party conformity assessment, this AI system should therefore be classified as high-risk. Strict obligations are placed on the providers of the high-risk AI system, such as responsibility for their compliance with the AI Act, drawing-up the technical documentation, keeping logs automatically generated by their systems or affixing the CE marking to their high-risk AI systems (Articles 16 to 23).

Even when a high-risk AI system is related to products, such as a robot, the robot manufacturer must fulfill the same obligations as the provider (Article 24). The future regulation is also aimed at high-risk AI systems professional users. It therefore concerns the doctor who implements an AI robot-assisted flexible cystoscope or the establishment the doctor belongs to. The professional user shall comply with a series of obligations (Article 29), which are more or less difficult to apply in practice. For example, using these systems in accordance with the instructions for use accompanying the systems, implementing the human oversight measures indicated by the provider, exercising control over the input data, monitoring the AI system operation, keeping the automatically generated logs, or interrupting the use of the AI system in the event of doubt as to its conformity. Non-compliance of the AI system with these requirements or obligations shall be subject to administrative fines of up to €20,000,000 or, if the offender is a company, up to 4% of its total worldwide annual turnover for the preceding financial year, whichever is higher (Article 71, paragraph 4).

The resolution on a framework of ethical aspects [53] also calls for important outcomes concerning the expected degree of explainability. Indeed, the latter "should depend on the context of those technical processes, and on the severity of the consequences of an erroneous or inaccurate output, and needs to be sufficient for challenging them and for seeking redress" (paragraph 19 of the proposal for a regulation contained in this resolution). In terms of AI robot-assisted cystoscopy, the doctor should not only need to provide information regarding the medical exploration and its risks, but also all the explanations concerning AI robotics so that the patient understands how the technology works, what the advantages and disadvantages are, or even the limits of the results obtained by the AI [55].

Consequently, the most relevant driver for XAI research is the growing relevance of legal aspects along with AI-assisted decision-making [6, 56]. Ultimately, it is mandatory in these areas that responsibility remain on the medical doctors, hence keeping the human-in-control [57]. Human control over the AI system is a central principle associated with responsibility. This is the case with the Montreal Declaration for the Responsible Development of Artificial Intelligence of 2018, which asserts that "only human beings can be held responsible for decisions stemming from recommendations made by AIS (Artificial Intelligence Systems), and the actions that proceed therefrom" (paragraph 9.1). The European Union has adopted a similar principle in its AI Act [54]. This proposal outlines that high-risk AI systems "shall be designed and developed in such a way, including with appropriate human–machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use" (Article 14, paragraph 1). The provider shall then identify measures to ensure human control, build them into the AI systems before they are placed on the market when technically feasible, or entrust their implementation to the user (Article 14, paragraph 3). The doctors, or their establishment, therefore have a real role to play in the implementation of the human control principle.

It is important to add that the 20 October 2020 European Parliament resolution with recommendations to the Commission on a civil liability regime for artificial intelligence [58] puts in place a strict liability for high-risk AI systems (Article 4).

Future work

This new XAI robot-assisted cystoscopy solution holds the potential to not only detect cancer, but to detect benign lesions as well. This capability is currently missing in biomarkers. However, it is still unclear how cost effective the

new XAI solution will be. Nevertheless, comparative studies using simulation analysis models (e.g. Monte Carlo [59]) could be carried out to mathematically build a case for the cost effectiveness of this new XAI solution.

Conclusion

Visual diagnostic issues in cystoscopy can occur during live operations when the urologist is distracted or over focused on wielding and navigating the cystoscope in the bladder (i.e., the urologist fails to detect a malignant lesion). The development of miniaturised (reduced to 8–12 Fr) XAI robot-assisted cystoscopes will improve the performance of flexible cystoscopy in terms of technical and diagnostic parameters. So far, there are still no biomarkers available that can compare to the accuracy of cystoscopy. Nevertheless, non-invasive biomarkers (e.g. K17) could still be used to stratify patients on waiting lists, to prioritise cases for XAI robot-assisted cystoscopy and avoid unnecessary hospital visits for other patients during a pandemic (e.g. COVID-19) [60].

The following two points highlight the uniqueness of the XAI solution: (i) this novel technology will detect and reduce cases of false negatives and false positives in this diagnostic procedure; and (ii) it gives the patient (in-the-loop) more control than ever before during this diagnostic procedure. This semi-autonomous procedure will be controlled by the nurse as well. This is the best way to guarantee and maximise patient safety. This new solution has the potential to dramatically increase patient acceptance and adherence, making cystoscopy a much less dreaded procedure. With the introduction of dedicated evaluation robots, patients will benefit from a more comfortable, quicker, thorough, and effective diagnostic scan. However, the definitive decision to perform another transurethral resection of bladder tumors (TURBT) or schedule a follow-up cystoscopy will be made by the urologist. But with XAI, that decision will rest on a more solid diagnostic foundation.

