Journal of Clinical and Medical Images

Research Article

ISSN: 2640-9615 | Volume 8

Improvement of the Quality of Prostate Cancer Diagnosis Using AI In Digital Pathology Short Heading: Quality Improvement by Using AI In Pathology

Alexi Baidoshvili[1,2], Mariam Khacheishvili[2*], and Christina Hulsbergen-Van De Kaa[1]

[1]Laboratory of Pathology East Netherlands (LabPON), Hengelo, The Netherlands

[2]David Tvildiani Medical University, Tbilisi, Georgia

[*]Corresponding author:

Mariam Khacheishvili,
David Tvildiani Medical University, Tbilisi, Georgia

Received: 10 Nov 2024
Accepted: 16 Dec 2024
Published: 22 Dec 2024
J Short Name: JCMI

Citation:

Mariam Khacheishvil, Improvement of the Quality of Prostate Cancer Diagnosis Using AI In Digital Pathology Short Heading: Quality Improvement by Using AI In Pathology. J Clin Med Img. 2024; V8(7): 1-6

Keywords:

Prostate cancer; AI; Digital pathology; Ibex prostate galen; Machine learning,

Abbreviations:

Labpon: Laboratorium Pathologie Oost-Nederland**;** WSI: Whole Slide Image; UFS: Ultra-Fast Scanner; PCNB: Prostate Core Needle Biopsy**;** AUC: Area Under the Receiver Operating Characteristic Curve**;** ML: Machine Learning**:** ADC: Adenocarcinoma**;** ANN: Artificial Neural Network**;** PIN: Intraepithelial Invasion

1. Abstract

1.1. Aim

Automated AI-based tools are becoming increasingly important in modern medicine, including pathology, by significantly supporting pathologists' diagnoses and reducing human biases. Histopathological evaluation of prostate biopsies plays a crucial role in diagnosing prostate cancer. Pathologists assess tumor type, grade (Gleason Grade), and tumor extension to determine the management plan. Diagnosis accuracy, particularly in tumor grading, can be affected by inter- and intraobserver variability among pathologists.1 Due to the increased incidence of prostate cancer and subsequent workload on pathologists,2 an AI-based tool like Ibex Prostate Galen, can potentially reduce pathologists' workflow and enhance diagnostic accuracy.3 This study aimed to retrospectively compare histologically diagnosed prostate cancer by pathologists to the AI-based algorithm, Ibex Prostate Galen. The study evaluates the algorithm's impact on laboratory workflow and diagnostic accuracy.

Methods: The study was conducted at the Laboratory of Pathology East Netherlands (Lab PON, Hengelo, The Netherlands), using hematoxylin and eosin-stained (H&E) Whole Slide Images (WSI) from 2021. A total of 169 randomly selected and de-identified prostate biopsy cases, consisting of 809 slides and 701 parts, were used. Slides were digitized using a Philips Ultrafast Scanner (UFS). Of these, 674 parts from 168 cases were used for the study, while 33 slides were excluded: 16 slides lacked a definitive diagnosis from the original report, and 17 slides were out-of-focus. According to pathologists' diagnoses, 391 parts (58%) were benign, and 283 (42%) contained carcinoma. Ibex Prostate Galen, a validated and FDA-approved AI tool developed using advanced machine learning techniques, particularly convolutional neural networks (CNNs), assessed slide-level scores for cancer probability, Gleason grading, and perineural invasion. The algorithm's performance was evaluated using the area under the receiver operating characteristic curve (AUC).

1.2. Results

The algorithm demonstrated high accuracy with an AUC of 0.996. Ibex Prostate Galen correctly classified 391 parts as benign, with a negative predictive value (NPV) of 99.7%, and 283 parts as cancerous, with a positive predictive value (PPV) of 99.6%. For Gleason grading, the AI and pathologists agreed on the same grade group for 61.6% of cases and had only one grade group difference in 96.7% of cases. Expert review of discrepancies revealed 4 cases (2%) with differing results between pathologists and AI. These results indicate that using the Ibex AI algorithm in daily laboratory

activities is beneficial and increases diagnostic accuracy. The AI platform also demonstrated broad capabilities, including the detection of benign features such as atrophy and inflammation, allowing for efficient and comprehensive reporting.

