



**Supplemental Figure S5.** Comparison of ground truth labels and AI outputs longitudinally on 3 typical eyes. The light gray shade indicates the interval diagnosed with nGA by human experts, and dark gray indicates cRORA (also considered as nGA in this context). The red curve plots the predicted confidence score of predicting being nGA by the fully automated AI approach. The blue curve plots the number of ground truth nGA B-scans. (A) (B) and (C) are cases where the fully automated AI approach correctly detected nGA onset, late and early. The two numbers in parentheses on each data point are the number of ground truth nGA B-scans vs. AI predicted nGA B-scans. A linear mixed model reveals that the number of OCT B-scans with nGA over time increase significantly based on the manual grading (2.0 B-scans/year, 95% CI = 0.7 to 3.4,  $P = 0.004$ ) and model output (2.0 B-scans/year, 95% CI = 1.0 to 3.1,  $P < 0.001$ ) amongst the 40 eyes that developed nGA in this study.