



OPEN

DATA DESCRIPTOR

The PAR dataset: Prostate biopsy whole slide images from an underrepresented Middle Eastern population

Peshawa J. Muhammad Ali^{1,2,12}, Navin Vincent^{3,12}, Saman S. Abdulla^{4,5}, Han N. Mohammed Fadhl⁶, Anders Blilie^{7,8}, Kelvin Szolnok⁹, Julia Anna Mielcarz³, Xiaoyi Ji⁹, Kimmo Kartasalo^{3,10,11}, Abdulbasit K. Al-Talabani^{1,13} & Nita Mulliqi^{3,13} ✉

Artificial intelligence (AI) is increasingly used in digital pathology. Publicly available histopathology datasets remain scarce, and those that do exist predominantly represent Western populations. Consequently, the generalizability of AI models to populations from less digitized regions, such as the Middle East, is largely unknown. This motivates the public release of our dataset to support the development and validation of pathology AI models across globally diverse populations. We present 1,017 whole slide images by digitizing 339 glass slides of prostate core needle biopsies from a consecutive series of 185 patients collected in Erbil, Iraq. Each glass slide was scanned by three different whole-slide scanners. The dataset also includes the corresponding Gleason Scores and the International Society of Urological Pathology grades assigned independently by three pathologists. All slides were de-identified and are provided in their native formats without further conversion. The dataset enables grading concordance analyses, color normalization, and cross-scanner robustness evaluations. The dataset is publicly available through the BioImage Archive and released under the CC BY 4.0 license.

Background & Summary

Prostate cancer diagnosis relies on microscopic assessment of hematoxylin and eosin (H&E) stained biopsy slides. In routine practice, core needle biopsy is the most common method for diagnosing prostate cancer, where a needle is used to take a small tissue sample from the prostate, guided by ultrasound and potentially MRI. After extraction, tissue cores are processed through routine histopathological steps to produce a glass slide for each biopsy. Each slide is subsequently reviewed and graded by a pathologist according to the Gleason scoring (GS) system¹. The GS consists of primary and secondary grades based on the proportion of tissue with different Gleason patterns, e.g., a score of $4 + 3 = 7$, where the first number represents the predominant pattern². In 2014, the International Society of Urological Pathology (ISUP) introduced the ISUP grades, pooling GS into five ordinal categories, from 1 to 5, with grade 1 being the least aggressive and grade 5 the most aggressive cancer³. Both GS and ISUP grades remain in parallel use in clinical practice. Despite the central clinical role of Gleason grading, its reproducibility remains limited. Numerous studies have demonstrated substantial intra- and inter-pathologist variability leading to under- and over-treatment of patients^{4–6}. Digital pathology and artificial

¹Department of Software Engineering, Faculty of Engineering, Koya University, 44023, Kurdistan Region, Koya, Iraq. ²Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Koya University, 44023, Kurdistan Region, Koya, Iraq. ³Department of Medical Epidemiology and Biostatistics, SciLifeLab, Karolinska Institutet, Stockholm, Sweden. ⁴College of Dentistry, Hawler Medical University, Erbil, Kurdistan Region, Iraq. ⁵PAR Private Hospital, Erbil, Kurdistan Region, Iraq. ⁶College of Dentistry, University of Sulaimani, Sulaymaniyah, Kurdistan Region, Iraq. ⁷Department of Pathology, Stavanger University Hospital, Stavanger, Norway. ⁸Faculty of Health Sciences, University of Stavanger, Stavanger, Norway. ⁹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. ¹⁰ELLIS Institute Finland, Espoo, Finland. ¹¹Department of Computing, Faculty of Technology, University of Turku, Turku, Finland. ¹²These authors contributed equally: Peshawa J. Muhammad Ali, Navin Vincent. ¹³These authors jointly supervised this work: Abdulbasit K. Al-Talabani, Nita Mulliqi. ✉e-mail: nita.mulliqi@ki.se

intelligence (AI)-based analysis of whole-slide images (WSIs) offer opportunities to improve diagnostic consistency⁷, and many studies now show that AI systems can diagnose and grade prostate cancer with performance comparable to expert pathologists^{8–11}.

