ACE BRIEF FOR NEW AND EMERGING HEALTH TECHNOLOGIES

Paige Prostate Detect to Assist the Diagnosis of Prostate Cancer

Document Number: HSB-M 07/2023

Date: May 2023

ACE agency for care effectiveness

Summary of Key Points

- Prostate cancer (PCa) is one of the most common malignancies and a leading cause of death in men globally. It is the second most common cancer in Singapore.
- Prostate needle core biopsy is the gold standard for the diagnosis of PCa although it is limited by grading subjectivity, human error and scarcity of pathologists in various healthcare systems.
- Paige Prostate Detect (PPD; Paige.AI, Inc.) is an artificial intelligence (AI) software that can support pathologists in identifying suspicious foci for cancer during review of whole slide images (WSIs) from prostate needle biopsies.
- Overall, PPD was found to be safe and may improve sensitivity of diagnoses, turnaround time and healthcare resource utilisation.
    - There were no major safety issues related to the use of the software.
    - Pathologists assisted by PPD exhibited good accuracy (sensitivity, 90% to 96.6%; specificity, 92.8% to 98%) in the diagnosis of PCa. Compared to unassisted reads, pathologists assisted by PPD reported significant sensitivity gain in some studies ($p<0.001$). However, this was not consistently demonstrated, indicating the need for further validation on accuracy.
    - There is some evidence that significant sensitivity gains were observed in both highly experienced genitourinary (GU) and non-GU pathologists, and across all tumour grades and sizes.
    - Compared to unassisted reads, there is some evidence indicating the ability of PPD to improve turnaround time by 21% to 65.5%, although the impact of the software on pathologist efficiency remains ambiguous.
    - The potential healthcare system benefits reported included reduced resource consumption, such as immunohistochemistry (36% *vs.* 46%; $p<0.001$) and second opinion requests (7% *vs.* 12%; $p=0.006$).
- Key limitations include the lack of prospective validation when implemented into routine clinical practice and unclear impact of the software on patient outcomes.
- The cost-effectiveness of PPD remains uncertain, with the National Institute for Health and Care Excellence (NICE) indicating a potential for it to release resources and produce overall cost savings.
- Based on a subscription-based service, the cost per slide starts at around £1 (S$2) with a one-time integration fee starting from £15,000 (S$25,000).
- Key implementation considerations include proper information technology (IT) infrastructure, shift towards digital pathology, oversight of AI medical devices and pathologists' training and acceptance of PPD.

I. Background

Prostate cancer (PCa) is an adenocarcinoma that develops primarily in the glandular part of the prostate. The tumour initially spreads to the adjacent prostate tissues, and may remain localised within the prostate for decades (where it is considered curable), or metastasise to the bone and lymph nodes.[1] Major risk factors of PCa include older age, ethnicity, obesity and family history.[1] Early PCa usually presents as asymptomatic, with most PCa cases being low-grade with relatively low risk and limited aggressiveness.[1] Late symptoms may manifest in patients with spinal metastasis including bone pain and paralysis. Renal failure may occur in those with bilateral ureteral obstruction.[1]

Globally, PCa remains one of the most common malignancies and a leading cause of death in men.[2] In Singapore, it is the second most common cancer in men, with 6,283 cases diagnosed from 2016 to 2020.[3] PCa typically occurs in men over the age of 50, with cancer aggressiveness decreasing with age although it can be aggressive in younger men.[1,3] Patients with localised PCa have a five-year survival rate of more than 99%. This is reduced to around 30% in those with distant metastasis.[4]

The gold standard diagnosis of PCa is prostate needle core biopsy, which involves manual interpretation of biopsy slides by pathologists. This presents several limitations, including inter- and intra-reader variability due to grading subjectivity, and human error. The latter can lead to missed or misinterpreted foci and subsequently discrepancies in diagnosis and treatment plans.[5] In addition, manual interpretation of biopsy slides requires highly skilled pathologists which can be costly. The scarcity of pathologists in many healthcare systems further limits the throughput of slides that can be interpreted in a timely fashion.[5] As such, there is a clinical unmet need for an objective, reproducible and efficient method of analysing biopsy slides that allows better management of PCa.

II. Technology

Paige Prostate Detect (PPD; Paige.AI, Inc.) is an artificial intelligence (AI)-based software that can support pathologists in identifying suspicious foci for cancer during review of whole slide images (WSIs) from prostate needle biopsies.[6] The software automatically analyses WSIs and detects a region of interest with the highest likelihood of harbouring cancer for further review by a pathologist, while providing a binary classification of suspicious or not suspicious for cancer (Figure 1).[6]

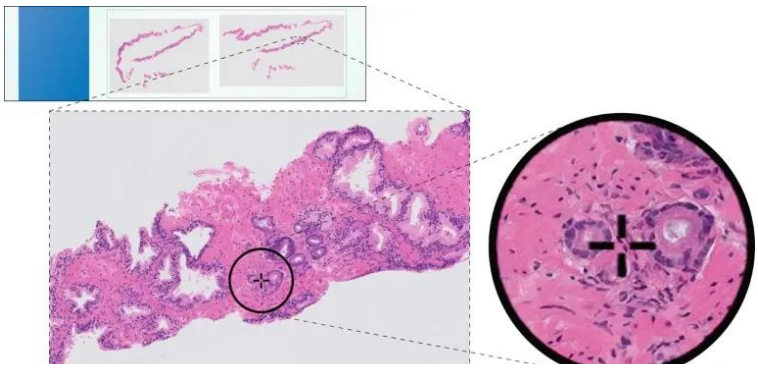

Figure 1: Illustration of Paige Prostate. Paige Prostate identifying an area of prostate tissue likely for harbouring cancer. The software allows pathologists to view digitised versions of traditional glass tissue slides (top left) and identify foci that could indicate cancer (bottom right). Image adapted from https://www.nature.com/articles/d43747-021-00041-x