2. Introduction

In recent years, cancer prevalence has reached approximately 1.7 million cases and is responsible for 0.6 million deaths annually in the US. With increasing cancer incidence and mortality, accurate diagnosis and effective management become imperative. Prostate adenocarcinoma is the second most common noncutaneous cancer diagnosis in men, with up to 1 million new diagnoses annually. Histopathological evaluation of prostate biopsies is critical for accurate diagnosis. However, pathologists face increased workloads due to rising tumor prevalence and complexity, which can lead to missed diagnoses and reduced confidence in treatment decisions. Automated AI-based tools are becoming increasingly important in modern medicine, including pathology, by potentially enhancing diagnostic accuracy and reducing human biases. Technical advances in machine learning (ML), a subset of artificial intelligence (AI), have enabled the development of sophisticated models, such as convolutional neural networks (CNNs), which replicate human neural processes. The use of ML in pathology has shown impressive results in improving diagnostic accuracy and reducing false positives and negatives. Our study aimed to compare histologically diagnosed prostate cancer by pathologists to the AI-based algorithm, Ibex Prostate Galen. This validated and FDA-approved AI tool, which has also been validated across various institutions, showcases its effectiveness in enhancing diagnostic accuracy and workflow efficiency in pathology laboratories.

3. Materials and Methods

This study conducted a blinded comparison of the AI-based algorithm for prostate adenocarcinoma detection and Gleason grading. The retrospective study design was based on validation of the algorithm conducted in the Laboratory for Pathology East Netherlands (Lab PON, Hengelo, The Netherlands), using randomly chosen hematoxylin and eosin (H&E) Whole slide images (WSI) of Prostate Core Needle Biopsies, with an original sign-out date in 2021. Since 2019 Lab PON has no general pathologists but all pathologists [18] are exclusively subspecialized in a few [3 -7] pathology fields with 7 uropathologists evaluating all uropathology. H&E stained slides were scanned using a Philips Ultra-Fast Scanner (Philips Digital Pathology Solutions; Best, Netherlands) at 40× magnification (resolution of 0·25 μm/pixel). The biopsies were presented in parts (I – VI) taken from several regions of the prostate gland. Each part was presented in separate WSIs, averaging 6 WSI per prostate case. The randomly selected dataset consisted of 169 cases, 701 parts, and 809 slides. Of these, 676 slides with 674 parts of 168 cases were used for the study, while 33 slides were ignored: 16 slides without a definite diagnosis from the original report and 17 slides out-of-focus. The set consisted of all grades of prostate cancer and included high–grade prostatic intraepithelial neoplasia, inflammation, and atrophy. Ibex Prostate Galen, the software that was integrated with IMS Proscia Concentriq, assessed slide-level scores for the probability of cancer, Gleason grading, perineural invasion, and calculation of cancer percentage for each prostate core needle biopsy (Figure1, 2, 3). All data was anonymized, and the corresponding metadata was recorded. The algorithm was run on the External Blind validation set. Three expert uropathologists with approximately from 20 to 29 years of experience reviewed discrepancies in the original diagnosis and the AI reading. Algorithm performance and grading were assessed with the area under the receiver operating characteristic curve (AUC). The concordance agreement between AI-based algorithms and pathologists was also evaluated.

