Reliability of retinal pathology quantification in age-related macular degeneration: Implications for clinical trials and machine learning applications

– Supplementary Material –

Philipp L. Müller[1,2,3], Bart Liefers[1,4,5], Tim Treis[6], Filipa Gomes Rodrigues[1,2], Abraham Olvera-Barrios[1,2], Bobby Paul[7], Narendra Dhingra[8], Andrew Lotery[9], Clare Bailey[10], Paul Taylor[11], Clarisa I. Sánchez[4,5,12], Adnan Tufail[1,2,*]

[1] Moorfields Eye Hospital NHS Foundation Trust, London, UK
[2] Institute of Ophthalmology, University College London, London, UK
[3] Department of Ophthalmology, University of Bonn, Bonn, Germany
[4] Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands
[5] Department of Ophthalmology, Radboud University Medical Center, Nijmegen, The Netherlands
[6] BioQuant, University of Heidelberg, Heidelberg, Germany
[7] Barking, Havering and Redbridge University Hospitals NHS Trust, Romford, UK
[8] Mid Yorkshire Hospitals NHS Trust, Wakefield, UK
[9] University Hospital Southampton NHS Foundation Trust, Southampton, UK
[10] University Hospitals Bristol NHS Foundation Trust, Bristol, UK
[11] Institute of Health Informatics, University College London, London, UK
[12] Informatics Institute, Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands

Supplementary Figure S1: Annotation tool
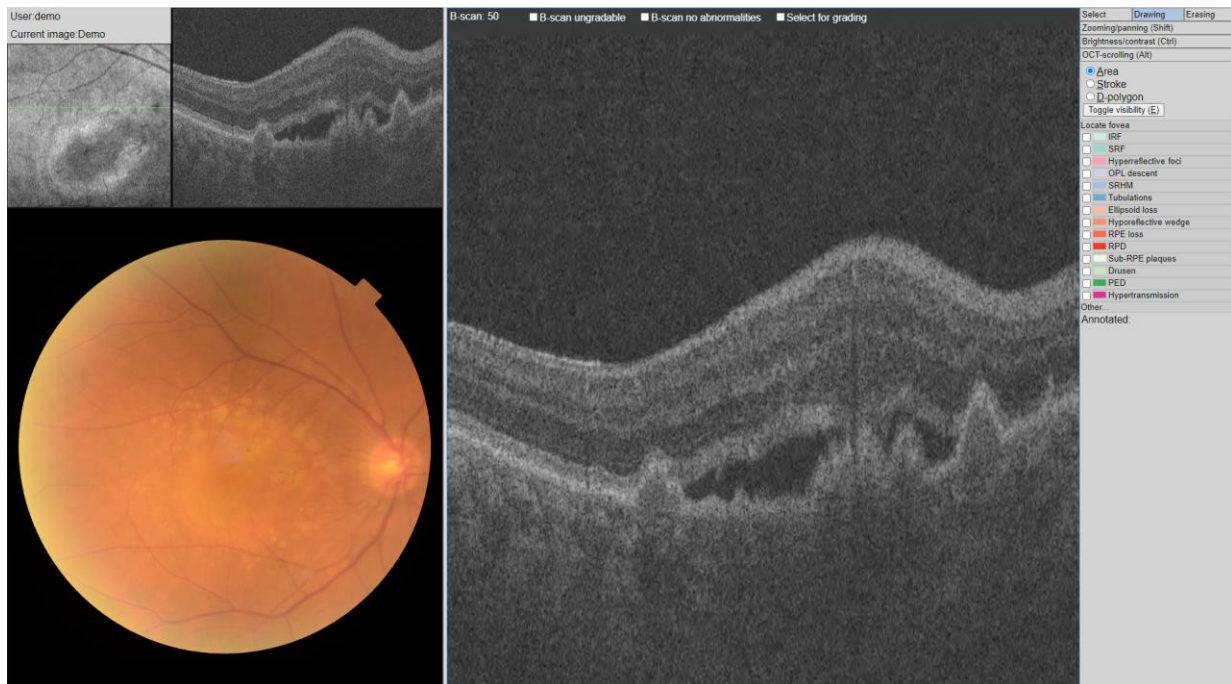
The custom-build platform for annotations is demonstrated. The annotations were made in the highlighted optical coherence tomography (OCT) B-scan (middle) by choosing the respective label (right). Additional labels not covered by the default setup could be added under "Other…" as free text. According to the label, the features were either annotated as area, stroke (i.e., line or point), or D-polygon. Wrong annotations could be deleted by selecting "Erasing" in the top-right corner. Ungradable B-scans or those without abnormalities could be labeled using the boxes at the top. On the top left the infrared reflectance image showed the location of the respective B-scan as green line next to a small preview of the B-scan. During the annotations, the reader could scroll through the other scans of the eye for context. The according color fundus photography was visualized at the bottom left.