To further accelerate AI development, several efforts have been made to release publicly available digitized prostate pathology datasets. The PANDA dataset released 5,160 digitized core needle biopsies from the Netherlands and 5,456 from Sweden with associated ISUP grades and, in some cases, tissue masks annotating the cancerous regions¹². The SPROB20 dataset released 2,611 digitized biopsies collected in Sweden with associated ISUP and other clinical variables¹³. The SICAPv2 published 155 digitized biopsies collected in Spain with associated GS, provided as .jpg patches instead of the full WSIs^{14,15}. The UKK and WNS datasets comprise 100 digitized biopsies collected in Germany and Austria, with associated ISUP grades from 10 different pathologists. These two datasets are ISUP-balanced and do not contain benign samples¹⁶. The DiagSet dataset includes digitized prostate biopsies collected in Poland as (i) 2.6 million patches with patch-level custom labels (e.g., benign, cancer of lower/higher severity), (ii) 4,675 WSIs with cancer/non-cancer labels, and (iii) 46 WSIs labeled with cancer/non-cancer/ambiguous (requires immunohistochemistry (IHC)) labels provided by nine pathologists¹⁷. The Gleason grading challenge 2022 includes 144 digitized prostatectomies and 53 biopsies collected in Singapore, with associated Gleason pattern pixel-level annotations¹⁸. The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA-PRAD) project released 499 digitized prostatectomies collected in the US, accompanied by GS, genomics, and other clinical data^{19,20}. The EMPaCT tissue microarray (TMA) dataset comprises 210 digitized TMAs collected in Switzerland, associated with binary labels for overall survival and disease progression status. It also includes consecutive tissue sections of H&E staining and additional immunohistochemical markers. The authors released the tissue cores as .png images²¹. Zhong *et al.*²² introduced a Swiss dataset derived from 73 unique patients comprising 71 TMA cores, two WSIs and 6,000 derived patches, with consensus annotations from two pathologists. The Tissue Microarrays Zürich (TMAZ) includes five prostate cancer TMAs, each containing approximately 200–300 cores from prostatectomy specimens collected in Switzerland, each associated with Gleason pattern annotations from two pathologists²³. The Gleason 2019 Challenge dataset released 331 TMA cores collected in Canada as .jpg images²⁴ with six separate expert annotation maps. In addition to these publicly available resources, many AI studies rely on proprietary industry datasets⁸ or large single-institution archives^{25,26}, which are not publicly accessible.

Despite efforts in the field to publish publicly available data, challenges still remain. First, given the substantial variability in Gleason grading, studies need to rely on datasets that provide reference grading from more than one pathologist. The available datasets typically contain single reference standards. This severely limits reproducibility and benchmarking. Secondly, most AI studies are developed and validated predominantly in Caucasian populations. This raises concerns about model generalization to underrepresented regions where pathology services are not yet digitized. Full workflow digitization requires substantial financial investment, limiting its feasibility in low-resource or low-volume settings. This represents a chicken-and-egg problem—lack of digital pathology precludes the collection of data and evidence for the effectiveness and generalizability of AI pathology in these populations, and the absence of this evidence in part contributes to the lack of investment in digitization. One potential solution would be to use cost-effective compact scanners for validation and pilot testing of AI in currently non-digital settings. No public datasets currently include data digitized with compact scanners or originating from underrepresented populations. Thirdly, numerous studies have demonstrated that AI model performance varies across whole-slide scanners and that scanner-induced differences impact model generalization^{27,28}. Existing public datasets typically provide WSIs from only one scanner type, limiting the ability to study cross-scanner robustness.

