Supplementary Methods and Data

This document supplements the article *"Inter-ocular Circumpapillary Retinal Nerve Fiber Layer Thickness Asymmetry and its Relationship to Age and Scanning Radius"* (submitted for publication)

Neda Baniasadi[1], *...*[1]

[1]Schepens Eye Research Institute, Harvard Medical School, Boston, MA

Calculation of the norms

All calculations described below have been performed with the statistical programming language R[1] and its `gamlss` library.[2,3] In order to generate the normative models of the inter-ocular retinal nerve fiber layer thickness (RNFLT) differences (RNFLT of left eye minus right eye), to each of the 768 circumpapillary measurement locations, a *Gaussian distribution* was fitted. The two parameters of this distribution, namely mean μ and standard deviation σ, are represented as functions of age a and the difference of the estimated true scanning radii of left minus right eye Δr. In particular, the optimal functions $\mu(a, \Delta r)$ and $\sigma(a, \Delta r)$ at each of the 768 locations were determined by a model comparison that included all models generated by any combination of third grade polynomials of a and Δr.

The simplest model is the "intercept only" model, i.e., for the mean, $\mu(a, \Delta r) = \beta_0$, which means, RNFLT difference at that location would not depend on age or radius difference. The most complex model for the mean, on the other hand, is the following six-parametric model:

$$\mu(a, \Delta r) = \beta_0 + \beta_1 a + \beta_2 a^2 + \beta_3 a^3 + \beta_4 \Delta r + \beta_5 \Delta r^2 + \beta_6 \Delta r^3$$

The models for σ are analogous. This means, at each of the 768 models, all models consisting of any combination of these six parameters were fitted to μ and σ. Therefore, in total, $2^{6+6} = 4,096$ models were calculated at each of the 768 locations. For each location, the optimal model out of these 4,096 models was determined by the Akaike information criterion (AIC).[4]

In other words, mean and standard deviation of the RNFLT difference distributions could be independent of age or radius difference, or could linearly, quadratically, or cubically depend on any of these two parameters. Model selection by AIC aimed to determine the respective optimal dependency of the RNFLT difference norms separately for each of the 768 locations.

Each row of table `rnflt_difference_norms.csv` contains the optimal model for the respective angular location, with those parameters set to 0 that were not part of the respective optimal model. Using the respective table row, μ and σ of the normative distribution at angle `angle` for a given age a and radius difference Δr are determined by

$$\begin{aligned}\mu = {} & \text{mu.intercept} + a \cdot \text{mu.age} + a^2 \cdot \text{mu.age2} + a^3 \cdot \text{mu.age3} \\ & + \Delta r \cdot \text{mu.rad} + \Delta r^2 \cdot \text{mu.rad2} + \Delta r^3 \cdot \text{mu.rad3}\end{aligned}$$

and

$$\begin{aligned}\sigma = {} & \exp(\text{sigma.intercept} + a \cdot \text{sigma.age} + a^2 \cdot \text{sigma.age2} + a^3 \cdot \text{sigma.age3} \\ & + \Delta r \cdot \text{sigma.rad} + \Delta r^2 \cdot \text{sigma.rad2} + \Delta r^3 \cdot \text{sigma.rad3})\end{aligned}$$

respectively.

In addition to the signed inter-ocular differences at the 768 circumpapillary locations, we also modeled the *absolute* differences for global thickness and the six sectors distinguished by the Spectralis machine (Heidelberg Engineering, Heidelberg, Germany). In contrast to the signed differences, which we approximated by a Gaussian distribution, absolute differences are non-negative, so that a distribution should be chosen that is bounded between