Supplementary Figure S2: Inter-reader agreement for the measures of the area of drusen
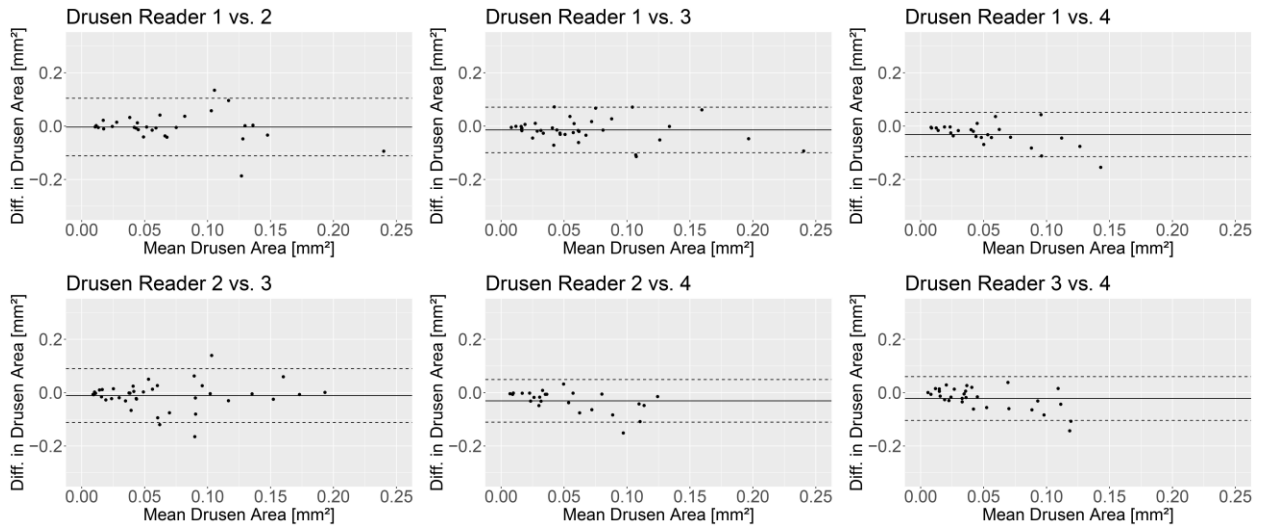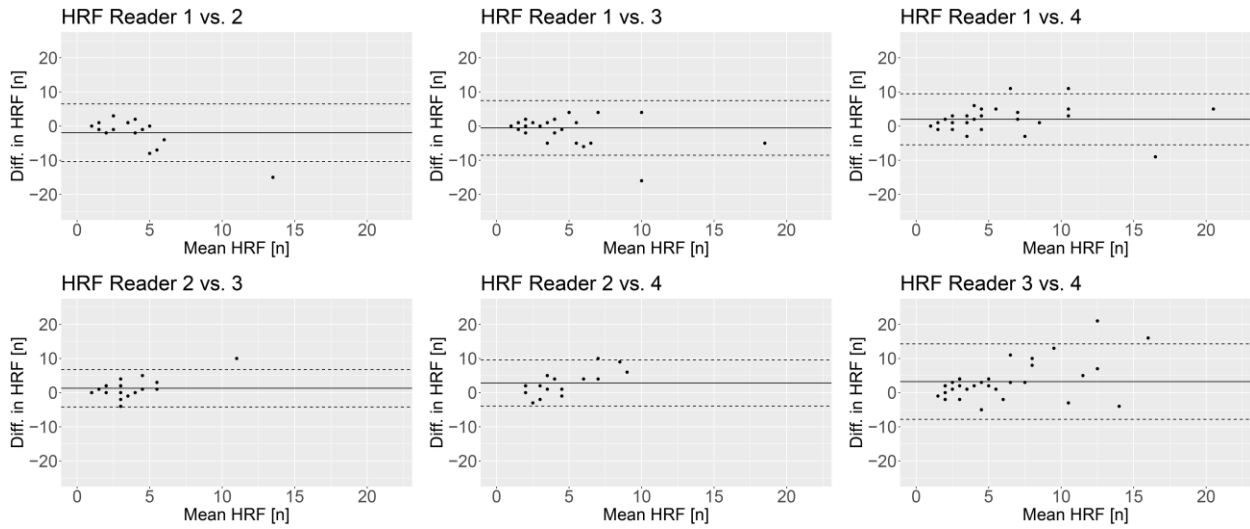