PPD is intended to be used with slide images digitised with Philips Ultrafast Scanner in the US or the Philips Ultrafast and Leica AT2 scanners in the European Union and UK, and visualised with Paige FullFocus WSI viewing software.[6] PPD is offered as a software as a service (SaaS) product that involves a cloud-based system. Digital images of histopathology slides are stored on the cloud and outputs from the AI software are displayed to pathologists through the Paige FullFocus web-based WSI viewer.[7] It is important to note that PPD should not be used as the sole basis for diagnosis and must be used alongside a comprehensive standard of care evaluation of the WSI by pathologists.[8]

PPD represents one of three software modules in the Paige Prostate suite. The other two are Paige Prostate Grade and Quantify for automatic grading with Gleason scoring, and Paige Prostate Perineural Invasion (PNI) to identify the presence of suspicious foci around nerve fibres within the prostate to detect for PNI in prostate biopsies.[6] However, these are outside the scope of this brief.

PPD represents an innovative technology that can augment the traditional slide interpretation workflow performed by pathologist. It has the potential to reduce subjectivity, lower variability and increase speed of grading compared to unassisted reads without AI.

III. Regulatory and Subsidy Status

In March 2019, Paige.AI was granted the breakthrough device designation by the US Food and Drug Administration (FDA) for its AI system in cancer diagnosis.[8] PPD has received de novo clearance (DEN200080) from the FDA in September 2021 and has been Conformité Européene (CE) and UK Conformity Assessed (UKCA) marked.

On the other hand, the other two modules (Paige Prostate Grade and Quantify and Paige Prostate PNI) are CE and UKCA marked.[6]

IV. Stage of Development in Singapore

☒	Yet to emerge	☐	Established
☐	Investigational / Experimental (subject of clinical trials or deviate from standard practice and not routinely used)	☐	Established *but* modification in indication or technique
☐	Nearly established	☐	Established *but* should consider for reassessment (due to perceived no/low value)

V. Treatment Pathway

According to the Pan-Asian adapted European Society for Medical Oncology (ESMO) guidelines for the management of PCa,[9] the diagnostic workup for patients suspected with PCa involves a multi-step process. First, an assessment of prostate-specific antigen (PSA) level is performed, with patients harbouring elevated PSA levels from repeated tests referred for multiparametric magnetic resonance imaging (mpMRI) to identify the presence of PCa.

Findings from mpMRI scans are assessed with the Prostate Imaging Reporting and Data System (PI-RADS) scoring that determines the likelihood of clinically significant PCa (score of 1 [very low] to 5 [very high]), with no further testing required for those with PI-RADS score of ≤2. Taking into consideration several diagnostic factors such as digital rectal examination findings, ethnicity, age and history of previous biopsy, patients with PI-RADS score of ≥3 (intermediate to high likelihood of clinically significant PCa) are referred for biopsy. The biopsy slides are manually assessed by pathologists for any abnormal findings, with those with foci indicative of cancer referred for further staging and risk assessment for optimal treatment.

The integration of AI-based solutions, such as PPD, into clinical workflows can enable a synergistic PCa care.[10] The software serves as a complementary tool in addition to standard care, supporting pathologists in detecting suspicious foci on WSI that are indicative of cancer.

VI. Summary of Evidence

The assessment was conducted based on the Population, Intervention, Comparator and Outcome (PICO) criteria presented in Table 1. Literature searches were conducted in health technology assessment (HTA) databases, Cochrane Library, PubMed and Embase. The key evidence base consists of three studies including: one HTA report from the National Institute of Health and Care Excellence (NICE; MIB280)[7] comprising three published studies[5,11,12], and two additional diagnostic studies[13,14]. The FDA Summary of Safety and Effectiveness Data (SSED)[8] document served as supplementary evidence. In most studies,[5,8,11,12,14] the software module involved in automated detection of suspicious foci in WSI was referred to as Paige Prostate which is the same as PPD. The study design and characteristics of the key and supplementary evidence sources are presented in Tables A1 and A2 (Appendix A).

Table 1: Summary of PICO criteria

Population	Patients suspected of prostate cancer
Intervention	Paige Prostate Detect (PPD) to assist in the diagnosis of prostate cancer
Comparator	Manual assessment of prostate biopsy slides without AI assistance
Outcome	Safety, clinical- and cost-effectiveness
Abbreviation: AI, artificial intelligence.	