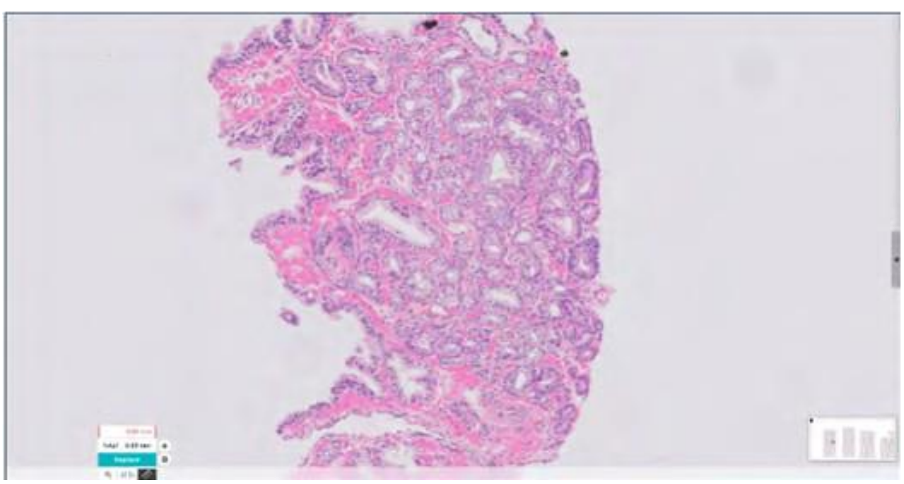

Figure 1: AI algorithm showing the tumour tissue (G3+3: atypical ducts) on WSI.

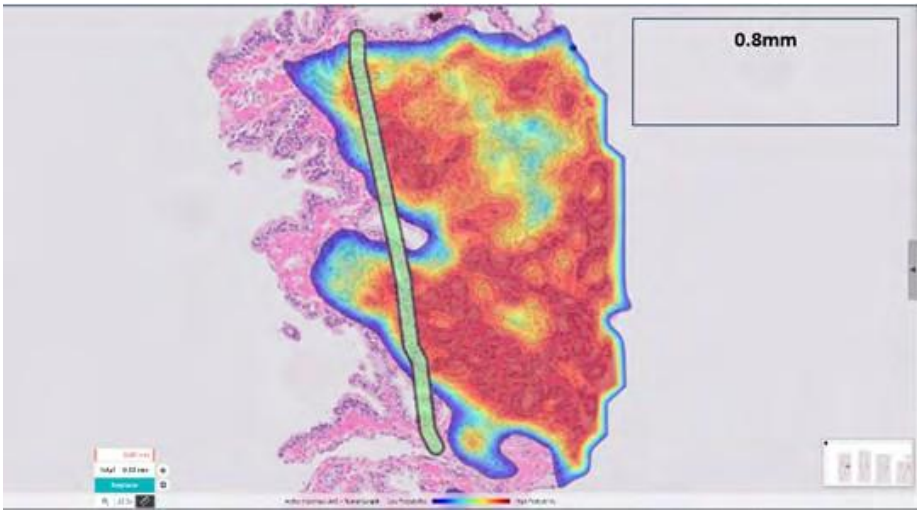

Figure 2: Showing the tumour length (0.8mm) measured by AI algorithm.

