Photoreceptor Cell Death Analysis and Replacement Therapies: Transplantation, Neuroprotection, and Immune Modulation.

Daniel Enrique Maidana

DOCTORAL THESIS

Photoreceptor Cell Death Analysis and Replacement Therapies: Transplantation, Neuroprotection, and Immune Modulation.

Daniel Enrique Maidana, M.D.

UNIVERSITAT DE BARCELONA

2020

Thesis for the Degree of Doctor of Philosophy

Photoreceptor Cell Death Analysis and Replacement Therapies: Transplantation, Neuroprotection, and Immune Modulation.

Medicine and Translational Research PhD Program

University of Barcelona School of Medicine

Student

Daniel Enrique Maidana, M.D.

Thesis Advisors

Josep Maria Caminal Mitjana, M.D., Ph.D.

Department of Ophthalmology, Bellvitge University Hospital, University of Barcelona

Demetrios G. Vavvas, M.D., Ph.D.

Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School

Acknowledgements

This doctoral thesis is the result of the most extraordinary period of my life. Unique not only for the outcomes but for the experiences and lessons learned. Even more, I was fortunate enough to see these lessons shape my principles and values. Thus, I would like to thank the following people who have contributed to this experience:

Foremost, I would like to thank Bayer Healthcare for their generosity in funding this thesis and for giving me a unique research opportunity.

I would like to express my sincere gratitude to Dr. Demetrios G. Vavvas, M.D., Ph.D, for opening my eyes to not only a world of science but one of principles. His patience and support have inspired me to accomplish this work and continue to motivate me to face any challenge in work and life.

My appreciation to Dr. Josep María Caminal Mitjana, M.D., Ph.D., for inspiring me in my early residency days, to pursue a path of dedication and integrity in all our endeavors.

My immense gratitude to my colleagues, Hidetaka Matsumoto, Shoji Notomi, Takashi Ueta, Pavlina Tsoka, Peggy Bouzika, and Alberto Caligiana, for their stimulating discussions and valuable input. My sincere thanks to Christine Roberge, to facilitating my research. Finally, my sincere gratitude to Lucía Gonzalez-Buendía, for her critical help, wholehearted support, and endless patience.

I am indebted to my parents Ramon and Irene, and my sisters Carolina and Florencia, for their unconditional support. I'm extremely grateful to my brothers-in-law Juan and Gustavo. I would also like to extend my sincere gratitude to my partner Alissa Andraski, for graciously and patiently supporting my efforts. I am indebted to them for inspiring and showing me, that without passion, dedication, and generosity, life is simply an ephemeral walk through these lands.

Finally, this work is a product of my conviction to improve current therapies for the treatment of retinal degenerations, which leads thousands of people to blindness. Their struggle has been an inspiration to carry this work forward. Undoubtedly, this work had very challenging days with many shortcomings, but never as many as these patients must go through day by day. To them, I dedicate these efforts.

Daniel E. Maidana, M.D

Tuesday, September 29, 2020

"Science means constantly walking a tightrope between blind faith and curiosity,
between expertise and creativity,
between bias and openness,
between experience and epiphany,
between ambition and passion,
and between arrogance and conviction,
in short, between an old today and a new tomorrow."