Safety

As PPD is used to assist pathologists in identifying suspicious foci on WSI from previously acquired prostate biopsy, no major safety issues are expected. However, there may be a risk of false positive or false negative results that may lead to unnecessary or delayed treatment, although these may not be easily quantified.[8]

Effectiveness

Accuracy

Based on one study[12] reviewed in NICE MIB280[7] and two additional diagnostic studies,[13,14] pathologists assisted by PPD demonstrated good sensitivity (90% to 96.6%) and specificity (92.8% to 98%) in the diagnosis of PCa (Table 2). In two studies,[12,14] pathologists assisted by PPD performed better than unassisted pathologists, demonstrating significant improvement in sensitivity ($p<0.001$; Table 2). However, compared to unassisted reads, Eloy et al. (2023)[13]

did not find any improvement in diagnostic performance of pathologists assisted by PPD (Table 2). The reason for the discrepancies in these findings is not clear, which may partially be influenced by different patient population, differences in pathologists' experience, their confidence over the software output, and varying range of cancer subtypes and variants across studies.[13,14] Aside from sensitivity, no overall significant difference in specificity was observed between AI-assisted and non-assisted reads.

One study further reported on the software's potential to improve sensitivity in highly experienced genitourinary (GU) and non-GU pathologists (see Figure B1 in Appendix B).[14] Sensitivity gains in pathologists assisted by PPD were also observed across all tumour grades and sizes, with one study demonstrating greater gains in the smallest tumour size (≤0.6mm; from 46% to 83%) and lowest grade group (Gleason grade 1; from 69% to 89%; see Tables B1 and B2 in Appendix B).[12,14] This corroborated findings by Eloy et al. (2023)[13] where the number of ambiguous atypical small acinar proliferation diagnosed by pathologists assisted by PPD reduced by 32% (p<0.001).

Table 2: Diagnostic accuracy of AI-assisted *vs.* unassisted read of prostate needle biopsies

Study	N*	Sensitivity			Specificity		
		AI-assisted, % (95% CI)	Unassisted, % (95% CI)	p-value	AI-assisted, % (95% CI)	Unassisted, % (95% CI)	p-value
Raciti et al. (2020)[12]	304	90.0 (NR)	73.8 (NR)	**<0.001**	95.2	96.6	0.33
Raciti et al. (2022)[14]	610	96.6 (94.6 to 98.6)	88.7 (84.5 to 92.8)	**<0.001**	98.0 (97.0 to 98.9)	97.3 (96.2 to 98.4)	**0.02**
Eloy et al. (2023)[13]	105	95.5 (NR)	96.8 (NR)	NR	92.8 (NR)	93.9 (NR)	NR

* Refers to the number of prostate needle core biopsy whole slide images.
Abbreviations: AI, artificial intelligence; CI, confidence interval; NR, not reported.

Raciti et al. (2022)[14] demonstrated that in reads (n=797) where assessment changed between initially unassisted and subsequently AI-assisted pathologists, 100% of those that became correct (n=341) were Paige driven, defined as those in which PPD classification was correct and matched the reads by pathologists assisted by PPD (see Figure B2 in Appendix B). In this same study, among a small number of reads that became incorrect (n=54), 85.2% were also Paige driven, indicating the accuracy of PPD needs further validation. This highlights the need for pathologists to remain active in the diagnostic process and understand situations where AI may underperform.[14]

Although not the intended use, compared to diagnosis by pathologists as the reference, PPD as a standalone software also demonstrated good diagnostic accuracy in classifying WSIs, with consistent performance across patient age, race and ethnicity (see Tables C1 and C2 in Appendix C).[5,11,12,14]

Impact on pathologist workflow

While no studies reported on patient health outcomes, some studies found PPD impacted pathologist workflow by improving turnaround time, with some indications of improved efficiency. Compared to unassisted reads, pathologists assisted by PPD had significantly reduced turnaround time to review WSIs in both benign and malignant cases (p<0.001; Table

3).[12,13] Further, as reported across three studies,[5,12,13] the software accounted for a 21% to 65.5% reduction in pathologists' time to review WSIs.

Table 3: Turnaround time of AI-assisted *vs.* unassisted read of prostate needle biopsies

Study	Turnaround time for review of WSI*		p-value
	AI-assisted	Unassisted	
Raciti et al. (2020)[12]			
Overall	55s ± 43s	63s ± 39s	**<0.001**
Cancer cases	48s ± 41s	61s ± 34s	**<0.001**
Cancer cases <1mm	52s ± 42s	64s ± 33s	**0.026**
Benign cases	59s ± 44s	64s ± 43s	0.086
Eloy et al. (2023)[13]			
Overall	108.5s	139s	**<0.001**
Cancer cases	206s	253.5s	**<0.001**
Benign cases	82s	100.5s	**<0.001**

* Mean time was reported by Raciti et al. (2020)[12] while median time was reported by Eloy et al. (2023)[13].
Abbreviation: WSI, whole slide image.

While there is some evidence suggesting that PPD may improve efficiency, it remains uncertain. Among 797 discordant assessments between unassisted and AI-assisted reads, Raciti et al. (2022)[14] showed that the software was responsible for similar proportion (99.7%) of efficiency gain, as measured by correct assessment in 287 of 288 of reads that were initially deferred without AI assistance, and of efficiency loss, where it led to a deferral in 98.2% (112 of 114) of cases that were initially correct without AI assistance (see Figure B2 in Appendix B).[14]

Healthcare system benefit

As reported by Eloy et al. (2023)[13], compared to unassisted reads, the software led to significantly reduced immunohistochemistry (IHC; 36% *vs.* 46%; p<0.001) and second opinion requests (7% *vs.* 12%; p=0.006) by pathologists (see Table B3 in Appendix B). This indicates the potential of Paige-driven reduced consumption of healthcare resources.