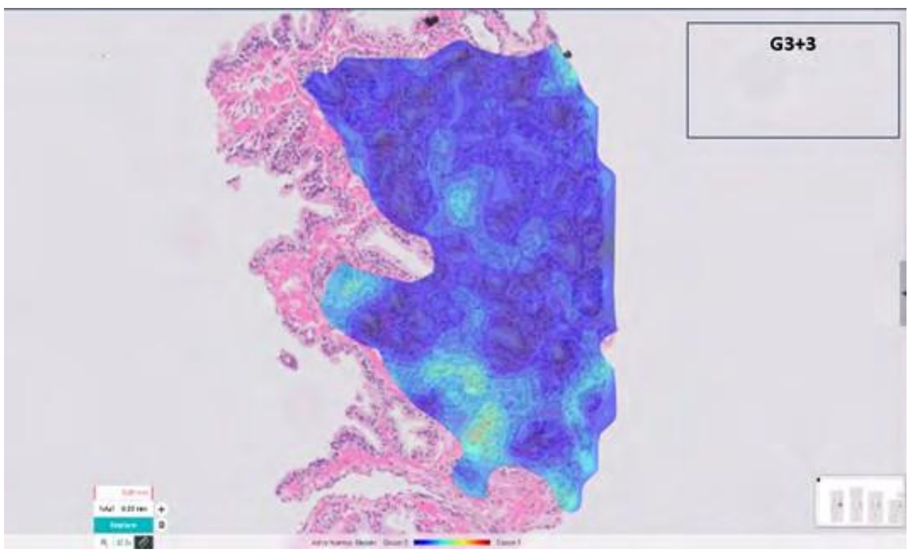

Figure 3: Showing the grading of the tumour accordingly to Gleason grading assessed by the AI algorithm (G3+3).

4. Results

The study showed an AUC for cancer detection of 0,996 (95% Cl 0.995 – 0.998) in the validation set. According to pathologists' diagnoses, 391 parts (58%) were benign, and 283 (42%) were classified as carcinoma. The Ibex AI algorithm classified 337 parts as benign with a negative predictive value (NPV) of 99.7% and 236 as cancer with a positive predictive value (PPV) of 99.6%. The Ibex AI algorithm classified 101 parts as «suspicious/undetermined», of which 47 parts were classified as cancer and 54 parts as benign by the pathologists.

The concordance agreement on grade group level between the AI algorithm and the uropathologists was also analyzed. Following the validation process, 18 cases (11%) were again checked by 3 expert uropathologists for reviewing the primary diagnosis, leading to a total of 4 cases (2%) showing differences between pathologists & AI. Therefore, three expert uropathologists studied these 4 discrepant cases for ground truth establishment. The main discrepancies between the AI algorithm and the pathologist's reports in the 4 cases are given in Table 1. Case 153 (part I), where the primary diagnosis was benign, was classified as cancer by the IBEX AI algorithm, with cancer scoring 0.999. Ibex highlighted a small focus on being highly suspicious of cancer. The expert uropathologist review result was a high-grade PIN and the outcome for the diagnosis changed from benign to non-malignant (high-grade PIN). The primary report for case 107 (part II) was diagnosed as benign. Again, IBEX highlighted a small focus on being highly suspicious of cancer. As a result of the uropathologists review, diagnosis remained inconclusive, but further testing was recommended for definitive diagnosis. Figure 4 and 5 are showing histopathologic pictures and heatmaps of these 2 cases. Case 136 (part V) where the original diagnosis was cancer and it was classified as "Undetermined" by the IBEX Galen AI-based tool after expert review corrected to benign and the final diagnosis changed from malignant to benign. Regarding the Case 79 (part I) the primary report was cancer G4+4 which was classified as benign by IBEX Galen AI-based tool, after expert reviewing was revised as intraductal adenocarcinoma or invasive cribriform adenocarcinoma and as a result the case outcome changed from malignant to a different type of malignancy. The percent of full agreement in grade group between the original pathology report and Ibex AI was 61.6 %, and the rate of grade group difference up to 1 = 96.7%.

