

The uropathologist of the future: getting ready with intelligence for the prostate cancer tsunami

Alessia Cimadamore^{1*}, Rodolfo Montironi^{2*}, Liang Cheng³, Antonio Lopez-Beltran⁴, Eamonn T. Rogers⁵, Carmine Franzese⁶, Alessandro Crestani⁶, Gianluca Giannarini⁶

¹ Institute of Pathological Anatomy, Department of Medicine, University of Udine, Udine, Italy; ² Molecular Medicine and Cell Therapy Foundation, c/o Polytechnic University of the Marche Region, Ancona, Italy;

³ Department of Pathology and Laboratory Medicine, Department of Surgery (Urology), Brown University Warren Alpert Medical School, Lifespan Health, and the Legorreta Cancer Center at Brown University, Providence, RI, USA; ⁴ Department of Surgery, Cordoba University Medical School, Cordoba, Spain; ⁵ Department of Urology, National University of Ireland Galway, Galway, Ireland; ⁶ Urology Unit, University Hospital, Ospedale Santa Maria della Misericordia, Udine, Italy

*Contributed equally

Summary

According to the recently published paper by the *Lancet* Commission on prostate cancer (PCa) ¹, the projections of new cases of PCa will rise from 1.4 million in 2020 to 2.9 million by 2040. Such a rise cannot be prevented by public health interventions and lifestyle changes. Late diagnosis of PCa is “widespread worldwide but especially in low-income and middle-income countries” ¹. The best way to cope with the harm due to the increase in case numbers is to develop systems for earlier diagnosis. Early diagnosis systems will have to integrate the growing power of artificial intelligence (AI), including digital pathology (DP) diagnostics, to aid the interpretation of prostate tissue specimens ¹. This contribution aims to point out how DP and AI can help pathologists for the prostate cancer “tsunami” about to come.

Key words: prostate cancer, gleason digital pathology, artificial intelligence, patient advocacy

Digital pathology

Digital pathology started to be developed in the 1980s-1990s. It is based on glass slide scanners designed to obtain virtual slides (VSs) of entire tissue sections ². Virtual slides have three main uses:

- routine histopathologic reporting without microscope;
- inference network systems and automated reasoning systems (i.e., decision support systems, DSSs) for the definition of the Gleason patterns, based on non-numeric (descriptive, linguistic terms) and numeric data;
- AI, based on numeric data derived from virtual slides ³.

Prostate cancer morphology evaluation based on digital pathology and artificial intelligence

Multiple retrospective studies have shown the benefits of AI-based diagnostic solutions for PCa that includes improved cancer detection, quantification, grading, and a potential to reduce pathologists' workload and

Received: June 30, 2024
Accepted: September 3, 2024

Correspondence

Rodolfo Montironi
E-mail: rodolfo.montironi51@gmail.com
Alessia Cimadamore
E-mail: alessiacimadamore@gmail.com

How to cite this article: Cimadamore A, Montironi R, Cheng L, et al. The uropathologist of the future: getting ready with intelligence for the prostate cancer tsunami. *Pathologica* 2024;116:267-272. <https://doi.org/10.32074/1591-951X-1047>

© Copyright by Società Italiana di Anatomia Patologica e Citopatologia Diagnostica, Divisione Italiana della International Academy of Pathology



OPEN ACCESS

This is an open access journal distributed in accordance with the CC-BY-NC-ND (Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International) license: the work can be used by mentioning the author and the license, but only for non-commercial purposes and only in the original version. For further information: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>

enhance pathology laboratory workflow ⁴. Automated deep learning systems can assist pathologists by screening biopsies, providing Gleason score and grade group, and presenting quantitative measurements of volume percentage. Several investigators have dealt with PCa grading based on DP and AI ⁶. In the study of Lucas et al. ⁷, the automatic classification of the Gleason patterns reached an accuracy of 90%, with a specificity and sensitivity of 94% and 77%, respectively, in distinguishing between Gleason patterns 4 or higher and Gleason pattern 3. Concordance of their computer-based GG evaluation with the evaluation made by a pathologist was 65%, i.e., a substantial agreement. AI can create algorithms that can allow a generalist to function as a specialist. The

rate of agreement with subspecialists in grading prostate cancers is proved to be significantly higher for the deep learning system (71.7%) than for general pathologists (58%) ⁸.

