Supplementary

Section I:
Graph construction

Graphs are mathematical constructs comprising nodes and edges that enable modeling relationships between objects. Here, 'objects' refer to leakage patches, whose centroids are used as nodes for construction of different kinds fully connected graphs. Voronoi-, Delaunay- and Minimum Spanning Tree (MST)-based graph tessellations were used to describe the spatial architecture of leakage patterns. Details are provided in the Supplementary document (Section I). Once these graphs are constructed, features are derived from these fully connected network constructs to characterize the inter-leakage distances, leakage connectivity, and global compactness. For example, average of Voronoi area may capture the overall density of the leakage spots; variance of connecting edges may capture the disorder pattern of the leakage spots.

Given a set of leakage centroid in the plane, s_1, s_2, . . . s_n, the Voronoi diagram partitions the plane into regions where the region associated with centroid s_i is the set of centroids in the plane that are closer to site s_i than any other. The Delaunay triangulation of a set of sites is a specific triangulation that is the "dual" of the Voronoi diagram. Two sites are connected by an edge in the Delaunay triangulation only if they share a boundary in the Voronoi diagram. To calculate the MST, each edge weight is taken into account and a tree producing the smallest sum of edge weights is created. The vectors connecting the centroids of the leakage spots are the edges of the graphs, which traverse across the retinal region.

Feature Set	Number of Features	Feature Description
Leakage graphs	51	Voronoi Diagram: polygon area, perimeter, chord length; Delaunay Triangulation: Triangle side length, area; Minimum Spanning Tree: Edge length statistics; Nearest Neighbors: Density of leakage spots, distance too nearest leakage.
Leakage Morphology	100	Area, Mean Intensity/Intensity Range of leakage, Mean Intensity/Intensity Range Around leakage, Eccentricity Perimeter, Smoothness, Invariant Moment 1-7, Fractal Dimension, Fourier Descriptor 1-10 (Mean, Std Dev, Median, Range, Skewness, Kurtosis of Each

Supplementary Table I: *Leakage graph and morphology features*

Section II:
Vessel Tortuosity Computation

This method transforms the image space into an accumulator space, and was previously introduced to distinguish adenocarcinomas from granulomas, and to predict response to neoadjuvant chemotherapy in breast cancer.[1] The tortuosity features consist of the first order statistics (mean, median, variance, skewness and kurtosis) of maximum Hough

peak orientations computed in a sliding fashion across vessel projections summarizing vasculature orientation in the XY -plane.

Section III:

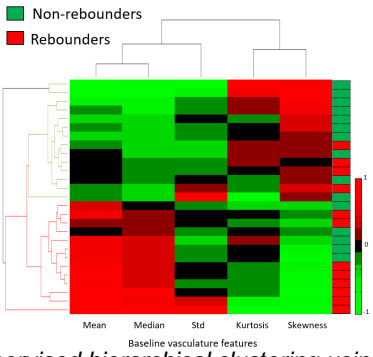
Statistical Analysis

Minimum redundancy maximum relevance (mRMR) test was implemented as a feature selection method. In the first experiment, features summarizing different morphological and graph-based attributes of leakage patches, were obtained for each leakage-segmented baseline FA image. To determine leakage features that best discriminated favorable responders from non-responders (at the first 8-week challenge), top 3 features were first determined using mRMR in a 3-fold cross validated setting over 100 runs. The top 3 features in each fold and run were then used to train a linear discriminant analysis (LDA) classifier[2] within the same run. Next, to identify quantitative vascular tortuosity measures that best discriminated favorable responders from non-responders on UWFA, similar to the first experiment, top 3 tortuosity features were used to train a LDA classifier in a cross-validated approach. Subsequently, in order to determine early response to therapy after administration of anti-VEGF for the first 4 months, we computed differences in tortuosity measures (delta-tortuosity features) between the first and the fourth visits. Delta leakage evaluation was also considered, however due to the dramatic reduction of leakage foci following therapy, this assessment was not feasible. The top discriminating delta-tortuosity measures were then used in conjunction with an LDA classifier to predict early response to anti-VEGF. Clinical parameters such as central subfield retinal thickness, macular volume, vessel area, vessel length, total leakage area, total number of leakage spots, and letter scores were also evaluated at baseline. The statistical significance of the different clinical features is computed for the two groups using a Wilcoxon rank-sum test.

