Supplemental Document

Biomedical Optics EXPRESS

# Analysis of correlations between local geographic atrophy growth rates and local OCT angiography-measured choriocapillaris flow deficits: supplement

ERIC M. MOULT,[1,2] YINGYING SHI,[3] QINQIN ZHANG,[4] LIANG WANG,[3] RAHUL MAZUMDER,[5] SIYU CHEN,[1] ZHONGDI CHU,[4] WILLIAM FEUER,[3] NADIA K. WAHEED,[6] GIOVANNI GREGORI,[3] RUIKANG K. WANG,[4] PHILIP J. ROSENFELD,[3] AND JAMES G. FUJIMOTO[1,*]

[1] *Department of Electrical Engineering and Computer Science, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA*
[2] *Health Sciences and Technology, Harvard & Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA*
[3] *Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL 33136, USA*
[4] *Department of Bioengineering, University of Washington, Seattle, WA 98104, USA*
[5] *Sloan School of Management, Operations Research Center and Center for Statistics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA*
[6] *New England Eye Center, Tufts Medical Center, Boston, MA 02111, USA*
[*] *jgfuji@mit.edu*

This supplement published with The Optical Society on 1 July 2021 by The Authors under the terms of the Creative Commons Attribution 4.0 License in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: https://doi.org/10.6084/m9.figshare.14851674

Parent Article DOI: https://doi.org/10.1364/BOE.427819

## 1. *POST HOC* STATISTICAL ANALYSIS OF COVARIATES

Let $S_0$ ($n$ = 99) be the group of visit pairs having Pearson correlations that were not rejected at a Benjamini-Hochberg false discovery rate (FDR) of 0.2, and let $S_1$ ($n$ = 15) be the group of visit pairs having Pearson correlations that were rejected at a Benjamini-Hochberg FDR of 0.2. Furthermore, let $S_1^-$ ($n$ = 4) and $S_1^+$ ($n$ = 11) be subgroups of $S_1$ comprised of those visit pairs with negative sample correlations (Pearson's $r < 0$) and positive sample correlations (Pearson's $r > 0$), respectively. As described in the *Discussion*, we performed a *post hoc* comparison of covariate values from visit pairs in $S_0$ and $S_1^-$ to those from visit pairs in $S_1^+$ (see Fig. 8). Specifically, for each covariate, we tested the null hypothesis that the mean of the covariate values from the visit pairs in $S_0$ and $S_1^-$ was equal to the mean of the covariate values from the visit pairs in $S_1^+$. However, statistical assessment of this hypothesis is complicated by the non-independence of the visit pairs comprising $S_0$, $S_1^-$, and $S_1^+$. In particular, non-independence arises due to the inclusion of multiple visit pairs from the same eye (i.e., Visit 1-Visit 2, Visit 2-Visit 3, and Visit 1-Visit 3) and inclusion of multiple eyes from the same subject (i.e., OD and OS). This non-independence results in substantial correlations between the covariate values of certain visit pairs. For example, the baseline lesion size of a Visit 1-Visit 2 visit pair and a Visit 2-Visit 3 visit pair from the same eye are clearly positively correlated. As another example, there is a perfect positive correlation between subject age (at enrollment in the study) between all visit pairs from a given patient. Thus, standard testing procedures, such as the two-sample t-test, may yield p-values with inflated Type I errors. To mitigate the effects of such correlations, we used a simple constrained Monte Carlo permutation scheme to estimate the null distribution of the two-sample t-test. In this scheme, the null distribution of the two-sample, two-sided t-test was estimated by: (1) randomly permuting, subject to a constraint described in the subsequent paragraph, the covariate data into two simulated set of visit pairs, $\left[S_0 \cup S_1^-\right]_{\text{sim}}$ and $\left[S_1^+\right]_{\text{sim}}$, where the former is comprised of 103 visit pairs and the latter of 11 visit pairs (note that the notation $\cup$ denotes set union, and the $\left[\,\cdot\,\right]_{\text{sim}}$ denotes a simulated partition, which is to be distinguished from the observed partition); (2) using $\left[S_0 \cup S_1^-\right]_{\text{sim}}$ and $\left[S_1^+\right]_{\text{sim}}$, computing the two-sample, two-sided t-test statistic; and (3) repeating steps 1-2 10,000 times. The estimated null distribution was then used to compute the p-value corresponding to the observed test statistic. The t-test statistic was computed using the MATLAB function `ttest2`, with flag `Vartype` set to `unequal`. No adjustment was made for multiple comparisons amongst the different covariates.