Declarations

Conflict of interest The authors have no conflict of interest to declare.

Research involving human participants This article does not contain any studies with human participants performed by any of the authors.

Informed consent Not applicable.

References

1. O'sullivan S et al (2019) Operational framework and training standard requirements for AI-empowered robotic surgery. Int J Med Robot 16(5):1–13. https://doi.org/10.1002/rcs.2020

2. Piccialli F et al (2021) A survey on deep learning in medicine: why, how and when? Information Fusion 66:111–137. https://doi.org/10.1016/j.inffus.2020.09.006

3. Bengio Y, Lecun Y, Hinton G (2021) Deep learning for AI. Commun ACM 64(7):58–65. https://doi.org/10.1145/3448250

4. Holzinger A et al (2019) Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev: Data Min Knowl Discov 9(4):1–13. https://doi.org/10.1002/widm.1312

5. Schneeberger D, Stoeger K, Holzinger A (2020) The European legal framework for medical AI. International cross-domain conference for machine learning and Knowledge extraction, fourth IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 international cross-domain conference, CD-MAKE 2020, Proceedings. Springer, Cham, pp 209–226. https://doi.org/10.1007/978-3-030-57321-8-12

6. Mueller H et al (2021) The ten commandments of ethical medical AI. IEEE Comput 54(7):119–123. https://doi.org/10.1109/MC.2021.3074263

7. Legido-Quigley H, Mckee M, Walshe K, Sunol R, Nolte E, Klazinga N (2008) How can quality of health care be safeguarded across the European Union? BMJ 336(7650):920–923. https://doi.org/10.1136/bmj.39538.584190.47

8. Stahl BC, Coeckelbergh M (2016) Ethics of healthcare robotics: Towards responsible research and innovation. Robot Auton Syst 86:152–161. https://doi.org/10.1016/j.robot.2016.08.018

9. Gunkel D (2012) The machine question: Critical perspectives on AI, robots, and ethics. MIT Press. https://doi.org/10.7551/mitpress/8975.001.0001

10. Hengstler M, Enkel E, Duelli S (2016) Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. Technol Forecast Soc Chang 105:105–120. https://doi.org/10.1016/j.techfore.2015.12.014

11. London AJ (2019) Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep 49(1):15–21. https://doi.org/10.1002/hast.973

12. Siegel RL et al (2019) Cancer statistics, 2019. CA A Cancer J Clin 69:7–34. https://doi.org/10.3322/caac.21551

13. Robert Koch-Institut (RKI) (2018) Cancer in Germany 2013/2014. RKI. https://doi.org/10.25646/5889

14. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer J Clin 68:394–424. https://doi.org/10.3322/caac.21492

15. Babjuk M et al (2017) EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder: update 2016. Eur Urol 71(3):447–461. https://doi.org/10.1016/j.eururo.2016.05.041

16. Mertens L, Neuzillet Y, Horenblas S et al (2014) Landmarks in non-muscle-invasive bladder cancer. Nat Rev Urol 11:476–480. https://doi.org/10.1038/nrurol.2014.130

17. Alfred Witjes J, Lebret T, Compérat EM et al (2017) Updated 2016 EAU guidelines on muscle-invasive and metastatic bladder cancer. Eur Urol 71(3):462–475. https://doi.org/10.1016/j.eururo.2016.06.020

18. Maas M, Bedke J, Stenzl A, Todenhöfer T (2019) Can urinary biomarkers replace cystoscopy? World J Urol 37(9):1741–1749. https://doi.org/10.1007/s00345-018-2505-2

19. Babu S, Roa-Peña L, Chan I, Kim N, Jahanfard S, Escobar-Hoyos L, Shroyer K (2020) Abstract B01: Validation of a novel cytologic biomarker for urothelial carcinoma. B01–B01. Am Assoc Cancer Res. https://doi.org/10.1158/1557-3265.BLADDER19-B01

20. Fradet Y et al (2007) A comparison of hexaminolevulinate fluorescence cystoscopy and white light cystoscopy for the detection of carcinoma in situ in patients with bladder cancer: a phase III, multicenter study. J Urol 178(1):68–73. https://doi.org/10.1016/j.juro.2007.03.028 (**discussion 73**)