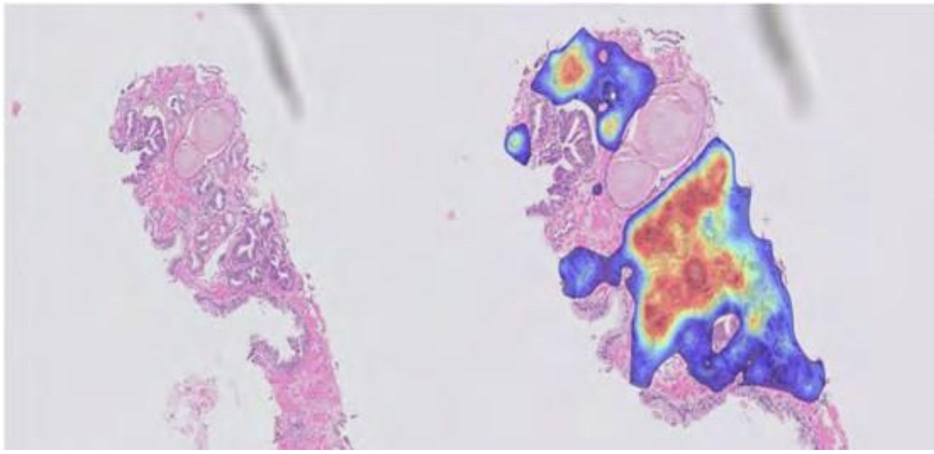

Figure 4: Case 153, part I: Original report was diagnosed as benign.
Ibex highlighted small focus highly suspicious of cancer. Review and ground truth adjudication: High-grade PIN

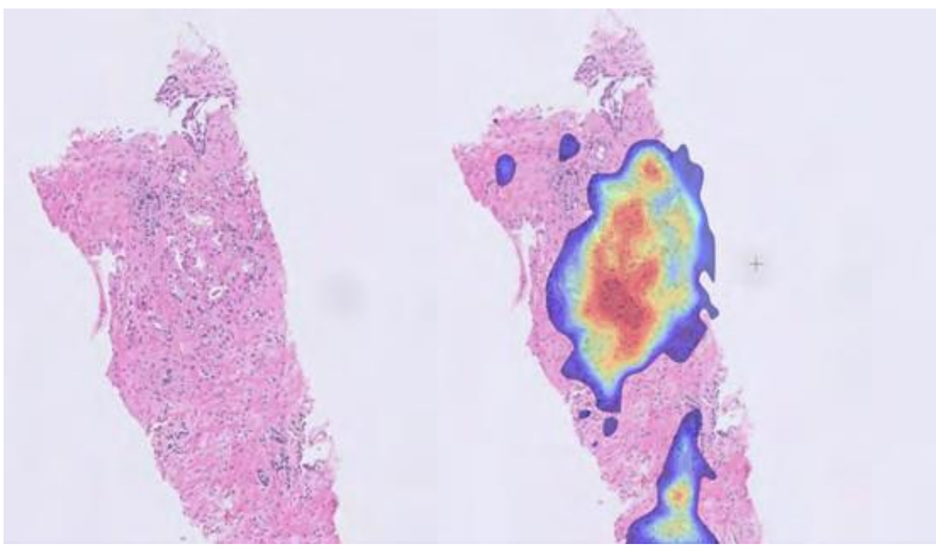

Figure 5: Case 107, part II: Original report was diagnosed as benign. Ibex highlighted small focus highly suspicious of cancer. Review and ground truth adjudication: IHC required for diagnosis.

Table 1: Reporting the top 4 discrepancies between pathologists and AI algorithm according to Gleason grade and detection of perineural invasion.

Case, Part	Original report	Ibex cancer score classification	Diagnosis after expert pathologist review
Case 153, part I	Benign	Cancer	Benign, HG PIN
Case 107, part II	Benign	Cancer	Suspicious, IHC required
Case 136, part V	AdC G5+5	Undetermined	Benign
Case 079, part I	AdC G4+4	Benign	AdC intraductal or invasive