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

Supplementary Figure S3: Inter-reader agreement for the measures of the extent of ellipsoid zone (EZ) loss
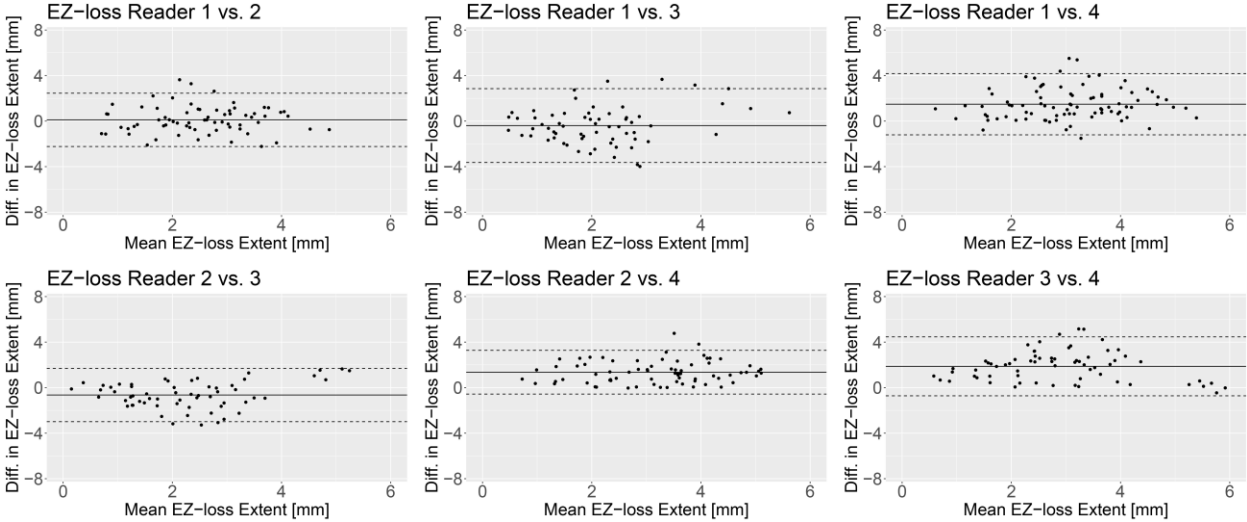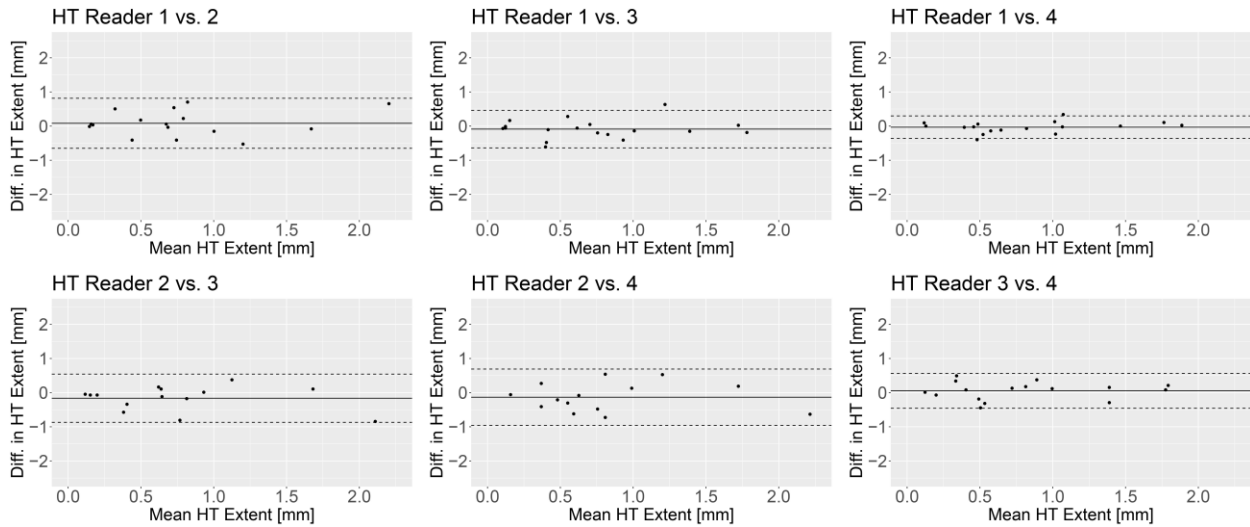