More recently, the results of the Prostate cANcer graDe Assessment (PANDA) challenge, a global AI competition, were published. In total, 12,625 publicly available whole-slide images (WSIs) of prostate biopsies were retrospectively collected from 6 different sites for algorithm development, tuning and independent validation. Around 1,290 developers joined the competition to catalyze development of reproducible AI algorithms for Gleason grading. Best algorithms achieved agreements of 0.862 (quadratically weighted κ , 95% confidence interval (CI), 0.840-0.884) and 0.868 (95% CI, 0.835-0.900) with ex-

Table I. Original research papers focusing on AI and Gleason grading with an external validation cohort.

First Author, Year Reference	DL Architecture	Training Cohort	IV Cohort	EV Cohort	Aim	Results
Källén, 2016 ¹⁰	OverFeat	TCGA	10-fold cross val	213 WSIs	GP tile classification Classify WSIs with a majority GP	ACC: 0.81 ACC: 0.89
Bulten, 2019 ¹¹	U-Net	62 WSIs	40 WSIs	20 WSIs	Benign vs. GP	IoU: 0.811 (IV) & 0.735 (EV) F1: 0.893 (IV) & 0.835 (EV)
Silva-Rodriguez, 2020 ¹²	Custom CNN	182 WSIs	5-fold cross-val	703 tiles from 641 TMAs	GP at the level	F1: 0.713 (IV) & 0.57 (EV) qKappa: 0.732 (IV) & 0.64 (EV)
					GG at the WSI level	qKappa: 0.81 (0.77 with method)
					Cribriform pattern detection at the tile level	AUC: 0.822
Nagpal, 2020 ⁸	Xception	524 WSIs	430 WSIs	322 WSIs	Malignancy detection at the WSI level	AUC: 0.981 Agreement: 0.943
					GG1-2 vs. GG3-5	AUC: 0.972 (EV) Agreement: 0.928
Pantanowitz, 2020 ¹³	IBEX	549 WSIs	2501 WSIs	1627 WSIs	Cancer detection at the WSI level	AUC: 0.997 (IV) & 0.991 (EV)
					Low vs. high grade (GS6 vs GS7-10)	AUC: 0.941 (EV)
					GP3/4 vs. GP5	AUC: 0.971 (EV)
					Perineural invasion detection	AUC: 0.957 (EV)

Table I. *Follows from the previous page.*

First Author, Year Reference	DL Architecture	Training Cohort	IV Cohort	EV Cohort	Aim	Results
Ström, 2020 ¹⁴	InceptionV3	6935 WSIs	1631 WSIs	330 WSIs	Malignancy detection at the WSI level	AUC: 0.997 (IV) & 0.986 (EV)
					GG classification	Kappa: 0.62
Bulten, 2020 ¹⁵	Extend Unet	5209 biopsies from 1033 patients	550 biopsies from 210 patients	886 cores	Malignancy detection at the WSI level GG > 2 detection	AUC: 0.99 (IV) & 0.98 (EV) AUC: 0.978 (IV) & 0.871 (EV)
			100 biopsies	None	GG classification	qKappa: 0.819 (general pathologist) & 0.854 (DL on IV) & 0.71 (EV)
Li, 2021 ¹⁶	Weakly supervised VGG11bn	13,115 WSIs	7114 WSIs	79 WSIs	Malignancy of slides	AUC: 0.982 (IV) & 0.994 (EV)
					Low vs. high grade at the WSI level	Kappa: 0.818 Acc: 0.927
Jung, 2022 ¹⁷	DeepDx Prostate	Pre-trained	Pre-trained	593 WSIs	Correlation with reference pathologist (pathology report comparison)	Kappa: 0.654 (0.576) qKappa: 0.904 (0.858)
Bulten, 2022 ⁹	Evaluation of multiple algorithms (PANDA challenge)	10,616 WSIs	545 WSIs	741 patients (EV1) 330 patients (EV2)	GG classification	qKappa: 0.868 (EV2) qKappa: 0.862 (EV1)