21. Hermann GG et al (2011) Fluorescence-guided transurethral resection of bladder tumours reduces bladder tumour recurrence due to less residual tumour tissue in Ta/T1 patients: a randomized two-centre study. BJU Int 108(8 Pt 2):E297-303. https://doi.org/10.1111/j.1464-410X.2011.10090.x

22. Capece M et al (2020) Hexaminolevulinate blue light cystoscopy (Hal) assisted transurethral resection of the bladder tumour vs white light transurethral resection of the bladder tumour in non-muscle invasive bladder cancer (NMIBC): a retrospective analysis. Arch Ital Urol Androl: Organo Uff Soc Ital Ecografia Urol Nefrol 92(1):17–20. https://doi.org/10.4081/aiua.2020.1.17

23. Mayr R, Burger M (2013) Value of fluorescence cystoscopy in high risk non-muscle invasive bladder cancer. Curr Urol Rep 14(2):90–93. https://doi.org/10.1007/s11934-013-0306-0

24. Storz K. Blue Light Cystoscopy with Cysview® System PDD. Available at www.karlstorznetwork1.com/blue-light-cystoscopy

25. Kim SB et al (2018) Detection and recurrence rate of transurethral resection of bladder tumors by narrow-band imaging: Prospective, randomized comparison with white light cystoscopy. Investig Clin Urol. https://doi.org/10.4111/icu.2018.59.2.98

26. Negassi M et al (2020) Application of artificial neural networks for automated analysis of cystoscopic images: a review of the current status and future prospects. World J Urol 38:2349–2358. https://doi.org/10.1007/s00345-019-03059-0

27. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-fei L (2009) Imagenet: a large-scale hierarchical image database. IEEE Conf Comput Vis Pattern Recogn 2009:248–255. https://doi.org/10.1109/CVPR.2009.5206848

28. Eminaga O et al (2018) Diagnostic classification of cystoscopic images using deep convolutional neural networks. JCO Clin Cancer Inform 2:1–8. https://doi.org/10.1200/CCI.17.00126

29. Henry M, Patil D, Filson C, Mehta A (2020) Antibiotic stewardship for ambulatory urologic procedures: cystoscopy and vasectomy. J Urol 7(4):272–278. https://doi.org/10.1097/UPJ.0000000000000112

30. Lotan Y et al (2019) Blue light flexible cystoscopy with hexaminolevulinate in non- muscle-invasive bladder cancer: review of the clinical evidence and consensus statement on optimal use in the USA—update 2018. Nat Rev Urol 16(6):377–386. https://doi.org/10.1038/s41585-019-0184-4

31. Gezginci E et al (2018) Three Distraction Methods for Pain Reduction During Cystoscopy: A Randomized Controlled Trial Evaluating the Effects on Pain, Anxiety, and Satisfaction. J Endourol. https://doi.org/10.1089/end.2018.0491

32. EAU (2020) EAU Guidelines. EAU Annual Congress, Amsterdam. Available at https://uroweb.org/wp-content/uploads/EAU-MIBC-Guidelines-2018V2.pdf

33. Freitas NR, Vieira PM, Lima E, Lima CS (2018) Automatic T1 bladder tumor detection by using wavelet analysis in cystoscopy images. Phys Med Biol 63(3):035031. https://doi.org/10.1088/1361-6560/aaa3af

34. Pearce S, Daneshmand S (2018) Enhanced endoscopy in bladder cancer. Curr Urol Rep 19(10):84. https://doi.org/10.1007/s11934-018-0833-9

35. Tsuchida S, Sugawara H (1973) A new flexible fiber cystoscope for visualization of the bladder neck. J Urol 109(5):830–831. https://doi.org/10.1016/s0022-5347(17)60554-8

36. Casteleijn NF et al (2017) The effect of office based flexible and rigid cystoscopy on pain experience in female patients. Investig Clin Urol 58(1):48–53. https://doi.org/10.4111/icu.2017.58.1.48

37. The American Cancer Society medical and editorial content team: patient info side on "Cystoscopy", last updated Jan 14th 2019 and American Urological Association. Cystoscopy. Accessed at www.urologyhealth.org/urologic-conditions/cystoscopy on November 9, 2018)

38. MacKenzie KR, Aning J (2016) Defining competency in flexible cystoscopy: a novel approach using cumulative sum analysis. BMC Urol. https://doi.org/10.1186/s12894-016-0143-9