Together, these gaps underscore the need for datasets that enable cross-population, multi-scanner, and multi-pathologist AI validation to support broad clinical adoption. To address these gaps, we present 1,017 WSIs of prostate core needle biopsies derived from 339 glass slides from 185 patients collected at the PAR Hospital in Erbil, Iraq. Each glass slide was digitized using three different scanners, resulting in three scan versions per glass slide. The consecutive patient series is representative of the cases encountered in routine clinical practice between 2013–2024. The slides are associated with slide-level GS and ISUP grades assigned independently by three pathologists blinded to each other and to any AI grading. The slides were digitized using three different whole-slide scanners: Leica Aperio GT 450 DX, Hamamatsu NanoZoomer HT 2.0, and the Grundium Ocus40 compact scanner. This work also addresses the broader benchmarking crisis in biomedical AI, where progress is constrained by limited public datasets, institution-specific biases, and inconsistent validation standards²⁹. By making this dataset available, we aim to help bridge long-standing inequities in both population representation and technological access, moving the field closer to inclusive and reliable AI for pathology.

Methods

Ethical considerations. The patient sample collection was approved by the Ethical Committee of PAR Hospital in Kurdistan, Iraq (permit 1002/07072024). Digital scanning and processing of the samples at Karolinska Institutet was approved by the Swedish Ethical Review Authority (permit 2019-05220). The samples were de-identified before shipping to Karolinska Institutet. Informed consent was waived by the ethical review board due to the use of de-identified prostate specimens in a retrospective setting.

Study design. This work consists of four main steps: slide selection, slide preparation in the laboratory, whole slide scanning, and annotation by pathologists to establish a reference standard.

Slide selection. The study cohort was derived from the pathology archive at the PAR Hospital (Erbil), which contained 30,056 cases assessed between 2013 and 2024. The archive reflects consecutive routine diagnostic cases. The study cohort was extracted by multiple manual steps. An initial filtering of this digital archive,

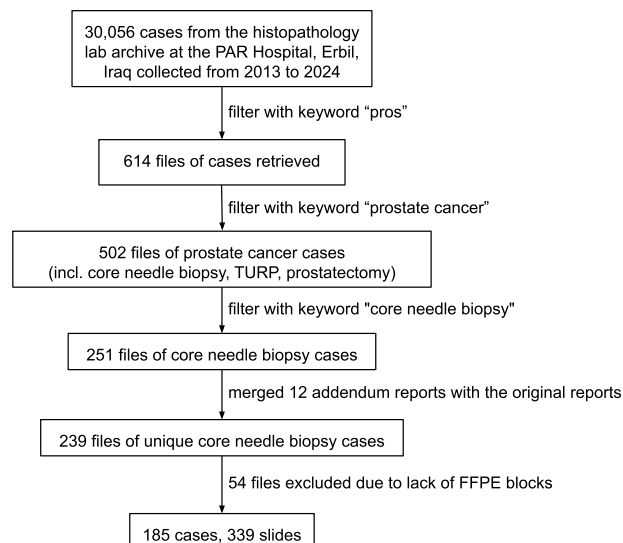


Fig. 1 Diagram illustrating dataset curation. The diagram outlines the process of identifying 30,056 pathology cases from the PAR Hospital archive, filtering prostate-related reports, removing addenda, confirming availability of archival FFPE blocks, and deriving the final set of 185 eligible prostate core needle biopsy cases for slide preparation and scanning.

consisting of reports in .docx format, was performed using the keyword “pros”. This primary search yielded 614 potentially relevant files. This subset underwent further cleaning to identify 502 relevant prostate cases, which included needle biopsies, transurethral resections of the prostate, and prostatectomies. For this study, the focus was narrowed to the 251 cases identified as “core needle biopsy files”. This group was refined further by merging 12 pairs (24 files) identified as “addenda”. These addenda files, which noted additional tests like IHC, were merged to avoid duplication, leaving 239 unique needle biopsy cases. The final eligibility criterion was the availability of the original formalin-fixed and paraffin-embedded (FFPE) tissue blocks in the lab archive. It was found that 54 tissue blocks were lost (mostly from the early years of the collection). After confirming eligibility, the final dataset consisted of 185 prostate cancer needle biopsy cases, from which new H&E slides were prepared from the existing blocks. In summary, the eligibility criteria were: (a) cases identified as prostate core needle biopsy, (b) non-duplicate/non-addenda reports, and (c) availability of the archival FFPE block (see Fig. 1).

