



**Supplementary Figure 2:** Performance of single timepoint FLAIR lesion prediction models using an alternate deeper U-Net architecture. (A) Schematic of alternate deeper three-dimensional U-Net architecture. (B) Model performance assessed by lesionwise sensitivity and PPV statistics across the test set, compared to the default disease-invariant model (green). (Error bars in each bar graph represent +/- 1 standard error of the mean across patients. \* =  $p < 0.05$  in comparison with the default FLAIR model after multiple comparison correction. # =  $p < 0.05$  in comparison with alternative training paradigm (fine-tune vs de-novo model) using the same training dataset size, after multiple comparison correction. FLAIR = fluid attenuation inversion recovery; PPV = positive predictive value.