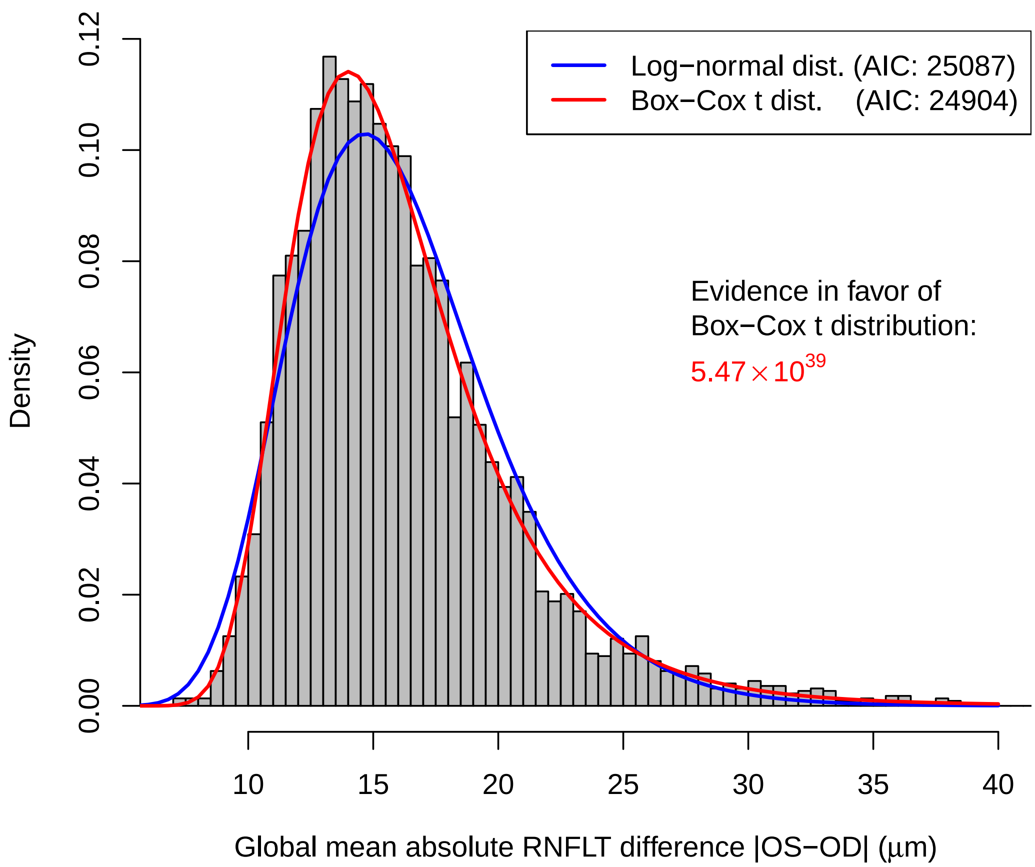

Figure 1. Comparison of distribution fits of log-normal (blue) and Box-Cox t distribution (red) to absolute global RNFLT differences. The gray histogram represents the empirical distribution of the absolute global RNFLT difference (i.e. the absolute RNFLT difference averaged over the entire measurement circle). The two curves illustrate the best fits for each distribution.

0 and $+\infty$. As a pre-step, we fitted two candidate distributions to the absolute difference of global RNFLT, namely the log-normal distribution and the Box-Cox t distribution,[5] a distribution in $\Re^+$ with the following four parameters: location μ, dispersion σ, a parameter τ controlling skewness,

and a parameter ν controlling kurtosis. Analogous to modeling the signed differences, μ and σ were modeled as polynomial functions of age and radius difference, whereas τ and ν were represented by intercept-only models.

Fig. 1 illustrates the best fits of log-normal (blue) and Box-Cox t distribution (red) to absolute global RNFLT differences. The evidence ratio (see Burnham & Anderson,[6] p. 78) in favor of the Box-Cox t distribution, i.e. $e^{(\text{AIC lognormal}-\text{AIC Box-Cox})/2}$, exceeds 10^{39}, which means, the Box-Cox t distribution is a substantially better model. Hence, the Box-Cox t distribution was chosen for the normative models of absolute RNFLT difference.

