

Supplementary Discussion

Feature importance

In inspecting the feature importances across models, it is unsurprising and still notable that PSA should be the most heavily weighted feature in models M0, M1 and M3. Radiomic features across M2 and M3 demonstrate some overlapping prioritizations. M2 heavily weights *GLSZM – Zone Variance*, capturing variability in size of homogenous zones within the lesion. Additional texture features include *GLSZM – Small Area Emphasis* and *GLRLM – Run Length Non-Uniformity* which relate to textural irregularity and non-uniformity of the lesion potentially contributing to meaningful predictions. M3 included *GLSZM - Gray Level Variance*, quantifying variability in gray level intensities across homogenous zones. Also potentially indicating that the model is capturing differences in tumor heterogeneity. Meanwhile, other features in M2 (such as *Shape Sphericity* and *First Order Range*) and in M3 (such as *Least Axis Length*, *Maximum 2D Diameter*, and *Mesh Volume*) also plausibly relate to things like tumor size, bulk, and heterogeneity which intuitively relate to disease burden and invasiveness. Although the exact features differ between the models, we observe that a core set of radiomics domains (texture, shape and intensity) demonstrate a shared reliance on biologically meaningful signals a radiologist would rely on. The convergence of these features provides evidence of clinical relevance of these radiomic features for assessment of individualize risk in PCa.