Multi-stage AI analysis system to support prostate cancer diagnostic imaging

J Suchanek [1], AW Rix [2]*, AH Mehan [3], CJL Doran [4], AR Padhani [5,6], C Kastner [7,8], T Barrett [8,9], E Sala [2,8,9]

(1) Department of Engineering, University of Cambridge, Cambridge, UK
(2) Lucida Medical Ltd, University of Cambridge Biomedical Innovation Hub, Cambridge, UK
(3) School of Clinical Medicine, University of Cambridge, Cambridge, UK
(4) Sidney Sussex College, University of Cambridge, Cambridge, UK
(5) Institute of Cancer Research, London, UK
(6) Paul Strickland Scanner Centre, Mount Vernon Cancer Centre, Northwood, Middlesex, UK
(7) Department of Urology, Addenbrooke's Hospital, Cambridge, UK
(8) CamPARI Prostate Cancer Group, Addenbrooke's Hospital and University of Cambridge, Cambridge, UK
(9) Department of Radiology, Addenbrooke's Hospital and University of Cambridge, Cambridge, UK

* Corresponding author: https://lucidamedical.com/contact/

Abstract

An Artificial intelligence (AI) system was developed to support interpretation of pre-biopsy prostate multiparametric MRI (mpMRI), aiming to improve patient selection for biopsy, biopsy target identification, and productivity of segmentation and reporting, in the prostate cancer diagnostic pathway.

For segmentation, the system achieved 92% average Dice score for prostate gland segmentation on held-out test cases from the PROMISE12 dataset (10 patients).

For biopsy assessment, the system identified patients with Gleason ≥3+4 clinically significant prostate cancer (csPCA) with sensitivity 93% (95% CI 82-100%), specificity 76% (64-87%), NPV 95% (88-100%), and AUC 0.92 (0.84-0.98), using biparametric MRI (bpMRI) data from the combined PROSTATEx development validation and test sets (80 patients). Performance on the held-out PROSTATEx test set (40 patients) was higher. Radiologists in major studies achieved 93% per-patient sensitivity at specificity from 18-73%. Equivalent sensitivity is reported for comparable AI/CAD systems at specificity from 6%-42%.

For biopsy targeting, the system identified lesions containing csPCa in the PROSTATEx blinded test set (208 lesions, 140 patients) with AUC 0.84/0.85 with bpMRI/mpMRI data respectively.

The AI system shows promising performance compared to radiologists and the literature. Further development, regulatory approvals, and evaluation with larger, multi-centre datasets are now planned.

Background

While pre-biopsy multiparametric magnetic resonance imaging (mpMRI) substantially improves detection of clinically significant prostate cancer (csPCa) and reduces unnecessary biopsies and diagnoses of insignificant cancer, there remain ongoing challenges with underdiagnosis, biopsy rates, and overdiagnosis. Major studies indicate that 21-49% of patients may still undergo a potentially avoidable biopsy [Ahmed 2020, Kasivisvanathan 2018, Rouvière 2019, van der Leest 2019], and up to 12% of csPCa may be missed [Drost 2019].

AI has the potential to support clinical interpretation and improve accuracy of pre-biopsy MRI for prostate cancer to address these concerns, and could also help improve productivity. We compare a new artificial intelligence (AI) based system for detecting Gleason ≥3+4 csPCa using MRI, with human readers and existing computer aided diagnosis (CAD) literature.

Methods

An AI diagnostic aid for prostate cancer detection was developed using a multi-stage architecture designed to produce three outputs: prostate segmentation, for PSA density estimation and fusion

biopsy; cancer risk calculation, to help reduce unnecessary biopsies; and lesion identification, to support biopsy targeting.

Data was obtained from open, anonymised prostate MRI datasets, and divided into training, development validation, and held-out test sets. Held-out test data was not used in model training or optimisation. Segmentation models were trained on axial T2 weighted imaging (T2WI) MRI data and accompanying prostate annotations from the PROMISE12 and NCI-ISBI 2013 Challenge datasets [Litjens 2014a, Bloch 2015], acquired using a variety of 1.5T and 3T scanners. Models for cancer assessment (risk calculation and lesion identification) were trained on PROSTATEx [Litjens 2014b], a mpMRI dataset acquired at one centre on two 3T scanners (Siemens MAGNETOM Trio and Skyra), with histopathology findings from MR-guided biopsy as ground truth.

Because some centres acquire bpMRI sequences only, two cancer assessment models were trained and evaluated on the PROSTATEx dataset, one using axial bpMRI data only, and a second using axial mpMRI data. bpMRI comprised b1400 diffusion-weighted imaging (DWI) and apparent diffusion coefficient (ADC) sequences computed by the scanner software, and T2WI. mpMRI additionally includes dynamic contrast-enhanced (DCE) axial T1 weighted sequences captured following injection of a gadolinium-based contrast agent.

Performance was evaluated after model development was completed. For prostate gland segmentation using axial T2WI, the Dice coefficient was calculated using held-out test cases from the PROMISE12 dataset.