In addition, the software may also have wider implications as it drives the adoption of digital pathology, enabling cross-institutional consultations. Along with the SaaS nature of the software, the move towards digital pathology may also allow remote work, thus freeing up laboratory space and resources.[15,16]

Cost-effectiveness

No cost-effectiveness studies for PPD were identified. NICE assessed that while adopting the technology is likely to cost more than standard care, there may be a potential for it to release resources and produce overall cost savings.[7]

Ongoing trials

One ongoing trial was identified from the ScanMedicine database (NIHR Innovation Observatory; Table 4). Funded by the National Health Service (NHS) Accelerated Access Collaborative and jointly conducted by Oxford University and regional NHS partners, the prospective study investigates the real-world use of Paige Prostate.

Table 4: Ongoing trial

Study (Trial ID)	Estimated enrolment	Brief description	Estimated completion date
ARTICULATE PRO (ISRCTN91685765)	1,500	This study examines if and how pathologists' diagnoses are changed when they see and use the information made available by Paige Prostate in a real-world clinical setting. In addition, the study investigates the impact of Paige Prostate on resource utilisation and patient management.	January 2024

Summary

Overall, PPD was found to be safe and may improve the sensitivity of pathologist reads, turnaround times and healthcare resource utilisation. There were no major safety issues related to use of the software. Pathologists assisted by PPD exhibited good accuracy (sensitivity, 90% to 96.6%; specificity, 92.8% to 98%) in the diagnosis of PCa. Compared to unassisted reads, some studies demonstrated that pathologists assisted by PPD performed better than unassisted pathologists, with gain in sensitivity ($p<0.001$). However, this was not consistently demonstrated across the limited evidence base, and no significant difference in specificity was observed, indicating the need for further validation on the accuracy of PPD. In addition, there was some evidence of improved sensitivity observed in both highly experienced GU and non-GU pathologists, and across all tumour grades and sizes, with the greatest gains reported for the smallest tumours (≤0.6mm; from 46% to 83%) and lowest grade group (Gleason grade 1; from 69% to 89%).

There is also some evidence indicating an improved turnaround time for PPD-assisted pathologist reads by 21% to 65.5% in comparison with unassisted pathologists, although the impact of the software on their efficiency remains ambiguous. The software was also found to bring healthcare system benefits where it reduced consumption of healthcare resources. The cost-effectiveness of PPD remains uncertain.

However, findings from these studies must be interpreted with caution. NICE assessed the evidence to be of low to moderate methodological quality.[7] Other limitations include the lack of prospective validation when implemented into routine clinical practice and unclear impact of the software on patient outcome.

VII. Estimated Costs

As reported by NICE, the Paige Prostate pricing model will be a subscription-based service. The cost per slide starts at around £1 (S$2)[a] and can vary depending on the laboratory's volume of prostate biopsies, number of biopsies and slides per case and usage of cloud storage and archiving services.[7] This fee includes both detection and grading and quantification software modules with outputs displayed on the Paige FullFocus viewer.[7] It is unclear if additional costs may be incurred for the web-based Paige FullFocus viewer.

[a] Based on the Monetary Authority of Singapore exchange rate as of 12 May 2023: £1=S$1.6673. Figures were rounded to the nearest dollar.

This is in addition to a one-time fee to integrate Paige Prostate into laboratory information management system (LIMS), which starts from £15,000 (S$25,000)[a] and depend on the level and type of integration and LIMS provider.[7]

VIII. Implementation Considerations

The adoption of PPD involves a range of implementation considerations. At the institutional level, proper information technology (IT) infrastructure is required for the deployment of the software in pathology laboratories. Locally, digital pathology workflow for image reporting is not widely used at most pathology laboratories (Personal Communication, Senior Consultant from National University Hospital, 12 May 2023). This presents an adoption barrier to integrating the Paige Prostate software within existing pathology workflows, due to the need for additional cost in IT infrastructure and the transition towards digital pathology. IT and cost considerations may also arise from integrating the software with current LIMSs and electronic health records systems.

In addition, there is a need for regulatory oversight in the introduction of such AI medical device (AI-MD) into clinical practice. Based on the Ministry of Health (MOH) Artificial Intelligence in Healthcare Guidelines (AIHGle),[17] key considerations include a clear understanding of the intended use and purpose of the software in clinical pathways, risk assessment to anticipate software failure and mitigation measures, assessment of cybersecurity vulnerabilities and performance tracking to ensure similar performance of the AI-MD in local settings. In particular, it was reported that the software performance is insensitive to staining and tissue preparation variabilities, negating the need for on-site calibration to ensure optimal performance.[14] In the long term, monitoring of the software performance and ensuring that it remains clinically relevant will be required post-implementation.[17]

The introduction of PPD into clinical workflow may also entail pathologist training and acceptance. As mentioned by a NICE-consulted clinical expert, training may also be required for healthcare professionals who convey results to patients – to support patient understanding around decision making for their clinical diagnosis and management.[7] Furthermore, as the software can disrupt existing well-established clinical workflow (e.g. shift towards digital pathology), the inertia of healthcare professionals towards its adoption should be considered. To note, a survey conducted on pathologists who took part in a clinical study reported that they would consider digitally reviewing WSIs for primary diagnosis if such system includes Paige Prostate.[12]

IX. Concurrent Developments

There are several other AI technologies like PPD that assist pathologists in reading prostate WSI for the diagnosis of PCa (Table 5).