IV: Internal Validation. EV: External Validation. TMA: Tissue MicroArray. CNN: Convolutional Neural Network. GG: Gleason ISUP Group, GS: Gleason Score, GP: Gleason Pattern. AUC: Area Under the Curve. ACC: Accuracy. F1: F1-score, combination of precision and recall. IoU: Intersection over Union. q/wKappa: quadratic/weighted Cohen Kappa. Sens: Sensitivity. Spec: Specificity. metric implies the mean of this metric (e.g., AUC).

pert uropathologists on United States and European external validation sets, respectively ⁹. Most relevant studies on AI and Gleason grading with an external validation cohort are reported in Table I.

In 2021, FDA approved the first AI product for digital pathology, namely Paige Prostate. This software showed a sensitivity of 97.7% and positive predictive value of 97.9%, and a specificity of 99.3% and negative predictive value of 99.2% in identifying core biopsies with cancer in a data set derived from an independent institution ¹⁸. However, the FDA specified that pathologists should use Paige Prostate in conjunction with their com-

plete standard of care evaluation of slide images and the software cannot be used as the primary diagnosis.

Critical issues concerning artificial intelligence application in pathology

Most AI applications involve Deep learning (DL) networks, that consists of an input layer, multiple hidden layers, and an output layer, recapitulating the human neural architecture. The hidden layers can recreate newer patches of the image and allow for the differ-

entiation between specific features, and ultimately for image pattern recognitions capable to predict disease diagnosis, prognosis, and therapeutics. These processes work as “black box”, where even the AI’s designers cannot explain why the DL algorithm arrives at a specific decision. To solve this, a new field of AI, named Explainable AI (XAI), is becoming a potential tool to also address the legal aspects of AI implementation. XAI will help users of AI-powered systems perform more effectively by improving their understanding of how those systems reason, creating algorithms that follow the three principles of transparency, interpretability, and explainability. This solution aims to ensure that the system operates in accordance with ethical and legal standards, and that its decision-making processes are transparent and accountable^{19,20}.

Another issue raised against the use of AI in pathology is that it does not provide information about disease understanding. However, the algorithms are generated by data input provided by humans. So far, the research on AI applications in prostate cancer has focused on algorithms generated by input data consisting in digital images of prostatic tissue. The next step might be to generate algorithms based on molecular alterations along with the image of the corresponding tissue. Combining these different types of data may lead to a better understanding of disease biology, and pave the way to a more personalized approach, not only based on morphology, but also on molecular features.

Limited experiences have been reported in the use of AI in evaluation of immunohistochemistry (IHC) in doubtful prostate cases. AI has been used to quantify IHC staining of predictive markers such as HER-2, PD-L1, Ki-67, with excellent results. However, in order to use AI for the differential diagnosis of difficult lesions, i.e. atypical small acinar proliferations, a specific algorithm should be trained with specific image datasets along with the corresponding IHC stainings. If properly trained and validated, the algorithm will recognize such ambiguous lesions.

The future in prostate cancer pathology

The future of PCa pathology, including grading, based on DP and AI is uncharted. For instance, AI could grade PCa in a manner no longer constrained by the limit of established pathologist’s knowledge. The current approach adopted in the AI program relies on features conceived by pathologists. Future AI systems could be entirely the product of AI training and evaluate new tissue morphologic features not included in the Gleason system. The result could be

a correlation with a prognosis far more accurate than when based on the traditional morphologic features. This will bear the stamp of another form of learning and logical evaluation.

Could artificial intelligence replace the pathologist or require a new type of pathologist?

AI will usher pathologists in a world in which decisions, basically diagnoses, are made in three primary ways: by pathologists (which is familiar), by machine (which is becoming familiar), and by collaboration between pathologists and machines (which is not only unfamiliar but also unprecedented).

A fundamental question is whether the pathologist with a traditional type of training can cope successfully with the development and applications of AI in the routine, i.e., a world in which diagnoses are made by a machine or in collaboration with a machine. The pathologist has not received proper training to being involved in the development and utilization of AI systems.