Section IV:

Unsupervised clustering analysis

In addition to the supervised classification analysis, we employed unsupervised clustering approaches to measure the efficacy of the features in distinguishing the two groups. Specifically, hierarchical clustering was used to evaluate the discriminative ability of the different groups of features. Using the 51 baseline leakage features, the unsupervised hierarchical clustering did not result in clean clusters. Using the baseline vasculature features, comprising the first order statistics (mean, median, variance, skewness, and kurtosis) of maximum Hough peak orientations, for hierarchical clustering (Figure I), two clusters stand out: Cluster 1 (green) has a preponderance of non-rebounders (72%), Cluster 2, similarly, is the group corresponding to rebounders (62%). This gives an overall accuracy of 67% - which is lower than the supervised classification results, but fairly close. As may be observed, the rebounder group exhibits higher mean and median peak orientations as compared to the non-rebounders. Kurtosis and skewness values are higher for the non-rebounders while not much difference is observed in the standard deviation values between the two groups.



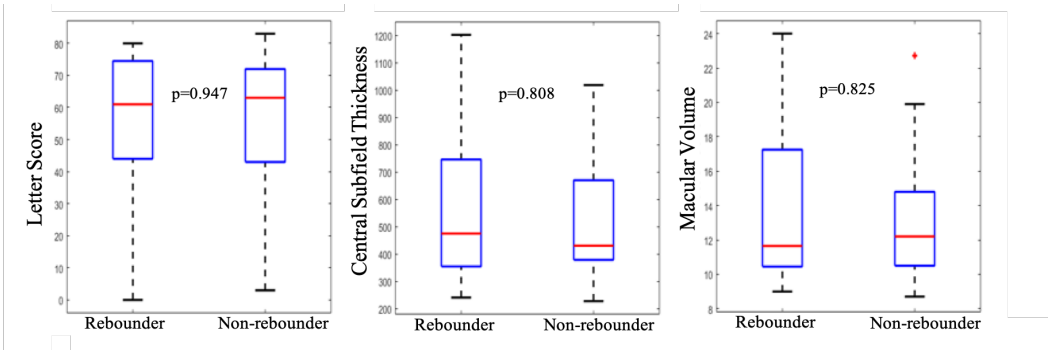
Supplementary Figure I: *Unsupervised hierarchical clustering using baseline vasculature tortuosity*

Section V:

Comparative Assessment with Clinical Parameters (letter score, central subfield thickness and macular volume) for Predicting Interval Tolerance

When evaluating the role of baseline clinical and more traditional imaging metrics, there were no significant differences in baseline features that were associated with rebound behavior, including ischemic index and underlying diagnosis. As shown in Supplementary Figure II, none of the three clinical parameters *(letter score, central subfield thickness and macular volume)* were statistically significantly different between the two groups of patients, with p-values of 0.947, 0.825 and 0.808, respectively. Upon running an LDA in conjunction with the three clinical parameters, we obtained an AUC of .42 ± .09 and an accuracy of .51 ± .05. A similar analysis was performed on other clinical measurements taken at baseline, including: vessel area, vessel length, total leakage area and total number of leakage spots. The LDA classifier yielded results of an AUC of .59 ± .07, and an accuracy of .66 ± .03.

Change in letter score was the only longitudinal clinical feature that showed a statistically significant difference between the two groups. The AUC and accuracy using these three features were found to be 0.48±0.03 and 0.58±0.08, respectively. Supplementary Table II provides the performance metrics obtained in the three different experiments.



Supplementary Figure II: Conventional Clinical Feature Evaluation. *Box and whisker plots of the following baseline clinical parameters (from left to right): Letter Score, Central Subfield Thickness and Macula Volume*

Features		AUC	Accuracy	Specificity	Sensitivity
Graph and morphological features	Baseline	0.77 ± 0.14	0.77 ± 0.11	0.80 ± 0.13	0.75 ± 0.15
Vasculature Features	Baseline	0.73 ± 0.10	0.74 ± 0.08	0.73 ± 0.18	0.74 ± 0.21
	Delta	0.73 ± 0.08	0.77 ± 0.07	0.81 ± 0.09	0.74 ± 0.10
Clinical Features	Baseline	0.42 ± 0.09	0.51 ± 0.05	0.78 ± 0.23	0.32 ± 0.23
	Delta	0.48 ± 0.03	0.58 ± 0.08	0.30 ± 0.25	0.79 ± 0.30

Supplementary Table II. *Performance metrics using image-derived markers and clinical features*

References

1. Braman N, Prasanna P, Alilou M, Beig N, Madabhushi A. Vascular Network Organization via Hough Transform (VaNgOGH): A Novel Radiomic Biomarker for Diagnosis and Treatment Response. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Lecture Notes in Computer Science. Springer International Publishing; 2018:803-811. doi:10.1007/978-3-030-00934-2_89

2. Park CH, Park H. A comparison of generalized linear discriminant analysis algorithms. *Pattern Recognit*. 2008;41(3):1083-1097. doi:10.1016/j.patcog.2007.07.022