Analogous to table `rnflt_difference_norms.csv`, each row of table `sector_abs_rnflt_difference_norms.csv` contains the optimal model for the respective sector and for global absolute RNFLT difference. To construct the four parameters of the Box-Cox t distribution from the table, the calculations of μ and σ follow exactly the corresponding descriptions for the signed differences and the Gaussian distribution above. Parameters ν and τ are directly represented in the table as nu.intercept and tau.intercept, respectively.

Usage of the R functions

The R functions for generating age and radius difference specific norms and for generating the color plots inspired by the Spectralis printout can be found in `sources.R`. They require the R libraries `plotrix` and `gamlss.dist`. If these libraries are not installed, the software will prompt and offer to install them.

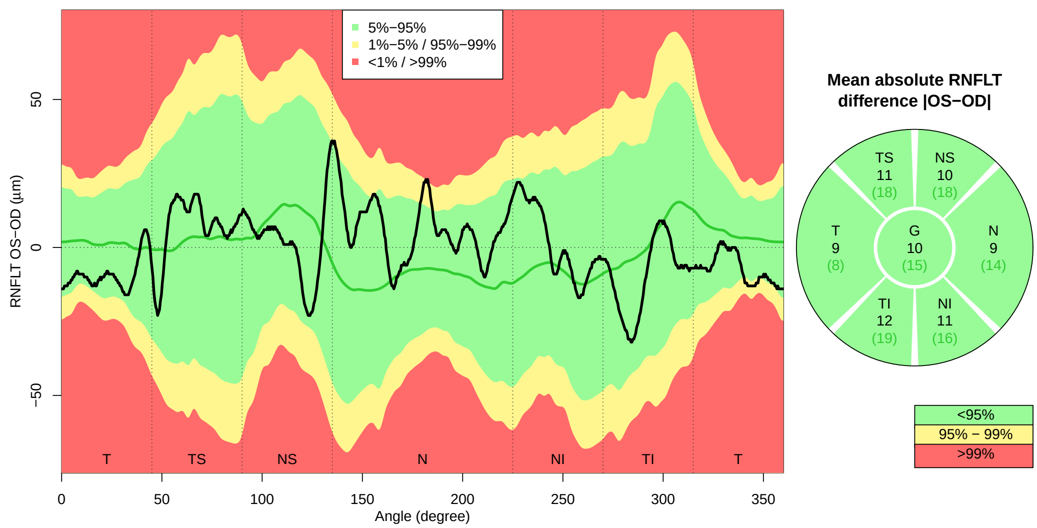

FIGURE 2. Plot generated by function difference.colorplot (see text for details).

When running the functions in R, make sure to have the two normative tables, `rnflt_difference_norms.csv` and `sector_abs_rnflt_difference_norms.csv`, in the same directory (or modify the sources accordingly). All functions are comprehensively documented in the source code. In addition, an illustrative RNFLT difference measurement of a healthy subject is added (file `rnfltdiff_example.csv`) which allows to test these functions. Here is a simple test:

```
source("sources.R") # load the functions
meas = read.csv("rnfltdiff_example.csv") # load the example
rnfltdiff = meas[-c(1,2)] # the RNFLT differences
# this generates a color plot similar to the Spectralis printout:
difference.colorplot(rnfltdiff, meas$age, meas$radiusdiff)
```

The final `difference.colorplot` call generates Fig. 2, where the black line represents the actual RNFLT difference measurement and the dark green line and the shaded areas the age and radius difference specific norms for this subject.

References

1. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

2. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *J R Stat Soc Ser C Appl Stat*. 2005;54:507–544.

3. Stasinopoulos DM, Rigby RA. Generalized additive models for location scale and shape (GAMLSS) in R. *J Stat Softw*. 2007;23.

4. Akaike H. New look at statistical-model identification. *IEEE Trans Automat Contr*. 1974;AC19:716–723.

5. Rigby RA, Stasinopoulos DM. Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*. 2006; 6:209–229.

6. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference, 2nd Ed.* New York: Springer, 2002.