5. Discussion

We report on the evaluation of an AI-based algorithm to evaluate digital prostate CNB slides and the validation of this AI tool in routine clinical practice in a community pathology laboratory with uropathology specialized pathologists. We demonstrated the high accuracy on a large blinded validation dataset to identify and qualify prostate cancer and assess the accuracy and objectivity of the AI tool in the PCNB. Of course, there are several studies on the validation of this algorithm; however, this study is further proof that similar algorithms are effective for Gleason grade detection but also can be useful for other tumor features and implementation in clinical practice [8]. The development of such AI tool algorithms and their introduction into clinical practice is essential not only because prostate adenocarcinoma is one of the most common cancers among men but also because the prevalence of tumors is increasing daily. Consequently, the workload of pathologists is increasing, especially when the histopathological assessment becomes more and more complex according to the new guidelines and recommendations. If we compare the number of new cancer cases every year, it turns out that the workload of pathologists has increased, which is the reason for potentially delayed reports and increased medical errors.6

Because there is much research on medical errors due to the increased workload of pathologists as well as a shortage in pathologists and low concordance in Gleason grading, which might change the diagnosis and consequently the further treatment during prostate cancer, many clinics use the second-read option to eliminate similar misdiagnosis, which is mainly done manually by pathologists. Such chances further increase the workload of pathologists and are associated with additional costs and technical difficulties [9]. Our study showed that the IBEX Galen AI-based algorithm has a high diagnostic accuracy (AUC = 0.996) and therefore can offer a highly efficient second reading method. Overall, no large discrepancies were found in our research, but only 4 discrepant cases out of 168 cases (2%) were identified concerning small suspicious areas in only one biopsy of one part (out of 10-14 biopsies or 4 to 6 parts) with diagnostic discrepancy according to cancer or precancer diagnosis, these changes in diagnosis based on expert review highlight the importance of thorough evaluation and consideration of multiple factors in pathology diagnosis, as well as the potential for AI algorithms to aid in detection but also the need for human expertise in interpretation and final diagnosis. Such discrepancies could have significant clinical consequences for patient management and treatment decisions [10]. Several studies have focused solely on the Gleason grading aspect of AI-based algorithms; however, our research has revealed that these algorithms possess broader functionalities [11]. The IBEX Galen AI-based tool demonstrated the capability to simultaneously detect tumors and evaluate their grading. Furthermore, although not within the scope of this study, it also exhibited the ability to identify perineural tumor invasion. Beyond tumor detection and grading, this AI-based algorithm exhibits versatile capabilities, including the detection of high-grade prostatic intraepithelial neoplasia (PIN), atrophy, and inflammation, as well as assessing tumor length. While these aspects were not specifically addressed in our study, they contribute to enhanced efficiency and accuracy in diagnosis. Integration of the IBEX Galen AI-based platform into Lab PON laboratory workflows will facilitate the display of case lists with preselection of slides likely to contain malignant or benign cases. This enables quick triage and shorter turnaround times. The AI-based classification of blocks as suspicious for cancer also implies potential savings in immunohistochemistry (IHC) and further reduces turnaround times. Therefore, the implementation of AI-based tools in routine clinical practice warrants careful consideration and offers laboratories a significant opportunity to enhance efficiency and productivity among pathologists.

6. Conclusion

The emergence of AI-based tools in pathology marks a transformative change in diagnostic assessments, addressing challenges faced by pathologists, particularly with prostate adenocarcinoma, one of the most prevalent cancers among men. Accurate and timely diagnoses are crucial for patient management, and AI algorithms like Ibex Prostate Galen enhance cancer detection and grading, reducing diagnostic errors and treatment delays. Pathologists are experiencing increased workloads due to the rising prevalence of cancers and the complexity of modern histopathological assessments. The intricate Gleason grading system, requiring detailed evaluation, can lead to bottlenecks and human errors that negatively impact patient care. AI technologies, such as the Ibex Prostate Galen system, offer significant benefits, including high diagnostic accuracy and the ability to serve as a reliable second-read method, thus alleviating the burden of labor-intensive manual readings.