Table 5: Similar technologies in development

Technology (Manufacturer)	Brief description	Status
Galen Prostate (Ibex Medical Analytics)	Galen Prostate is a clinical-grade AI algorithm that assists pathologists in improving the detection and grading of prostate cancer.	CE marked

DeepDx Prostate (Deep Bio Inc.)	A medical software that classifies the histological severity of prostate cancer by analysing the WSI of prostate biopsy tissue with AI.	
Aiforia Clinical AI Model for Prostate Cancer (Aiforia)	The deep learning-based tool automatically detects and grades tumor areas in prostate tissue images from WSI to assess Gleason patterns.	
HALO Prostate AI (Indica Labs Inc)	A deep learning-based screening tool designed to assist pathologists in the identification and grading of prostate cancer in core needle biopsies. The algorithm automatically analyses all appropriate case slides and notifies pathologists of cases with suspected findings directly in their native workflow.	
QAi Prostate (Qritive)	Using machine learning algorithms, QAi Prostate can identify prostatic adenocarcinoma regions as well as classify malignant and benign tumour areas in biopsy tissue samples.	For research or LDT use
Lunit INSIGHT (Lunit Inc)	A deep learning-based software that assists radiologists or clinicians in the interpretation of medical images for lung and breast cancer, with plans from the company to expand to prostate cancer.	In development
Abbreviations: AI, artificial intelligence; CE, Conformité Européene; LDT, laboratory developed test; WSI, whole slide image.		

X. Additional Information

Local clinical experts shared that the technology may be more useful for non-experts than experts, while providing expert pathologists more time for complex cases. If used in a collaborative and inclusive manner, it may improve the standard for pathological assessment, reduce inter-reader variation, improve efficiency and training of younger pathologists although further validation of the software and implementation costs should be taken into consideration. However, such a technology would only work with the adoption of digital pathology workflows in local laboratories. Further, it was shared that pathologists would still have to examine all the cores in the slides despite AI assistance, rendering savings in terms of time and effort equivocal (Personal communications, Senior Consultant from National University Hospital, 12 May 2023; and Senior Consultant from National Cancer Centre Singapore, 30 May 2023).

References

1. Leslie SW, Soon-Sutton TL, R IA, Sajjad H, Siref LE. Prostate Cancer. StatPearls. Treasure Island (FL): StatPearls Publishing LLC.; 2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470550/

2. Eyrich NW, Morgan TM, Tosoian JJ. Biomarkers for detection of clinically significant prostate cancer: contemporary clinical data and future directions. Transl Androl Urol. 2021;10(7):3091-103. doi: 10.21037/tau-20-1151.

3. SingHealth. Prostate Cancer [Internet]. Singapore: SingHealth; c2021 [cited 10 May 2023]. Available from: https://www.singhealth.com.sg/patient-care/conditions-treatments/prostate-cancer

4. American Cancer Society. Survival Rates for Prostate Cancer [Internet]. Atlanta: ACS; c2023 [cited 10 May 2023]. Available from: https://www.cancer.org/cancer/types/prostate-cancer/detection-diagnosis-staging/survival-rates.html

5. da Silva LM, Pereira EM, Salles PG, et al. Independent real-world application of a clinical-grade automated prostate cancer detection system. J Pathol. 2021;254(2):147-58. doi: 10.1002/path.5662.

6. Paige.AI Inc. Built by Pathologists for Pathologists [Internet]. New York: Paige.AI; c2023 [cited 2023 May 10]. Available from: https://paige.ai/clinical/

7. National Institute for Health and Care Excellence. Paige Prostate for prostate cancer (MIB280). London: NICE; 2021 [cited 5 May 2023]. Available from: https://www.nice.org.uk/advice/mib280

8. US Food and Drug Administration. Evaluation of automatic Class III designation for Paige Prostate. United States: FDA; 2021 [cited 11 May 2023]. Available from: https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN200080.pdf

9. Kanesvaran R, Castro E, Wong A, et al. Pan-Asian adapted ESMO Clinical Practice Guidelines for the diagnosis, treatment and follow-up of patients with prostate cancer. ESMO Open. 2022;7(4):100518. doi: 10.1016/j.esmoop.2022.100518.

10. George RS, Htoo A, Cheng M, et al. Artificial intelligence in prostate cancer: Definitions, current research, and future directions. Urol Oncol. 2022;40(6):262-70. doi: 10.1016/j.urolonc.2022.03.003.

11. Perincheri S, Levi AW, Celli R, et al. An independent assessment of an artificial intelligence system for prostate cancer detection shows strong diagnostic accuracy. Modern Pathology. 2021;34(8):1588-95. doi: 10.1038/s41379-021-00794-x.

12. Raciti P, Sue J, Ceballos R, et al. Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. Modern Pathology. 2020;33(10):2058-66. doi: 10.1038/s41379-020-0551-y.

13. Eloy C, Marques A, Pinto J, et al. Artificial intelligence-assisted cancer diagnosis improves the efficiency of pathologists in prostatic biopsies. Virchows Arch. 2023;482(3):595-604. doi: 10.1007/s00428-023-03518-5.

14. Raciti P, Sue J, Retamero JA, et al. Clinical Validation of Artificial Intelligence-Augmented Pathology Diagnosis Demonstrates Significant Gains in Diagnostic Accuracy in Prostate Cancer Detection. Arch Pathol Lab Med. 2022. doi: 10.5858/arpa.2022-0066-OA.