Slide preparation. The standard biopsy protocol at the institution involved collecting six cores from each side of the prostate, though core fragmentation was occasionally observed during extraction. If biopsies from each side were submitted in separate containers, separate slides were generated. Conversely, if biopsies from both sides were submitted in a single container, they were combined to produce a single slide. After reception of the samples in the laboratory, they were placed in cassettes and underwent standard processing, including fixation in formalin to preserve morphology, dehydration through graded alcohols, clearing (often with xylene), and embedding in paraffin wax to form solid blocks (FFPE). These FFPE blocks were subsequently sectioned with a microtome into thin slices, typically 3–4 μm thick, with sections obtained at multiple levels to ensure adequate tissue representation. Sections were checked for quality, and additional deeper levels were obtained where it was deemed necessary by the pathologist. For 154 patients, biopsies from the left and right sides were submitted in separate containers, resulting in two slides per patient, whereas for the remaining 31 patients, both sides were submitted in a single container and therefore produced one slide per patient. The slides were stained with H&E and were finally coverslipped.

For this study, the original archived diagnostic H&E slides were not used, as many exhibited significant stain fading and degradation due to prolonged storage. Therefore, to ensure consistent and high-quality staining, all retrieved FFPE blocks were recut to achieve new sections, with prior re-embedding performed when deemed necessary, to generate fresh slides for analysis. The new slides were screened and assessed independently of the original pathology report.

Whole slide image scanning. All glass slides were scanned at the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, using three whole-slide scanners, Hamamatsu NanoZoomer 2.0 HT, Leica Aperio GT 450 DX and a compact, single-slide scanner, Grundium Ocus40. Slides were digitized by trained personnel using an automated scanning workflow. Manual rescanning of slides was done in cases where automated focusing was not successful. All slides were scanned at 40x, with resolutions of 0.22 $\mu\text{m}/\text{pixel}$ for Hamamatsu, 0.26 $\mu\text{m}/\text{pixel}$ for Aperio, and 0.25 $\mu\text{m}/\text{pixel}$ for Grundium. WSIs were stored in the scanners’ proprietary formats: .ndpi for Hamamatsu and .svs for Grundium and Aperio. See Fig. 2 for example WSIs of the dataset³⁰.