We now describe the permutation constraint, which was imposed to more closely match the correlation structures of the simulated partitions $\left[S_0 \cup S_1^-\right]_{\text{sim}}$ and $\left[S_1^+\right]_{\text{sim}}$ to those of the observed partitions $S_0 \cup S_1^-$ and $S_1^+$. Briefly, we permitted only those permutations generating partitions such that $\left[S_1^+\right]_{\text{sim}}$ had a subject and visit pair 'structure' matching that of $S_1^+$. More precisely, let $P_1$ be the set of 16 subjects who had only one of their eyes included in the study, and let $P_2$ be the set of the 11 subjects who had both eyes included. To form our simulated partitions, $\left[S_1^+\right]_{\text{sim}}$ was required to be composed of visit pairs satisfying the following 9 constraints: (1) the visit pairs come from 6 different subjects, which we denote Subject *A*-*E*; (2) Subject *A* is in $P_2$ and Subjects *B*-*E* are in $P_1$; (3) 4 visit pairs are from Subject *A*, 2 visit pairs are from Subject *B*, 2 visit pairs are from Subject *C*, 1 visit pair is from Subject *D*, 1 visit pair is from Subject *E*, and 1 visit pair is from Subject *F*; (4) of the 4 visit pairs from Subject *A*, 2 are Visit 1-Visit 2 and Visit 1-Visit 3 visit pairs from one eye (e.g., OD), and 2 are Visit 2-Visit 3 and Visit 1-Visit 3 visit pairs from the other eye (e.g., OS)—these visit pairs are 'similar' to those of Eye 24 and Eye 26; (5) the visit pairs from Subject *B* are Visit 1-Visit 2 and Visit 1-Visit 3 visit pairs—these visit pairs are similar to those of Eye 16; (6) the visit pairs from Subject *C* are Visit 2-Visit 3 and Visit 1-Visit 3 visit pairs—these visit pairs are similar to those of Eye 23; (7) the visit pair from Subject *D* is a Visit 1-Visit 2 visit pair—this visit pair is similar to that of Eye 2; (8) the visit pair from Subject *E* is a Visit 1-Visit 3 visit pair—this visit pair is similar to that of Eye 10; and, (9) the visit pair from Subject *F* is a Visit 1-Visit 2 visit pair—this visit pair is similar to that of Eye 15. These constraints are shown graphically in Fig. S1. Note that this permutation structure does not require that the distribution of OD and OS eyes comprising $\left[S_0 \cup S_1^-\right]_{\text{sim}}$ and $\left[S_1^+\right]_{\text{sim}}$ matches that of $S_0 \cup S_1^-$ and $S_1^+$; however, we do not have reason to suspect that any of the covariate values systematically vary with respect to laterality.

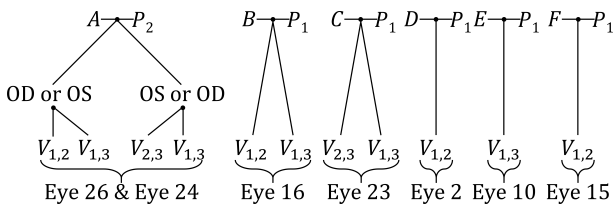


**Fig. S1.** Permutation constraint on group $\left[S_1^+\right]_{\text{sim}}$. In this diagram, $A$ through $F$ denote distinct subjects in the study; $V_{1,2}$, $V_{2,3}$, and $V_{1,3}$ denote the Visit 1-Visit 2, Visit 2-Visit 3, and Visit 1-Visit 3 visit pairs, respectively; $P_1$ denotes the set of 16 subjects who had only one of their eyes included in the study, and $P_2$ denotes the set of the 11 subjects who had both eyes included; and lines indicate correspondence/membership (e.g., between a visit pair that belongs to a given eye and subject). The leaf nodes of the diagram constitute a set of visit pairs having a 'structure' that is similar, in the sense described in the text, to the eyes of $S_1^+$; in particular, the corresponding eyes of $S_1^+$ are indicated underneath the brackets. A permissible permutation $\left[S_1^+\right]_{\text{sim}}$ is comprised of 11 visit pairs that can be arranged so as to satisfy the leaf nodes of this diagram.