15. Liscia DS, Bellis D, Biletta E, D'Andrea M, Croci GA, Dianzani U. Whole-Slide Imaging Allows Pathologists to Work Remotely in Regions with Severe Logistical Constraints Due to Covid-19 Pandemic. J Pathol Inform. 2020;11:20. doi: 10.4103/jpi.jpi_32_20.

16. Williams BJ, Brettle D, Aslam M, et al. Guidance for Remote Reporting of Digital Pathology Slides During Periods of Exceptional Service Pressure: An Emergency Response from the UK Royal College of Pathologists. J Pathol Inform. 2020;11:12. doi: 10.4103/jpi.jpi_23_20.

17. Ministry of Health Singapore. Artificial Intelligence in Healthcare Guidelines (AIHGle). Singapore: MOH; 2021 [cited 15 May 2023]. Available from: https://www.moh.gov.sg/licensing-and-regulation/artificial-intelligence-in-healthcare

Appendix

Appendix A: Studies identified and study design

Table A1: List of included studies

Type of study	Key evidence base	Supplementary evidence base
HTA report	1	—
Diagnostic studies	2	—
FDA SSED	—	1

Note:

1. Inclusion criteria
   a. Studies that fulfil the PICO criteria listed in Table 1.
2. Exclusion criteria
   b. Studies only available in the abstract form.

Abbreviations: FDA, United States Food and Drug Administration; HTA, health technology assessment; SSED, Summary of Safety and Effectiveness Data.

Table A2: Design and characteristics of included studies

Study	Number of WSIs	Number of patients	Study design	Assessment of diagnostic performance		Washout period between AI-assisted and unassisted read
				Standalone AI *vs.* unassisted pathologist	AI-assisted *vs.* unassisted pathologist	
HTA report (NICE MIB280)						
Da Silva et al. (2021)	600	100	Retrospective	✓		NA
Perincheri et al. (2021)	1,876	118	Retrospective	✓		NA
Raciti et al. (2020)	304	NR	Retrospective	✓	✓	4 weeks
Additional key evidence						
Raciti et al. (2022)	610	NR	Retrospective	✓	✓	Immediate
Eloy et al. (2023)	105	NR	Retrospective		✓	At least 2 weeks
Supplementary evidence						
FDA SSED	—	—	—	—	—	—

Abbreviations: AI, artificial intelligence; FDA, United States Food and Drug Administration; NA, not applicable; NR, not reported; SSED, Summary of Safety and Effectiveness Data; WSI, whole slide image.

Appendix B: List of supplementary tables and figures

Table B1: Summary of sensitivity gains across histological grade groups and tumour sizes between AI-assisted and unassisted reads

	Level	N	Observed Reader Average Sensitivity			Model Class III ANOVA p-value*		
			Assisted	Unassisted	Difference	Mode	Factor	Model/Factor interaction
Gleason Grade Group (N=190)	ASAP, Treated, or Unknown	15	74.6%	54.2%	20.4%	<0.001	<0.001	0.13
	1	110	98.1%	89.8%	8.4%			
	2	39	99.2%	95.4%	3.8%			
	3	10	99.4%	96.9%	2.5%			
	4	12	99.0%	90.6%	8.3%			
	5	4	98.4%	95.3%	3.1%			
Primary Gleason Grade (N=190)	ASAP, Treated, or Unknown	15	74.6%	54.2%	20.4%	<0.001	<0.001	0.04
	3	149	98.4%	91.2%	7.2%			
	4	26	99.0%	93.8%	5.3%			
Secondary Gleason Grade (N=190)	ASAP, Treated, or Unknown	15	74.6%	54.2%	20.4%	<0.001	<0.001	0.07
	3	120	98.2%	90.4%	7.9%			
	4	51	99.1%	94.2%	4.9%			
	5	4	98.4%	95.3%	3.1%			
Tumour Length (mm) Quartile (N=180)	0.1 to 0.35	45	96.8%	81.7%	15.1%	<0.001	<0.001	0.06
	0.35 to 0.55	45	98.3%	90.1%	8.2%			
	0.55 to 3	51	99.6%	95.3%	4.3%			
	3 to 42	39	99.4%	98.2%	1.1%			
Tumour Percentage Gleason 4/5 Quartile (N=51)	2 to 10	25	99.0%	97.5%	1.5%	0.99	0.93	0.52
	10 to 15	3	100.0%	87.5%	12.5%			
	15 to 45	11	99.4%	92.6%	6.8%			
	45 to 100	12	99.5%	97.4%	2.1%			
Tumour Percentage Carcinoma Quartile (N=180)	1 to 3	65	98.8%	84.5%	14.2%	<0.001	<0.001	<0.001
	3 to 5	36	96.0%	90.6%	5.4%			
	5 to 15	37	99.5%	94.8%	4.7%			
	15 to 95	42	99.6%	99.1%	0.4%			
ASAP (N=190)	No	180	98.5%	91.3%	7.3%	<0.001	<0.001	<0.001
	Yes	10	61.9%	41.9%	20.0%			
Origin of slide (N=190)	MSKCC	95	96.3%	88.4%	7.9%	<0.001	0.74	0.69
	Submitted	95	97.0%	88.9%	8.0%			

* Logistic model with random slide intercept, reader random effect, and factor and read type main effects.