The calculated risk scores were assessed for patient selection for biopsy by estimating per-patient sensitivity, specificity, and negative predictive value (NPV), all at optimum NPV, and receiver operating characteristic area under curve (AUC), with bootstrapped 95% confidence intervals, using the development validation and held-out test sets from the PROSTATEx dataset.

For lesion identification, submission was made to the PROSTATEx Grand Challenge, which evaluates AUC only for blinded test cases in the PROSTATEx dataset [Litjens 2014b]. In addition, per-lesion sensitivity, specificity, NPV and AUC were evaluated using the development validation and held-out test sets from the PROSTATEx dataset.

Results

The system achieved 92% average Dice score for prostate gland segmentation when compared with the benchmark examples from the PROMISE12 [Litjens 2014a] held-out test set (10 patients).

For the task of selecting patients with csPCA for biopsy, radiologists achieved per-patient sensitivity of 88%/93%/94% and specificity 45%/18%/73% respectively in three major studies [Ahmed 2018, Rouvière 2019, van der Leest 2019]. Comparable AI/CAD publications report 93% sensitivity using held-out test data at specificity ranging from 6% [Thon 2017] to 42% [Schelb 2020].

The AI system had per-patient sensitivity 93% (95% CI 82-100%), specificity 76% (64-87%), NPV 95% (88-100%), and AUC 0.92 (0.84-0.98), evaluated using bpMRI data from the combined PROSTATEx development validation and test sets (128 lesions, 80 patients). The mpMRI model performed similarly on the same data, with sensitivity 93% (83%-100%), specificity 78% (65-88%), NPV 95% (89-100%), and AUC 0.91 (0.84-0.97). On the PROSTATEx held-out test set (64 lesions, 40 patients) both models had higher performance, per-patient and per-lesion.

For identifying biopsy targets, the AI system detected lesions containing csPCa in the blinded PROSTATEx Grand Challenge test set (208 lesions, 140 patients) with AUC 0.84/0.85 with bpMRI/mpMRI data. In the combined PROSTATEx development validation and test sets (128 lesions, 80 patients), the bpMRI/mpMRI models had per-lesion sensitivity 94% (85-100%)/94% (85-100%), specificity 71% (61-89%)/69% (59-78%), NPV 97% (93-100%)/97% (93-100%), and AUC 0.89 (0.83-0.95)/0.90 (0.84-0.95). The blinded Grand Challenge test set appears to contain several extra-prostatic lesions that the model does not identify, and this may account for differences in performance between these test sets.

Conclusion

The AI system's performance is in line with central reporting by expert radiologists, providing a preliminary indication that it could be used to support accurate assessment and reporting of pre-biopsy MRI for prostate cancer. The system processes axial data only, enabling its use with abbreviated protocols. Similar performance with both bpMRI and mpMRI data suggests a potential

application to support reduction or avoidance of contrast agent sequences. Its accuracy appears to exceed published results for similar prostate CAD/AI systems. Methodological and dataset differences and small test set sizes limit these comparisons. Improved detection of extra-prostatic lesions, workflow integration, training and evaluation with larger, more diverse datasets, and prospective studies are recommended to evaluate this and related detection tasks further.

Impact

Artificial intelligence-based software could support exclusion of clinically significant prostate cancer with high NPV and AUC, assist with the identification of lesions to target for biopsy, and may facilitate avoidance of gadolinium-based contrast agents or use of faster MRI protocols.

Disclosures

Dr Antony W. Rix and Prof Evis Sala are co-founders and shareholders of Lucida Medical Ltd (UK), the company that develops the AI software described herein.

The AI software is not currently available and is described here in connection with research use. Regulatory approvals are in progress. Please contact the corresponding author for further details.

References

[Ahmed 2020] Ahmed HU, et al. Lancet 2017;389:815-822 https://doi.org/10.1016/S0140-6736(16)32401-1
[Bloch 2015] Bloch N, et al. The Cancer Imaging Archive http://doi.org/10.7937/K9/TCIA.2015.zF0vlOPv
[Drost 2019] Drost FJH, et al. Cochrane Sys Rev Diag. 2019 https://doi.org/10.1002/14651858.CD012663.pub2
[Kasivisvanathan 2018] Kasivisvanathan V, et al. New Engl. J. Medicine 2018; 378(19):1767-1777 https://doi.org/10.1056/NEJMoa1801993
[Litjens 2014a] Litjens G, et al. Medical Image Analysis 2014;18(2):359-37 https://doi.org/10.1016/j.media.2013.12.002
[Litjens 2014b] Litjens G, et al. IEEE Trans. Medical Imaging 2014;33:1083-1092 https://doi.org/10.1109/TMI.2014.2303821
[Rouvière 2019] Rouvière O, et al. Lancet Oncol. 2019;20(1):100-109 https://doi.org/10.1016/S1470-2045(18)30569-2
[Schelb 2020] Schelb P, et al. Eur. Radiol. 2020; in press https://doi.org/10.1007/s00330-020-07086-z
[Thon 2017] Thon A, et al. PLoS ONE 2017 12(10):e0185995 https://doi.org/10.1371/journal.pone.0185995
[van der Leest 2019] van der Leest M, et al. Eur. Urol. 2019;75(4):570–578 https://doi.org/10.1016/j.eururo.2018.11.023

Copyright notice