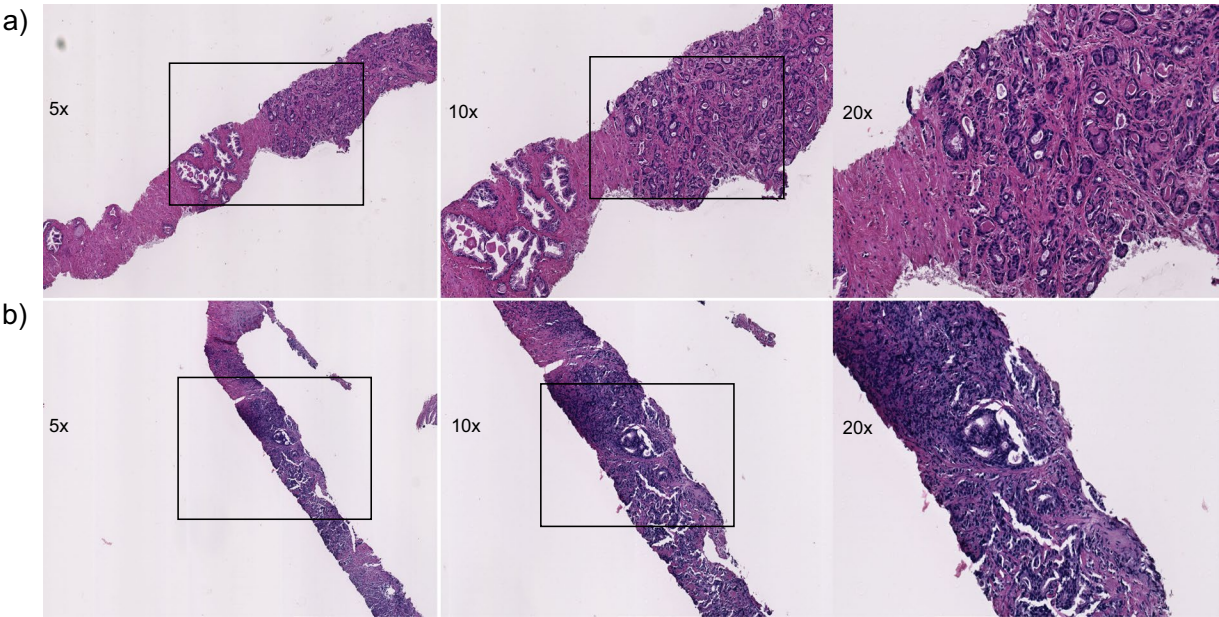


Fig. 2 Example whole slide images. The figure illustrates two prostate core needle biopsy slides (a,b) digitized with the Hamamatsu NanoZoomer 2.0 HT whole slide scanner. For each slide, representative zoom-in regions highlight tissue morphology at higher magnification. The respective slides and regions were selected by the pathologist (H.M.). The maximum available magnification, 40x, corresponds to 0.2266 micrometers per pixel.

Number of patients		Number of slides			
n = 185		n = 339			
Age, years		ISUP grade	Pathologist I	Pathologist II	Pathologist III
<= 49 yrs	5	Benign	164	208	11
50–54 yrs	9	ISUP 1 (3 + 3)	7	2	5
55–59 yrs	14	ISUP 2 (3 + 4)	38	7	13
60–64 yrs	26	ISUP 3 (4 + 3)	59	8	3
65–69 yrs	49	ISUP 4 (4 + 4, 3 + 5, 5 + 3)	30	25	6
>= 70 yrs	79	ISUP 5 (4 + 5, 5 + 4, 5 + 5)	41	86	21

Table 1. Cohort characteristics. The table outlines age distribution across the 185 patients, and ISUP grade distribution across 339 prostate biopsies from all pathologists.

Reference standard protocol. The WSIs were independently assessed by three pathologists. The first was S.A., a general pathologist with 22 years of experience and a special interest in prostatic pathology. He works as a lecturer at Hawler Medical University and as a consultant pathologist at PAR Hospital in Erbil, Iraq, and serves as a senior examiner for the Royal College of Pathologists, UK. The second, H.M., is a general pathologist with 15 years of experience, currently working at the University of Sulaimani in Sulaymaniyah, Iraq. The third, A.B., is a uropathologist with 6 years of experience at Stavanger University Hospital in Stavanger, Norway. The pathologists assessed the slides digitally by using Cytomine³¹ and assigned slide-level GSs (using 0 + 0 for benign slides). Scores were subsequently converted to ISUP grades and stored with the corresponding slide IDs. No pixel-level annotations for outlining regions of interest were marked during the assessment. The first pathologist (S.A.) graded all 339 slides, the second pathologist (H.M.) graded 337 slides, and the third uropathologist (A.B.) graded a random subset of 59 slides stratified by ISUP grades assigned by the first pathologist (see Table 1). Assessments were fully independent with no immediate discussion or reconciliation among pathologists. The distinction between uropathologists and general pathologists is not consistently defined across countries and clinics, and there are substantial differences in subspecialty training requirements and clinical responsibilities. This should be taken into account when interpreting agreement rates between general and specialized pathologists.