Abbreviations: ASAP, atypical small acinar proliferation; mm, millimetres, MSKCC, Memorial Sloan Kettering Cancer Center; WSI, whole slide images.

Table adapted from Raciti et al. (2022)[14].

Table B2: Average sensitivity by Gleason Grade group with and without Paige Prostate Alpha.

	Grade group 1	Grade group 2	Grade group 3	Grade group 4	Grade group 5
Without Paige Prostate Alpha	69%	85%	89%	90%	100%
With Paige Prostate Alpha	89%	97%	100%	90%	100%
Change	+20%	+13%	+11%	NC	NC

Note: Paige Prostate Alpha is an early version of Paige Prostate Detect.

Abbreviation: NC, no change.

Table adapted from Raciti et al. (2020)[12].

Table B3: Impact of AI-assisted read on immunohistochemistry and second opinion requests compared to unassisted read

	AI-assisted read, n (%)	Unassisted read, n (%)	Difference (%)	p-value
IHC requests				
All cases	153 (36.43)	193 (45.95)	-20.72	<0.001*
Cancer cases	67 (42.95)	89 (57.05)	-24.72	<0.001*
Negative cases	86 (32.58)	104 (39.39)	-17.31	<0.001*
Second opinion requests				
All cases	31 (7.38)	51 (12.14)	-39.21	<0.001†
Cancer cases	18 (11.54)	29 (18.59)	-37.93	0.001†
Negative cases	13 (4.92)	22 (8.33)	-40.91	<0.001†

* Chi-squared test

† Fisher exact test

Abbreviations: AI, artificial intelligence; IHC, immunohistochemistry.

Table adapted from Eloy et al. (2023)[13].

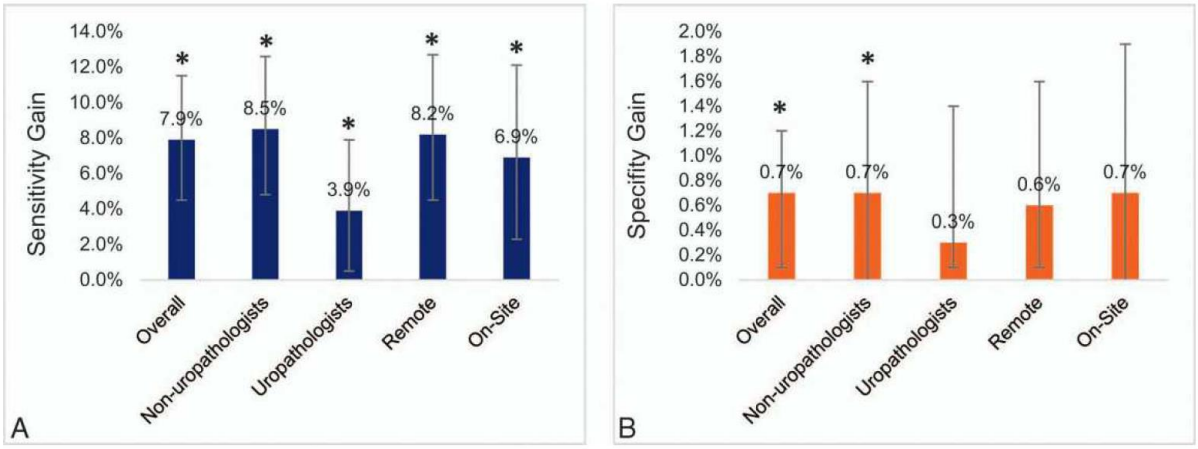

Figure B1: Gains in diagnostic performance by pathologists, stratified by experience and location. Note that asterisks (*) indicate statistically significant changes (p<0.05). (A) Statistically significant sensitivity gains were seen regardless of

pathologist type and location of slide review. **(B)** Statistically significant specificity gains were seen among all pathologists overall and among non-uropathologists. Figure adapted from Raciti et al. (2022)[14].

