

## **Automated AI Tumor Burden Analysis in PSMA & FDG PET and Lu-177 PSMA Quantitative SPECT with Global Threshold Regional Consensus Network (GTRC-Net)**

Price Jackson<sup>1,2</sup>, Lachlan McIntosh<sup>1</sup>, James Buteau<sup>1</sup>, Raghava Kashyap<sup>1</sup>, Kai Qin<sup>3</sup>, Ramin Alipour<sup>1</sup>, Sebastian Casanueva<sup>1</sup>, Yu Sun<sup>1</sup>, Shahneen Sandhu<sup>1,2</sup>, Michael Hofman<sup>1,2</sup>

Affiliations: <sup>1</sup>Peter MacCallum Cancer Centre, Melbourne Australia, <sup>2</sup>Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne Australia, <sup>3</sup>Swinburne University, Melbourne Australia

**Background:** Quantitative assessment of disease extent on PSMA & FDG PET imaging prior to therapy with <sup>177</sup>Lu-PSMA has yielded prognostic biomarkers to predict likelihood of treatment response with intensity of PSMA expression (SUV<sub>mean</sub>) and volume of FDG avid disease (total metabolic tumor volume, TMV) being most strongly correlated. Despite PSMA PET representing a strong predictive and prognostic biomarker, the manual overhead whole-body delineation and lack of robust standardisation currently hinders adoption in routine clinical use or as mandatory scoring criteria for early-stage clinical trials. Previous works in AI disease delineation on PSMA PET imaging have achieved volumetric accuracies of 65-70% (Jafari et al., 2023), 54-65% (Zhao et al., 2020), and 82% (Moazemi et al., 2021), but further refinement is warranted to achieve a clinically viable, fully automated workflow.

**Methods:** Expert annotations of total tumor volume were collected for a population of metastatic prostate cancer patients receiving <sup>177</sup>Lu-PSMA therapy. In total, 567 PSMA PET/CT, 109 PSR PET/CT, 390 FDG PET/CT, and 477 LuPSMA quantitative SPECT/CT series were utilised for AI training. Semantic segmentation models based on nnU-Net were developed and extended with a novel threshold-based consensus post-processing technique (GTRC-Net) for each of the three applications: PSMA PET/CT, FDG PET/CT, & <sup>177</sup>Lu-PSMA post-therapy SPECT/CT.

**Results:** Models trained using GTRC-Net framework achieved validation mean and median volumetric dice accuracy for PSMA/PSR PET of 93.7 & 97.5%, FDG PET of 84.5 & 96.0%, and LuPSMA SPECT of 97.0 & 99.6%. Volumetric agreement with expert annotations was consistently improved through the consensus post-processing method relative to the initial nnU-Net segmentation output. Correlation scores for AI-derived image biomarkers (Pearson  $r$ ) were 0.977 for PSMA SUV<sub>mean</sub> and 0.942 for FDG TMV. Figure 1 illustrates the accuracy of the AI model in an example case and reports the agreement of quantitative image biomarkers for all PSMA & PSR PET/CTs used in the training dataset. Higher accuracy was achievable with PSMA tracers compared to FDG; potentially related to the higher relative tumor-to-background uptake and larger typical total disease volume. Among PSMA modalities, the improved agreement in SPECT imaging was most likely attributed to the delayed (24-hour) scan time yielding a relative reduction in physiological uptake for discrimination though partial-volume effects and reduced resolution may also reduce the complexity of the segmentation task.

**Conclusions:** This work extends on previous investigations of AI-assisted disease analysis by optimising computational methods to suit threshold-based segmentation workflows. Models for the relevant modalities in mCRPC staging and follow-up were developed on a large institutional dataset using expert annotations. Machine learning models and training workflow code are made openly available for research purposes with the aim to accelerate the understanding of predictive imaging biomarkers in prostate cancer.

**Funding Acknowledgements:** Price Jackson is a 2021 PCF Young Investigator.

**Conflicts of Interest Disclosure Statement:** None

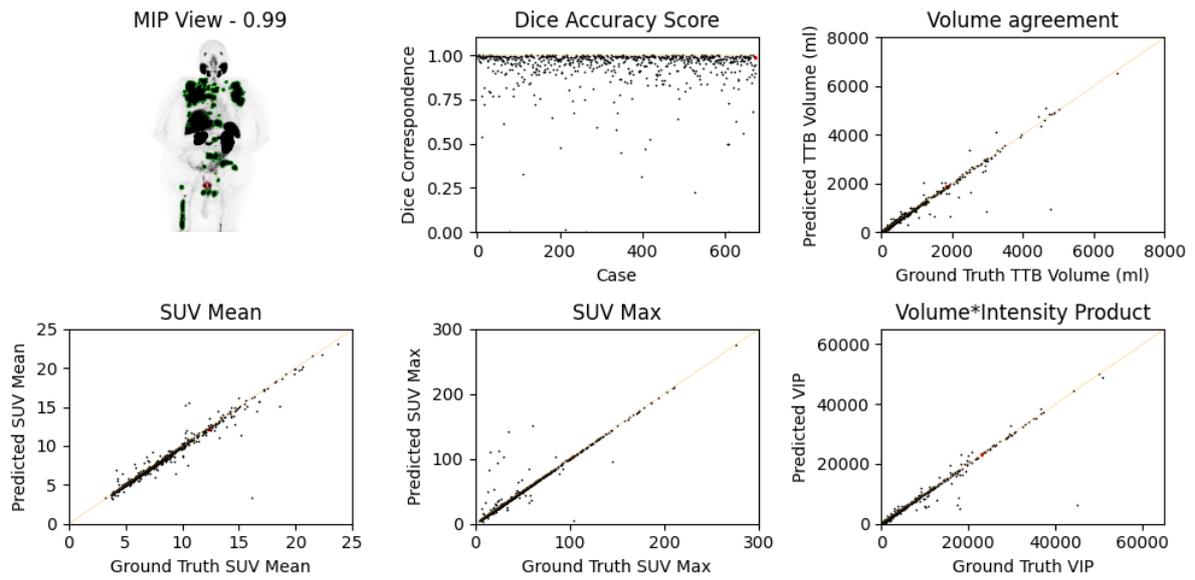


Figure 1 – Example PSMA PET accuracy illustration (top, left) showing areas of expert agreement in green and discordance in red; in this case representing 99% volumetric accuracy by Dice coefficient. Dice scores for all cases are shown (top, middle) and agreement between expert quantitative image biomarkers are illustrated in subsequent plots.