Data Records

The PAR dataset³⁰ has been deposited to the BioImage Archive under accession number S-BIAD2323 (<https://doi.org/10.6019/S-BIAD2323>). The dataset is deposited as a single root folder, “iraq_prostate_wsi”, containing six subfolders: “component_A_grundium_svs”, “component_B_hamamatsu_ndpi”, “component_C_leica_svs”, “annotations”, “file_lists”, and “docs”. The three component folders each hold 339 WSIs collected from 185 patients, where 154 have two slides, and 31 have only one slide per patient. Filenames share a consistent, non-identifying scheme: *c* < slide_id > < a | b | none > . < ext > , where *a* or *b* indicates when a patient has right/left slides, and

is omitted when only one slide exists, and the extension reflects the scanner format. The annotations folder provides a single table that lists slide names and the associated Gleason scores and ISUP grades assigned independently by the three pathologists. GS were assigned using ten classes (0 + 0, 3 + 3, 3 + 4, 4 + 3, 4 + 4, 3 + 5, 5 + 3, 4 + 5, 5 + 4, 5 + 5), and ISUP grades show the GS grouped in 6 ordinal values as: ISUP 0 (GS 0 + 0), ISUP 1 (GS 3 + 3), ISUP 2 (GS 3 + 4), ISUP 3 (GS 4 + 3), ISUP 4 (GS 4 + 4 or 3 + 5 or 5 + 3), and ISUP 5 (GS 4 + 5, 5 + 4, or 5 + 5). The “file_lists” folder documents how files are linked or paired across scanners, and “docs” contains dataset documentation such as the license and README files. No personal identifiers are included; file and label names are sufficient for users to pair slides with labels and to link the three scanner versions of each slide.

Technical Validation

We verified technical quality through practical, user-facing checks:

- (i). Completeness: each scanner component folder contains 339 WSIs, matching 185 patients, 154 with suffix *a* or *b*, and 31 single-slide cases.
- (ii). Openability: tissue was segmented using deep-learning-based algorithms, and patches were generated from all WSIs. This process involved using OpenSlide³² to access the WSIs and read the full image area. No issues were encountered when opening either file format (.svs or .ndpi). In addition, all WSIs were successfully viewed at multiple magnification levels in Cytomine during grading and separately in the Automated Slide Analysis Platform (ASAP)³³.
- (iii). Annotation conformance: the annotations table uses a fixed dictionary with 10 GS classes and 6 ISUP grades.

Usage Notes

Traditional image viewers cannot read WSIs; native formats like .svs and .ndpi can be opened with specialized viewers and toolkits such as Cytomine³¹, OpenSlide³², ASAP³³, and QuPath³⁴. While the annotation table does not contain a consensus label from the three pathologists, users can choose their own consensus rule explicitly, for instance, the majority, or the highest grade. The typical preprocessing for modeling includes foreground masking, tiling at different sizes, and stain/illumination normalization³⁵.

Data availability

All data supporting this publication are openly available through the BioImage Archive (BIA) under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. The dataset is delivered as a single folder named “iraq_prostate_wsi”, which contains six subfolders: three component folders with the WSIs from each scanner, an annotations folder with Gleason score and ISUP grade assignments from three pathologists, a “file_lists” folder documenting cross-scanner image pairing, and a “docs” folder with the dataset documentation. A detailed description of these components is provided in the Data Records section. The complete dataset is deposited in BIA under accession number S-BIAD2323 and can be accessed at: (<https://identifiers.org/biostudies:S-BIAD2323>).

Code availability

No custom code was used to generate or process this dataset. Slides were produced with standard laboratory scanner software. No analysis scripts or preprocessing pipelines are required for reuse.

Received: 4 December 2025; Accepted: 29 June 2026;