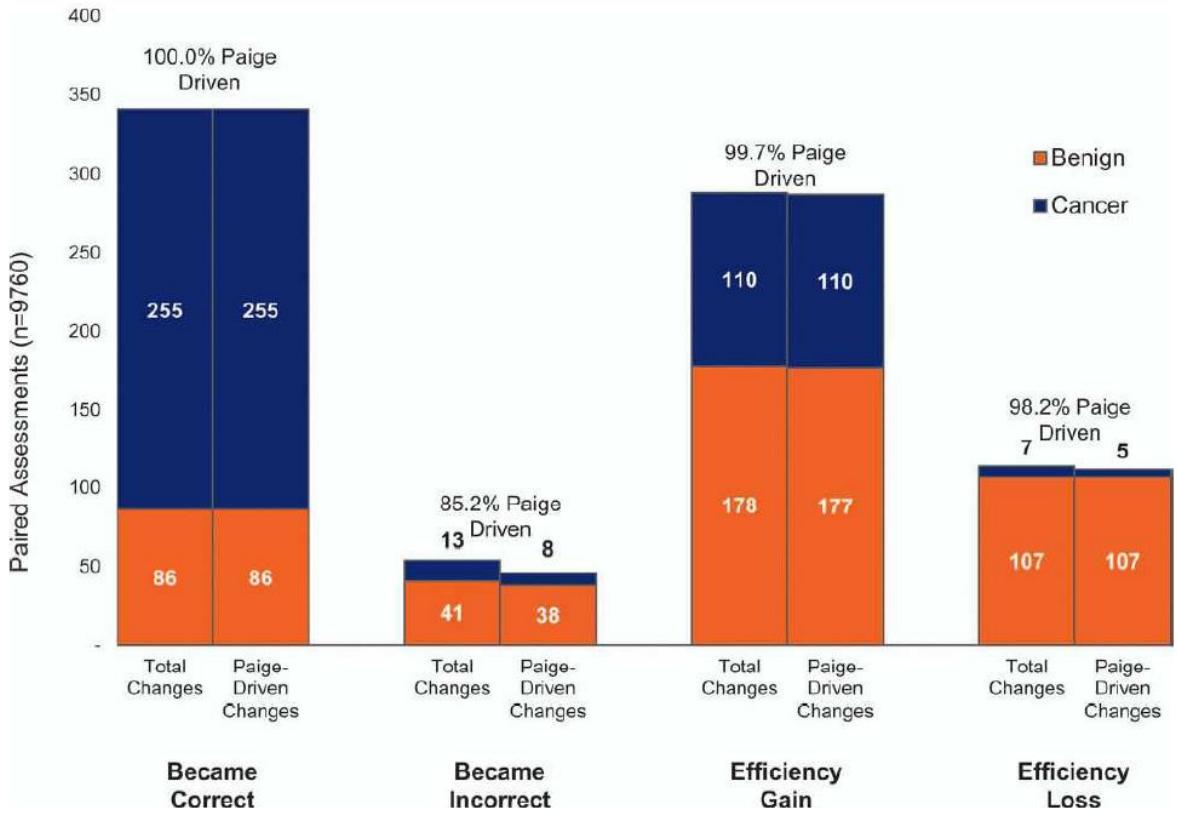

Figure B2: Change in pathologist's assessment driven by Paige Prostate. In 797 reads, Paige Prostate-assisted pathologist reads differed from unassisted reads. Paige-driven changes are defined as those in which Paige Prostate classification was correct and matched the Paige Prostate-assisted pathologist reads. All reads (100%) that became correct were Paige driven. Many reads (85.2%) that became incorrect were also Paige driven. Paige-driven reads that resulted in efficiency gains slightly outnumbered those that resulted in efficiency losses, but overall rates were similar. Table adapted from Raciti et al. (2022)[14].

Appendix C: List of supplementary tables supporting the standalone diagnostic performance of Paige Prostate Detect

Table C1: Standalone diagnostic accuracy of Paige Prostate Detect *vs.* pathologist rendered diagnosis of PCa

Study	N*	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
da Silva et al. (2021)[5†]	579	98.9% (95.9% to 99.9%)	93.3% (90.4% to 95.5%)	86.5% (81.0% to 90.9%)	99.5% (98.1% to 99.9%)
Perincheri et al. (2021)[11†]	1,857	97.7% (NR)	99.3% (NR)	97.9% (NR)	99.2% (NR)
Raciti et al. (2020)[12†]	304	96% (NR)	98% (NR)	NR	NR
Raciti et al. (2022)[14]	610	97.4% (94.0% to 99.1%)	94.8% (92.2% to 96.7%)	NR	NR

* Refers to the number of prostate needle core biopsy whole slide images.

† Reviewed in NICE MIB240.

Abbreviations: CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

Table C2: Diagnostic performance of Paige Prostate Detect as a standalone software across patient demographics

Outcome	Factor	Factor levels	Count information	Estimate (95% CI)	p-value*
Sensitivity	Age category	<60	36/38	94.7% (82.3% to 99.4%)	0.80
		60-69	71/73	97.3% (90.5% to 99.7%)	
		70-79	60/61	98.4% (91.2% to 100.0%)	
		>79	18/18	100.0% (81.5% to 100.0%)	
	Race	Asian-far East/Indian Subcont	8/8	100.0% (63.1% to 100.0%)	>0.99
		Black or African American	16/16	100.0% (79.4% to 100.0%)	
		Other	11/11	100.0% (71.5% to 100.0%)	
		Unknown	3/3	100.0% (29.2% to 100.0%)	
		White	147/152	96.7% (92.5% to 100.0%)	
	Ethnicity	Hispanic or Latino	17/17	100.0% (80.5% to 100.0%)	>0.99
		Not Hispanic	157/162	96.9% (92.9% to 99.0%)	
		Unknown	11/11	100.0% (71.5% to 100.0%)	
Specificity	Age category	<60	63/65	96.9% (89.3% to 99.6%)	0.27
		60-69	170/178	95.5% (91.3% to 98.0%)	
		70-79	147/156	94.2% (89.3% to 97.3%)	
		>79	18/21	85.7% (63.7% to 97.0%)	
	Race	Asian-far East/Indian Subcont	10/12	83.3% (51.6% to 97.9%)	0.55
		Black or African American	27/29	93.1% (77.2% to 99.2%)	
		Other	9/9	100.0% (66.4% to 100.0%)	
		Unknown	14/15	93.3% (68.1% to 99.8%)	
		White	338/355	95.2% (92.4% to 97.2%)	
	Ethnicity	Hispanic or Latino	20/22	90.9% (70.8% to 98.9%)	0.50
		Not Hispanic	358/376	95.2% (92.5% to 97.1%)	
		Unknown	20/22	90.9% (70.8% to 98.9%)	

* ANOVA p-value from logistic model.
Table adapted from Raciti et al. (2022)[14].