An image classification algorithm for prostate cancer diagnosis on multiparametric transrectal ultrasound – Preliminary results of a prospective multicenter diagnostic accuracy trial

Jager A.¹, Posterma A.W.¹, Hagens M.J.², Van Der Poel H.G.², Beerlage H.P.¹, Mischi M.³, Oddens J.R.¹

¹Amsterdam University Medical Centers, Dept. of Urology, Amsterdam, The Netherlands, ²Netherlands Cancer Institute, Dept. of Urology, Amsterdam, The Netherlands, ³Eindhoven University of Technology, Dept. of Electrical Engineering, Eindhoven, The Netherlands

Introduction & Objectives: The tendency towards early detection programs for prostate cancer (PCa) is expected to increase demand for pre-biopsy imaging. Meeting this demand with MRI while maintaining high quality imaging will be challenging. Ultrasound could be a cost-effective and more widely available alternative to MRI. This study reports the preliminary performance of a machine learning (ML) based image classification algorithm for PCa diagnosis on 3D multiparametric transrectal ultrasound (3D mpUS).

Materials & Methods: This is a phase II prospective multicenter diagnostic accuracy study. A total of 715 patients will be included in a period of 2 years. Patients are eligible in case of suspected PCa for which prostate biopsy is indicated or in case of biopsy proven PCa for which radical prostatectomy (RP) will be performed. All study participants undergo 3D mpUS, consisting of 3D greyscale, 3D contrast-enhanced ultrasound and 3D shear wave elastography. The study-arm included before RP provides the ground-truth to train the algorithm: whole mount RP histopathology. Preliminary diagnostic performance of the algorithm for detection of clinically significant PCa (csPCa) is reported as Area Under the Receiver Operating Characteristic Curve (AUROC) on a per voxel (0.75x0.75x0.75 mm) and a per region level (2 cm³), using full mount histopathology as the reference standard. csPCa is defined as the International Society of Urological Pathology (ISUP) Grade Group (GG) ≥ 2. Validation is based on a 5-fold nested cross-validation strategy.

Results: At the time of writing a total of 59 complete datasets (e.g. 3D mpUS and whole mount histopathology) were available for algorithm training and evaluation of diagnostic performance. Best classification performances were achieved using support vector machine and XGBoost classifiers. AUROC for csPCa detection on a per voxel and per region level were 0.84 and 0.90, respectively.
Conclusions: Preliminary evaluation of the ML-based image classification algorithm shows promising diagnostic accuracy for csPCa detection on 3D mpUS. Expanding the dataset and external validation will have to prove the robustness of the algorithm. If future clinical validation trials prove non-inferiority compared to MRI, 3D mpUS will represent a valuable diagnostic tool in early detection programs for PCa.