Published online: 16 July 2026

References

1. Gleason, D. F. Histologic grading of prostate cancer: a perspective. *Hum. Pathol.* **23**, 273–279, [https://doi.org/10.1016/0046-8177\(92\)90108-f](https://doi.org/10.1016/0046-8177(92)90108-f) (1992).
2. Epstein, J. I., Allsbrook, W. C. Jr., Amin, M. B., Egevad, L. L. & ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma. *Am. J. Surg. Pathol.* **29**, 1228–1242, <https://doi.org/10.1097/01.pas.0000173646.99337.b1> (2005).
3. Epstein, J. I. *et al.* The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am. J. Surg. Pathol.* **40**, 244–252, <https://doi.org/10.1097/PAS.0000000000000530> (2016).
4. Melia, J. *et al.* A UK-based investigation of inter- and intra-observer reproducibility of Gleason grading of prostatic biopsies. *Histopathology* **48**, 644–654, <https://doi.org/10.1111/j.1365-2559.2006.02393.x> (2006).
5. Egevad, L. *et al.* Standardization of Gleason grading among 337 European pathologists. *Histopathology* **62**, 247–256, <https://doi.org/10.1111/his.12008> (2013).
6. Ozkan, T. A. *et al.* Interobserver variability in Gleason histological grading of prostate cancer. *Scand. J. Urol.* **50**, 420–424, <https://doi.org/10.1080/21681805.2016.1206619> (2016).
7. Pantanowitz, L. *et al.* Twenty years of digital pathology: An overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. *J. Pathol. Inform.* **9**, 40, https://doi.org/10.4103/jpi.jpi_69_18 (2018).
8. Campanella, G. *et al.* Clinical-grade computational pathology using weakly supervised deep learning on whole-slide images. *Nat. Med.* **25**, 1301–1309, <https://doi.org/10.1038/s41591-019-0508-1> (2019).
9. Ström, P. *et al.* Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* **21**, 222–232, [https://doi.org/10.1016/S1470-2045\(19\)30738-7](https://doi.org/10.1016/S1470-2045(19)30738-7) (2020).
10. Bulten, W. *et al.* Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol.* **21**, 233–241, [https://doi.org/10.1016/S1470-2045\(19\)30739-9](https://doi.org/10.1016/S1470-2045(19)30739-9) (2020).
11. Mulliqi, N. *et al.* Foundation models – a panacea for artificial intelligence in pathology? *Preprint at* <https://arxiv.org/abs/2502.21264> (2025).
12. Bulten, W. *et al.* Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat. Med.* **28**, 154–163, <https://doi.org/10.1038/s41591-021-01620-2> (2022).
13. Walhagen, P. *et al.* Spear Prostate Biopsy 2020 (SPROB20). *AIDA Data Hub* <https://doi.org/10.23698/aida/sprob20> (2020).