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

Supplementary Figure S4: Inter-reader agreement for the measures of intraretinal hyperreflective foci (HRF)
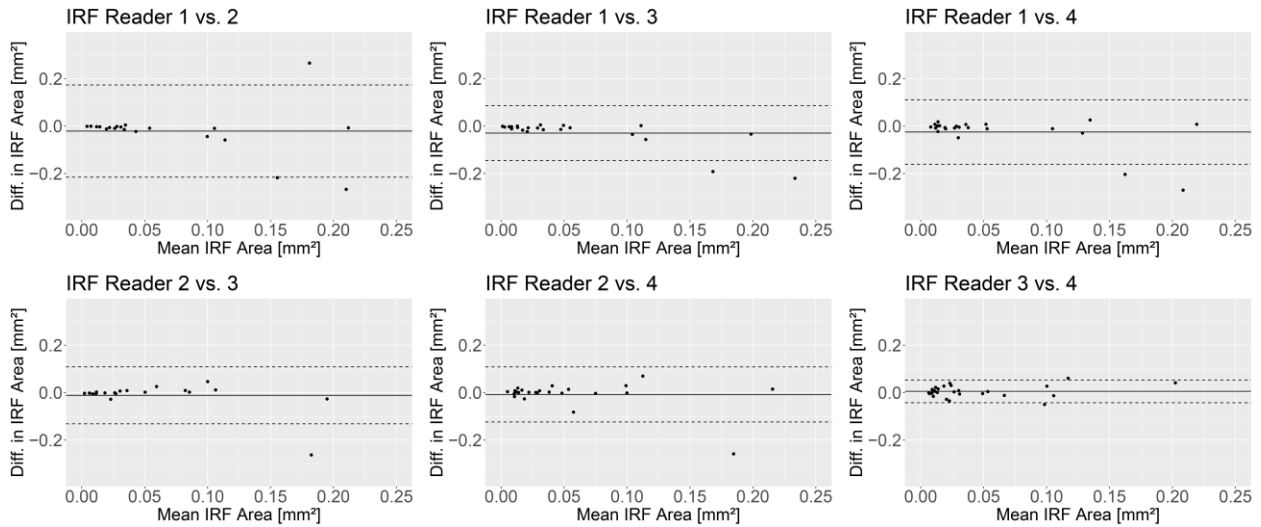

