Table S2. Detailed performance of the fully automated artificial intelligence (AI) approach for detecting nascent geographic atrophy (nGA) in the combined validation test sets.

AUC	AUPRC	Training and Tuning Specificity[#]	Precision	Accuracy	F1
All Visits					
		0.99	0.73 (0.54 to 0.89)	0.97 (0.95 to 0.98)	0.74 (0.61 to 0.84)
0.94 (0.90 to 0.97)	0.71* (0.56 to 0.87)	0.98	0.65 (0.48 to 0.81)	0.96 (0.94 to 0.98)	0.74 (0.61 to 0.85)
		0.95	0.50 (0.37 to 0.75)	0.94 (0.91 to 0.96)	0.65 (0.52 to 0.76)
At First Development					
		0.99	0.41 (0.23 to 0.65)	0.97 (0.95 to 0.98)	0.45 (0.31 to 0.60)
0.91 (0.86 to 0.96)	0.41 (0.27 to 0.63)	0.98	0.38 (0.23 to 0.57)	0.97 (0.95 to 0.98)	0.50 (0.35 to 0.65)
		0.95	0.27 (0.17 to 0.39)	0.94 (0.92 to 0.97)	0.40 (0.28 to 0.54)

All values are presented as the estimate (95% confidence interval); [#] = based on threshold determined from the training and tuning sets; Note: Recall is reported in Table 1 as Sensitivity. *0.36 if without imagenet pretraining.