P068 Development of a machine learning algorithm to accurately detect clinically significant prostate cancer on MRI

EUR Urol Suppl 2019;18(11):e3493

Dillon J.W.M.[1], Whish-Wilson T.I.[2], Riddell S.J.[2], Wong L-M.[2], Brotchie P.[3], Sutherland T.R.[1]

[1]St Vincent's Hospital Melbourne, Dept. of Medical Imaging, Melbourne, Australia, [2]St Vincent's Hospital Melbourne, Dept. of Surgery, Melbourne, Australia, [3]St Vincent's Hospital Melbourne, Dept. of Medical Imaging , Melbourne, Australia

Introduction & Objectives: Deep machine learning has proven capable of picking up features barely perceptible to radiologists, representing an exciting diagnostic aid. Prostate Imaging Reporting And Data System v.2 (PI-RADSv2) has been validated with a high accuracy and negative predictive value when detecting clinically significant prostate cancer (csPCa). The purpose of this study is to investigate whether it is possible to develop an artificial intelligence deep machine learning algorithm to detect csPCa on multiparametric prostate MRI studies (mp-MRIp), using PI-RADS 4 and 5 images as a small training set available only at a tertiary hospital.

Materials & Methods: Retrospective series of patients between 1 January 2017 to 1 July 2018, who had a mp-MRIp at a tertiary hospital that demonstrated a PI-RADS 4 or 5, and a transrectal/transperineal biopsy, or radical prostatectomy (RP) that demonstrated Gleason ≥3+4 disease. Lesions were manually traced and binary masks images were generated. Supervised deep machine learning was performed using a u-net model with images split into training and test sets. Inference scripts were determined using a weighting of 1.00. Training was conducted using 100-250 epochs and a data augmentation factor of 0 or 4.

Results: 62 patients were selected, giving a training set of 160 T2, 151 high b value DWI, and 146 ADC images. 33 images of tumour and 5 normal images were used as a test set. Optimal conditions were a data augmentation factor of 4, 100 test epochs and excluding all predictions under 5 x 5 mm. This resulted in a 100% accuracy for the high b value DWI model, 82.1% accuracy for the T2 model and 84.2% accuracy for the ADC model. All models correctly predicted no tumour on all normal images. This gave non-localisation within the prostate sensitivities of 1.00 for the high b value DWI and 0.82 for both T2 and ADC, with specificities of 1.00 for all. The DWI model was the best performed – with 100% accuracy for the 38 test images and 93.9% (31/33) of the predicted tumour areas closely correlating to the MRI abnormality. ADC and T2 models correctly had no predictions for the normal images whilst predicting 81.8% of the tumours (27/33). The ADC model had a closer correlation to the PI-RADS lesions - 96.2% (26/27), compared to 66.7% (18/27) for the T2 model. Additionally, all three sequences were able to identify bony mestastases present despite not being trained to do so.

Conclusions: Learning algorithms were successful for detecting PI-RADS 4 and 5 lesions which corresponded to csPCa with a small dataset of training images of less than 200 images for each sequence. The DWI model was the best performed with 100% accuracy for the 38 test images and 93.9% (31/33) of the predicted lesions sizes correlating closely with the size of the MRI abnormality. Deep machine learning is capable of detecting lesions from mpMRIp using limited datasets available locally to tertiary hospitals.