The solution would be that the pathologist receives additional training, whenever possible, in computer science. This is not easy to implement and, in many cases, nor feasible. Another solution would be that additional people are recruited from non-medical fields to work in parallel with pathologists in the developmental phase and validation of the algorithm. This can lead to a new approach to pathology and corresponds to what Dr James ND et al. referred to “incorporate novel mixes of personnel management”¹. With this approach, the risk is that AI systems could replace or limit to a certain extent the work done by pathologists.

How would clinicians use artificial intelligence-generated pathology data?

The clinicians must deal with data originating from multiple sources, including AI-based morphologic diagnostics, radiomics, diagnostic imaging and robotic surgery. They will have to follow a process of merging data “defined as multi-criteria decision making and information fusion”²¹. The resulting information, including the diagnostic and therapeutic decisions, will be far more accurate than when the various sources are evaluated separately and individually”²¹. All this involves the utilization of AI.

Empowering patients

The literature on patients and pathologists and more in general clinicians diagnosing a disease or recommending treatments based on AI is scant. Rodler et al. evaluated patients' trust in AI-based decision-making for localized prostate cancer. They found that AI-assisted physicians (66.74%) were preferred over physicians alone (29.61%), physicians controlled by AI (2.36%), and AI alone (0.64%) for treatment²². Patient advocacy groups should be involved in the development of AI systems to improve diagnoses, prognosis, and personalized treatments, optimizing care, and changing the doctor-patient relationship for the better. Individual patients should also be involved, to ensure that the decisions made by or with AI align with human values and priorities.

Conclusions

To sum up, the requirement for AI and human values to collaborate emerges from their unique and complementary strengths. AI can offer significant analysis and processing capabilities, while humans possess emotional intelligence, creativity, and ethical considerations. By combining these strengths, better outcomes can be achieved for individuals and society as a whole. It is important to remember that humans are the creators of AI, not the other way around.

Message from a patient (Anonymized) with PCa: He would support the adoption of AI if it can help prevent the development of PCa, in addition to improving diagnosis and therapy when already developed.

CONFLICT OF INTEREST STATEMENT

The authors have no relevant conflict of interest to declare.

FUNDING

The authors did not receive any funding.

AUTHORS' CONTRIBUTION

Conceptualization: M.R. and G.G.; data acquisition: A.Ci., C.F.; writing – original draft preparation: A.Ci.; writing – review and editing: M.R. and A.Ci.; and supervision: L.C., E.T.R., A.Cr, and A.L.B. All authors have read and agreed to the published version of the manuscript.