14. Silva-Rodríguez, J. SICAPv2 – Prostate whole slide images with Gleason grades annotations. *Mendeley Data*. <https://doi.org/10.17632/9xxm58dvs3.1> (2020).
15. Silva-Rodríguez, J., Colomer, A., Sales, M. A., Molina, R. & Naranjo, V. 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.* **195**, 105637, <https://doi.org/10.1016/j.cmpb.2020.105637> (2020).
16. Tolkach, Y. *et al.* An international multi-institutional validation study of the algorithm for prostate cancer detection and Gleason grading. *NPJ Precis. Oncol.* **7**, 77, <https://doi.org/10.1038/s41698-023-00424-6> (2023).
17. Koziarski, M. *et al.* DiagSet: a dataset for prostate cancer histopathological image classification. *Sci. Rep.* **14**, 6780, <https://doi.org/10.1038/s41598-024-52183-4> (2024).
18. Huo, X. *et al.* A comprehensive AI model development framework for consistent Gleason grading. *Commun. Med.* **4**, 84, <https://doi.org/10.1038/s43856-024-00502-1> (2024).
19. The Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell* **163**, 1011–1025, <https://doi.org/10.1016/j.cell.2015.10.025> (2015).
20. Heath, A. P. *et al.* The NCI Genomic Data Commons. *Nat. Genet.* **53**, 257–262, <https://doi.org/10.1038/s41588-021-00791-5> (2021).
21. Karkampouna, S. & Kruithof-de Julio, M. Dataset EMPaCT TMA. *Zenodo*. <https://doi.org/10.5281/zenodo.10066853> (2023).
22. Zhong, Q. *et al.* A curated collection of tissue microarray images and clinical outcome data of prostate cancer patients. *Sci. Data* **4**, 170014, <https://doi.org/10.1038/sdata.2017.14> (2017).
23. Arvaniti, E. *et al.* Replication Data for: Automated Gleason grading of prostate cancer tissue microarrays via deep learning. *Harvard Dataverse*. <https://doi.org/10.7910/DVN/OCYCMP> (2018).
24. MICCAI 2019 Gleason Challenge Organizers. Gleason 2019: Automatic Gleason grading of prostate cancer in digital histopathology. *MICCAI 2019 Conference - Challenge Website* <https://gleason2019.grand-challenge.org/> (2019; accessed 28 November 2025).
25. Chen, R. J. *et al.* Towards a general-purpose foundation model for computational pathology. *Nat. Med.* **30**, 850–862, <https://doi.org/10.1038/s41591-024-02857-3> (2024).
26. Vorontsov, E. *et al.* A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat. Med.* **30**, 2924–2935, <https://doi.org/10.1038/s41591-024-03141-0> (2024).
27. Duenweg, S. R. *et al.* Whole slide imaging (WSI) scanner differences influence optical and computed properties of digitized prostate cancer histology. *J. Pathol. Inform.* **14**, 100321, <https://doi.org/10.1016/j.jpi.2023.100321> (2023).
28. Ji, X. *et al.* Physical color calibration of digital pathology scanners for robust artificial intelligence-assisted cancer diagnosis. *Mod. Pathol.* **38**, 100715, <https://doi.org/10.1016/j.modpat.2025.100715> (2025).
29. Mahmood, F. A benchmarking crisis in biomedical machine learning. *Nat. Med.* **31**, 1060, <https://doi.org/10.1038/s41591-025-03637-3> (2025).
30. Muhammad Ali, P. J. The PAR dataset: Prostate biopsy whole slide images from an underrepresented Middle Eastern population. *BioImage Archive*. <https://doi.org/10.6019/S-BIAD2323> (2025).
31. Marée, R. *et al.* Collaborative analysis of multi-gigapixel imaging data using Cytomine. *Bioinformatics* **32**, 1395–1401, <https://doi.org/10.1093/bioinformatics/btw013> (2016).
32. Goode, A., Gilbert, B., Harkes, J., Jukic, D. & Satyanarayanan, M. OpenSlide: A vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* **4**, 27, <https://doi.org/10.4103/2153-3539.119005> (2013).
33. Litjens, G. ASAP – Automated Slide Analysis Platform. *GitHub* <https://computationalpathologygroup.github.io/ASAP> (2017).
34. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878, <https://doi.org/10.1038/s41598-017-17204-5> (2017).
35. Lu, M. Y. *et al.* Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat. Biomed. Eng.* **5**, 555–570, <https://doi.org/10.1038/s41551-020-00682-w> (2021).

Acknowledgements

We thank PAR Hospital histopathology lab for their support in preparing slides. High-performance computing was supported by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) at C3SE partially funded by the Swedish Research Council through grant agreement no. 2022-06725 and no. 2018-05973, by the supercomputing resource Berzelius provided by the National Supercomputer Centre at Linköping University and the Knut and Alice Wallenberg Foundation.

Author contributions

Study conception and design, software and data handling, operational workflows, and manuscript drafting: P.J.M.A., N.V., S.S.A., H.N.M.F., K.K., A.K.A.-T., N.M. Laboratory and pathology activities and grading: S.S.A. (full slide preparation and grading for all slides), H.N.M.F. (graded 337 slides), A.B. (graded 59 slides). K.S. set up and maintained Cytomine, J.A.M. did scanning and data management, and X.J. did data management and processing of the slides. All authors reviewed the manuscript, and K.K., A.K.A. and N.M. supervised the work.

Funding

K.K. received funding from the SciLifeLab & Wallenberg Data-Driven Life Science Program (KAW 2024.0159), the Instrumentarium Science Foundation, and the Karolinska Institutet Research Foundation. Open access funding provided by Karolinska Institute.

Competing interests

K.K. and N.M. are shareholders of Clinsight AB. Other authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to N.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026