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

4

Supplementary Figure S5: Inter-reader agreement for the measures of the extent of OCT signal hypertransmission (HT)
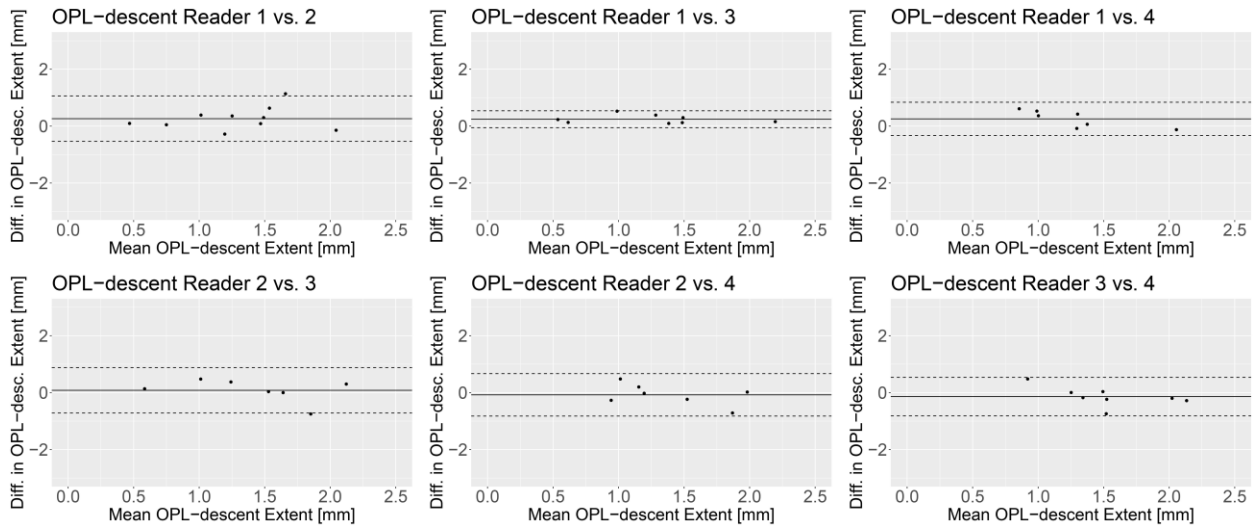

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

Supplementary Figure S6: Inter-reader agreement for the measures of the area of intraretinal fluid (IRF)

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.
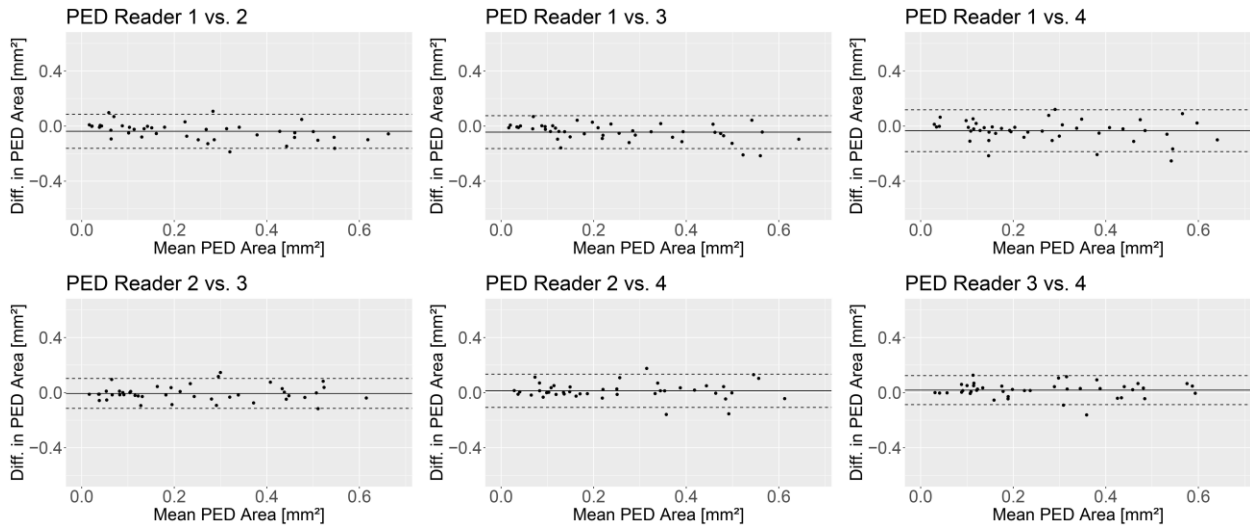
Supplementary Figure S7: Inter-reader agreement for the measures of the extent of outer plexiform layer (OPL) descent