References

- James ND, Tannock I, N'Dow J, et al. The Lancet Commission on prostate cancer: planning for the surge in cases. *Lancet*. 2024;403(10437):1683-1722. [https://doi.org/10.1016/S0140-6736\(24\)00651-2](https://doi.org/10.1016/S0140-6736(24)00651-2).
- Montironi R, Cimadamore A, Scarpelli M, et al. Pathology without microscope: From a projection screen to a virtual slide. *Pathol Res Pract*. 2020;216(11):153196. <https://doi.org/10.1016/j.prp.2020.153196>.
- Soori M, Arezoo B and Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. *Cognitive Robotics*. 2023(3):54-70. <https://doi.org/10.1016/j.cogr.2023.04.001>.
- Satturwar S, Parwani AV. Artificial Intelligence-Enabled Prostate Cancer Diagnosis and Prognosis: Current State and Future Implications. *Adv Anat Pathol*. 2024;31(2):136-144. <https://doi.org/10.1097/PAP.0000000000000425>.
- Cimadamore A, Cheng L, Scarpelli M, et al. Digital diagnostics and artificial intelligence in prostate cancer treatment in 5 years from now. *Transl Androl Urol*. 2021;10(3):1499-1505. <https://doi.org/10.21037/tau-2021-01>.
- Montironi R, Cimadamore A. Considerations on current and future issues related to reproducibility and accuracy in prostate cancer grading. *Virchows Arch*. 2021;478:375-377. <https://doi.org/10.1007/s00428-020-02913-6>.
- Lucas M, Jansen I, Savci-Heijink CD, et al. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. *Virchows Arch*. 2019;475:77-83. <https://doi.org/10.1007/s00428-019-02577-x>.
- Nagpal K, Foote D, Tan F, et al. Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer From Biopsy Specimens. *JAMA Oncol*. 2020;6(9):1372-1380. <https://doi.org/10.1001/jamaoncol.2020.2485>.
- Bulten W, Kartasalo K, Chen PC, et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med*. 2022;28(1):154-163. <https://doi.org/10.1038/s41591-021-01620-2>.
- Kallen H, Molin J, Heyden A, et al. Towards Grading Gleason Score Using Generically Trained Deep Convolutional Neural Networks. *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) 2016*, pp. 1163-1167. <https://doi.org/10.1109/ISBI.2016.7493473>.
- Bulten W, Bándi P, Hoven J, et al. Epithelium Segmentation Using Deep Learning in H&E-Stained Prostate Specimens with Immunohistochemistry as Reference Standard. *Sci Rep*. 2019;9(1):864. <https://doi.org/10.1038/s41598-018-37257-4>.
- Silva-Rodríguez J, Colomer A, Sales MA, et al. Going Deeper through the Gleason Scoring Scale: An Automatic End-to-End System for Histology Prostate Grading and Cribriform Pattern Detection. *Comput Methods Programs Biomed*. 2020;195: <https://doi.org/105637>. 10.1016/j.cmpb.2020.105637.
- Pantanowitz L, Quiroga-Garza GM, Bien L, et al. 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. [https://doi.org/10.1016/S2589-7500\(20\)30159-X](https://doi.org/10.1016/S2589-7500(20)30159-X).
- Ström P, Kartasalo K, Olsson H, et al. Artificial Intelligence for Diagnosis and Grading of Prostate Cancer in Biopsies: A Population-Based Diagnostic Study. *Lancet Oncol*. 2020;21(2):222-232. [https://doi.org/10.1016/S1470-2045\(19\)30738-7](https://doi.org/10.1016/S1470-2045(19)30738-7).
- Bulten W, Pinckaers H, van Boven H, et al. Automated Deep-Learning System for Gleason Grading of Prostate Cancer Using Biopsies: A Diagnostic Study. *Lancet Oncol*. 2020;21(2):233-241. [https://doi.org/10.1016/S1470-2045\(19\)30739-9](https://doi.org/10.1016/S1470-2045(19)30739-9).
- Li J, Li W, Sisk A, et al. A Multi-Resolution Model for Histopathology Image Classification and Localization with Multiple Instance Learning. *Comput Biol Med*. 2021;131:104253. <https://doi.org/10.1016/j.compbiomed.2021.104253>.

- ¹⁷ Jung M, Jin MS, Kim C, et al. Artificial Intelligence System Shows Performance at the Level of Urologists for the Detection and Grading of Prostate Cancer in Core Needle Biopsy: An Independent External Validation Study. *Mod. Pathol.* 2022;35:1449-1457. <https://doi.org/10.1038/s41379-022-01077-9>.
- ¹⁸ Perincheri S, Levi AW, Celli R, et al. An independent assessment of an artificial intelligence system for prostate cancer detection shows strong diagnostic accuracy. *Mod Pathol.* 2021;34(8):1588-1595. <https://doi.org/10.1038/s41379-021-00794-x>.
- ¹⁹ Roscher R, Bohn B, Duarte MF, Garcke J. Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access.* 2020;8:42200-42216. <https://doi.org/10.1109/ACCESS.2020.2976199>.
- ²⁰ Murdoch WJ, Singh C, Kumbier Karl, et al. Interpretable machine learning: definitions, methods, and applications. *Proceedings of the National Academy of Sciences of the United States of America. Proc Natl Acad Sci U S A.* 2019;116(44):22071-22080. <https://doi.org/10.1073/pnas.1900654116>.
- ²¹ Montironi MA. From image analysis in pathology to robotics and artificial intelligence. *Anal Quant Cytopathol Histol.* 2016;38:268-269.
- ²² Rodler S, Kopliku R, Ulrich D, et al. Patients' Trust in Artificial Intelligence-based Decision-making for Localized Prostate Cancer: Results from a Prospective Trial. *Eur Urol Focus.* 2023:S2405-4569(23)00237-7. <https://doi.org/10.1016/j.euf.2023.10.020>.