39. Okhunov Z, Hruby GW, Mirabile G et al (2009) Prospective comparison of flexible fiberoptic and digital cystoscopes. Urology 74(2):427–430. https://doi.org/10.1016/j.urology.2009.01.007

40. Watterson JD, Denstedt JD (2007) Ureteroscopy and cystoscopy simulation in urology. J Endourol. https://doi.org/10.1089/end.2007.9982.A

41. Schout BM et al (2008) Update on training models in endourology: a qualitative systematic review of the literature between January 1980 and April 2008. Eur Urol 54(6):1247–1261. https://doi.org/10.1016/j.eururo.2008.06.036

42. Naylor R et al (2009) Can medical students achieve skills proficiency through simulation training? Am J Surg 198:277–282. https://doi.org/10.1016/j.amjsurg.2008.11.036

43. Hu WG, Feng JY, Wang J et al (2015) (2015) Ureteroscopy and cystoscopy training: comparison between transparent and non-transparent simulators. BMC Med Educ 15:93. https://doi.org/10.1186/s12909-015-0380-8

44. Turkay C et al (2014) On computationally-enhanced visual analysis of heterogeneous data and its application in biomedical informatics. Lecture notes in computer science LNCS 8401. Springer, Berlin, Heidelberg, pp 117–140. https://doi.org/10.1007/978-3-662-43968-5_7

45. Merseburger AS, Herrmann TR, Shariat SF et al (2013) EAU guidelines on robotic and single-site surgery in urology. Eur Urol 64(2):277–291. https://doi.org/10.1016/j.eururo.2013.05.034

46. Shkolyar E et al (2019) Augmented bladder tumor detection using deep learning. Eur Urol. https://doi.org/10.1016/j.eururo.2019.08.032

47. Galgano SJ, Porter KK, Burgan C, Rais-Bahrami S (2020) The role of imaging in bladder cancer diagnosis and staging. Diagnostics 10:703. https://doi.org/10.3390/diagnostics10090703

48. Suarez-Ibarrola R et al (2020) A novel endoimaging system for endoscopic 3D reconstruction in bladder cancer patients. Minim Invasive Ther Allied Technol 3:1–8. https://doi.org/10.1080/13645706.2020.1761833

49. Suarez-Ibarrola R et al (2020) Current standards in the endoscopic management of bladder cancer: a survey evaluation among urologists in German-speaking countries. Urol Int 104(5–6):410–416. https://doi.org/10.1159/000506653

50. Waldbillig F, Hein S, Grüne B et al (2020) Current European trends in endoscopic imaging and transurethral resection of bladder tumors. J Endourol 34(3):312–321. https://doi.org/10.1089/end.2019.0651

51. The Medical Futurist (2019) FDA approvals for smart algorithms in medicine in one giant infographic", Available at https://medicalfuturist.com/fda-approvals-for-algorithms-in-medicine

52. FDA (2019) Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)—based software as a medical device (SaMD)—discussion paper and request for feedback, Available at https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf

53. European Parliament (2020) Resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)). Available at https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html

54. European Parliament and Council (2021) Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial intelligence Act) and amending certain union legislative acts, 21.4.2021, COM(2021) 206 final, 2021/0106 (COD). Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN

55. Nevejans N (2017) Treatise of Law and Ethics of Civil Robotics [Traité de droit et d'éthique de la robotique civile (in French only)], LEH Editions, esp. pp. 646–651. Available at https://www.leh.fr/edition/p/traite-de-droit-et-d-ethique-de-la-robotique-civile-9782848746685

56. O'sullivan S, Nevejans N, Allen C, Blyth A, Leonard S, Pagallo U, Holzinger K, Holzinger A, Sajid MI, Ashrafian H (2019) Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. Int J Med Robot Comput Assist Surg 15(1):1–12. https://doi.org/10.1002/rcs.1968

57. Wagner B (2019) Liable, but not in control? Ensuring meaningful human agency in automated decision-making systems. Policy Internet 11(1):104–122. https://doi.org/10.1002/poi3.198

58. European Parliament (2020) Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)). Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020IP0276&from=EN

59. Suryavanshi M, Mehta A, Jaipuria J, Kumar D, Vishwakarma G, Panigrahi MK, Verma H, Saifi M, Sharma S, Tandon S, Doval DC, Das BC (2018) Clinical utility of RT-PCR in assessing HER 2 gene expression versus traditional IHC and FISH in breast cancer patients. Breast Cancer 25(4):416–430. https://doi.org/10.1007/s12282-018-0840-1

60. NICE UK (2021) URO17 for detecting bladder cancer Medtech innovation briefing. Available at www.nice.org.uk/guidance/mib250