In addition to diagnostic improvements, AI serves an educational role, allowing pathologists to reassess their diagnostic processes and refine their skills through feedback on discrepancies between their interpretations and AI results. While minimal diagnostic discrepancies have been noted, they emphasize the complexity of pathology and the necessity for human expertise in final diagnoses, reinforcing that AI should complement rather than replace pathologists. Furthermore, the Ibex Galen algorithm is capable of detecting various pathological features, contributing to comprehensive patient management. AI can streamline workflows, reduce unnecessary tests, and ultimately lead to cost savings in healthcare, transforming laboratory practices to meet the increasing demands of modern healthcare systems.

7. Declarations

7.1. Ethics Approval and Consent to Participate

We want to inform the editorial team that this study was conducted in Lab PON's laboratory (Laboratorium Pathologie Oost-Nederland). The laboratory has a scientific committee meeting to discuss planned research issues (including bioethics). According to research, different technologies (AI software, digital microscope) are used to compare diagnostic reports during routine workflow. Even during the routine diagnostic process, patients who leave diagnostic material in the laboratory sign a consent implying that their tissues and biological material may be used in various studies. Furthermore, complete de-identification occurs when processing patient tissue and creating digital whole slide images or non-digital slides from it, which protects the principle of medical secrecy and confidentiality. Because the article addresses and discusses the characteristics of the patient's tissue and the diagnosis, the Ethics Committee discussed the current article and research. Also, the approval of the Lab PON Laboratory's Scientific Committee was sufficient.

8. Consent for Publication

We would also like to inform you that the study does not include the patient's personal information. Only statistics on the comparison diagnostic reports are used. The study compared the primary diagnosis of pathologists to diagnoses made by AI software. Additionally, patients sign a consent implying that their tissues and biological material may be used in various studies. Accordingly, the Ethics Committee and LabPON Laboratory's Scientific Committee discussed the current article and research, and there was no disagreement with the consent of the patient and their legal representatives.

References

1. Ozkan TA, Eruyar AT, Cebeci OO, Memik O, Ozcan L, Kuskonmaz I. Interobserver variability in Gleason histological grading of prostate cancer. Scand J Urol. 2016; 50(6): 420-424.
2. Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW. Artificial Intelligence in Pathology. J PatholTransl Med. 2019; 53(1): 1-12.
3. Strom P, Kartasalo K, Olsson H. Artificial intelligence for diagnosing and grading prostate cancer in biopsies: a population-based. Diagnostic Study, Lancet Oncol. 2020; 21: 222-232.
4. Bulten, Wouter, Kimmo Kartasalo, Po-Hsuan Cameron Chen, Peter Ström, Hans Pinckaers. "Artificial Intelligence for Diagnosis and Gleason Grading of Prostate Cancer: The PANDA Challenge." Nature Medicine. 2022; 28: 154-63.
5. Metter DM, Colgan TJ, Leung ST, Timmons CF, Park JY. Trends in the US and Canadian pathologist workforces from 2007 to 2017. JAMA Netw Open. 2019.
6. Kronz JD, Milord R, Wilentz R, Weir EG, Schreiner SR, Epstein JI. Lesions missed on prostate biopsies in cases sent in for consultation. Prostate. 2003; 54: 310-14.
7. Pantanowitz L, Quiroga-Garza GM, Bien L, Heled R, Laifenfeld D, Linhart C. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. Lancet Digit Health. 2020; 2(8): e407-e416.
8. Parwani AV. Commentary: Automated Diagnosis and Gleason Grading of Prostate Cancer - Are Artificial Intelligence Systems Ready for Prime Time? Journal of Pathology Informatics. 2019; 10: 41.
9. Zhang Lj, Wu B, Zha Zl. Perineural invasion as an independent predictor of biochemical recurrence in prostate cancer following radical prostatectomy or radiotherapy: a systematic review and meta-analysis. BMC Urol. 2018; 18: 5.
10. Nagpal K, Foote D, Liu Y. Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. npj Digit. Med. 2019.
11. Guy Nir, Soheil Hor, Davood Karimi, Ladan Fazli, Brian F. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts, Medical Image Analysis, Volume. 2018; 167-180.