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

Supplementary Figure S8: Inter-reader agreement for the measures of the area of pigment epithelial detachment (PED)

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.
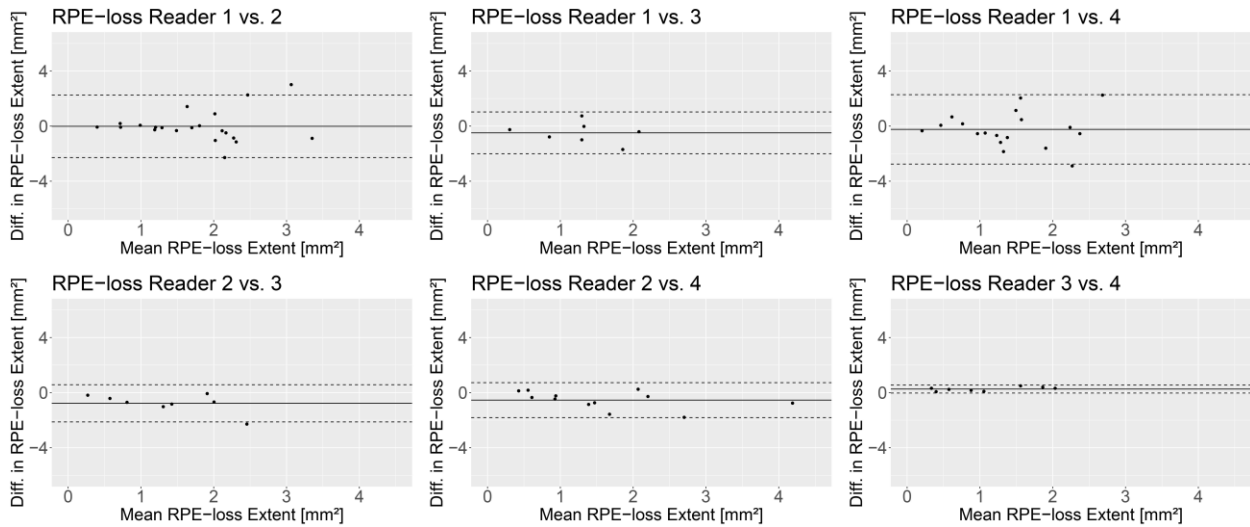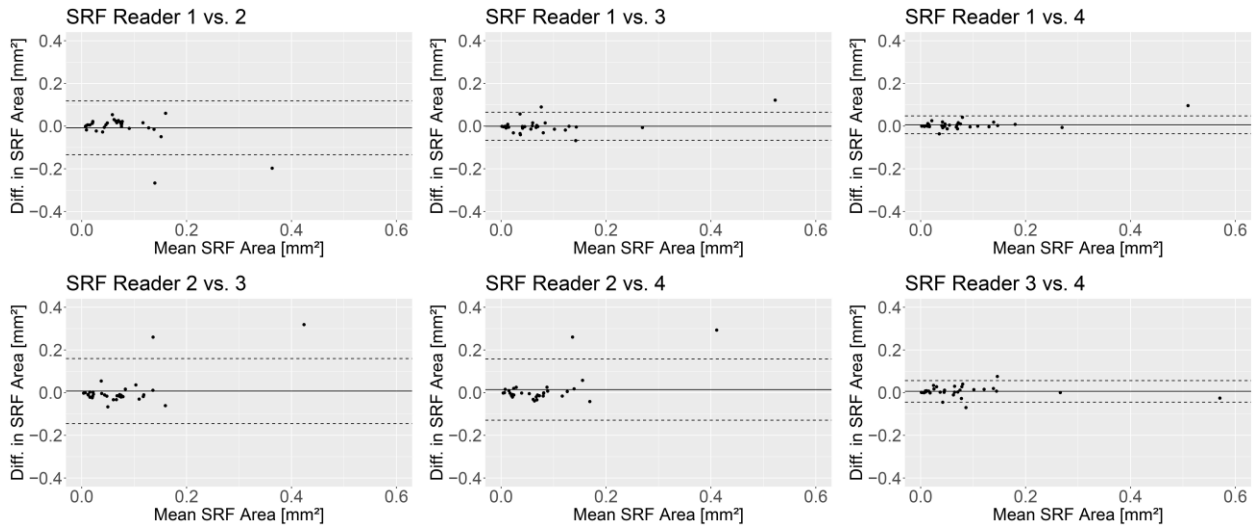
Supplementary Figure S9: Inter-reader agreement for the measures of the extent of retinal pigment epithelium (RPE) loss
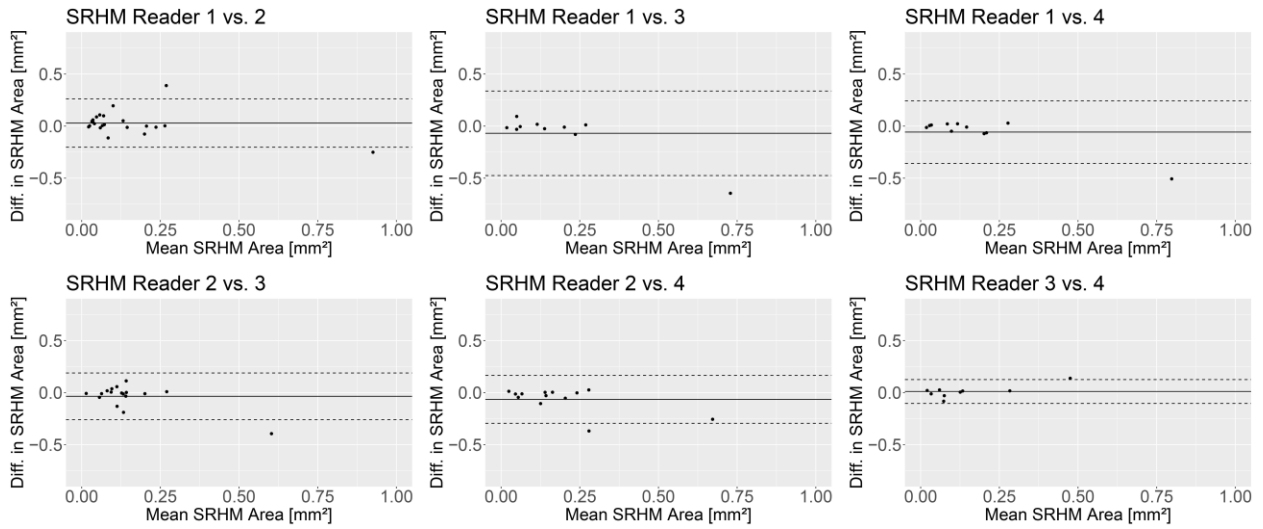

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

Supplementary Figure S10: Inter-reader agreement for the measures of the area of subretinal fluid (SRF)

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.

Supplementary Figure S11: Inter-reader agreement for the measures of the area of subretinal hyperreflective material (SRHM)

The Bland-Altman plots demonstrate the measurement differences (diff.) of two readers plotted against their mean. The solid line indicates the mean difference and the dashed lines indicate the 95% limits of agreement. There were no systematic differences between the readers.