Heinrich Rohrer
1986 Nobel Prize in Physics

List of Abbreviations

ILM: inner limiting membrane

RNFL: retinal nerve fiber layer

GCL: ganglion cell layer

IPL: inner plexiform layer

INL: inner nuclear layer

OPL: outer plexiform layer

ONL: outer nuclear layer

OLM: outer limiting membrane

IS: photoreceptor inner segments

OS: photoreceptors outer segments

RPE: retinal pigment epithelium

BM: basal membrane

BrM: Bruch's membrane

MC: Müller cells

RGC: retinal ganglion cells

HC: horizontal cell

BC: bipolar cell

rBC: rod bipolar cell

cBC: cone bipolar cell

cGMP: cyclic guanosine monophosphate

GDP: guanosine diphosphate

GTP: guanosine triphosphate

PDE: phosphodiesterase

CNGC: cyclic nucleotide-gated channels

dUTP: deoxyuridine triphosphate

TUNEL: terminal deoxynucleotidyl transferase (dUTP) nick end labeling

CoV: coefficient of variation

ID: index of dispersion

DNA: deoxyribonucleic acid

NMDA: N-methyl-D-aspartic acid

RGB: red green blue

TIFF: tagged image file format

JPEG: joint photographic experts group

ANOVA: analysis of variance

SD: standard deviation

SE: standard error

LoA = limit of agreement

TOST: two one-sided tests

DAPI: 4′,6-diamidino-2-phenylindole

TT: ThicknessTool

RIPK3: receptor-interacting protein kinase 3

AMD: age-related macular degeneration

RP: retinitis pigmentosa

TNF: tumor-necrosis factor

TLR: toll-like receptor

EGFP: enhanced green fluorescent protein

WT: wild-type

KO: knock-out

NOD: nucleotide-binding domain

NLR: NOD-like receptor

NALP3: NLR family pyrin domain containing 3

NRL: neural retina leucine zipper

MACS: magnetic-activated cell sorting

PDE6b: phosphodiesterase 6 beta

CRB1: crumbs cell polarity complex component 1

RPE65: retinal pigment epithelium-specific 65 kDa protein

BRN3A: brain-specific homeobox/POU domain protein 3 A

PKC-1α: protein kinase C 1 alpha

GFAP: glial fibrillary acidic protein

LDH: lactate dehydrogenase

IL: interleukin

CXCL: C-X-C motif chemokine

CCL: C-C motif chemokine

BMT: bone marrow transplant

FACS: fluorescence-activated cell sorting

HMGB1: high mobility group box 1

Outline

Table of Contents

List of Figures and Tables

Chapter I. Introduction

Chapter III. Automated Cell Death Analysis

Chapter IV. Photoreceptor Cell Death: Automated Retinal Layer Thickness Measurement

Chapter I

Photoreceptor Cell Death: Introduction

Chapter I. Introduction

1.1. Organization of the Vertebrate Retina

The vertebrate retina is comprised by two layers of distinct embryological original and function: the neural retina and the retinal pigment epithelium (RPE). The neural retina is a composed by the several layers of neurons and supportive glial cells. Seminal work from Richard Masland has thoroughly characterized the neuronal census of the retina.[1,2] Initial estimates of 55 cell types continues to increase, in parallel with our understanding of the visual processing performed by these cells.[2]

The neurosensory retina is formed by several layers formed by different cellular populations and supporting extracellular matrix: i) inner limiting membrane (ILM); ii) retinal nerve fiber layer (RNFL); iii) ganglion cell layer (GCL); iv) inner plexiform layer (IPL); v) inner nuclear layer (INL); vi) outer plexiform layer (OPL); vii) outer nuclear layer (ONL); viii) outer limiting membrane (OLM); ix) photoreceptor inner segments (IS) / outer segments (OS) layer.[3]

The inner limiting membrane is the outermost layer of the retina, which separates the neurosensory retina from the posterior hyaloid of the vitreous. Moreover, the ILM represents the basal membrane (BM) of Müller cells (MC).[4] As such, the end-feet of MCs are aligned across the ILM throughout the entire retina. The ILM is composed mostly of collagen IV, laminins, proteoglycans, and collagen XVIII.[4]

The retinal nerve fiber layer comprises the axons of the retinal ganglion cells (RGC) that will bundle up to form the optic nerve. In addition to the RGC axons, it contains retinal vessels, astrocytes, and Müller cell footplates.[5] The ganglion cell layer contains the nuclei of RGC which projects its axons to the RNFL and its dendrites to the IPL. In addition, several displaced amacrine cells can also be observed in this layer.[6]

The inner plexiform layer contains the amacrine cells, and the synapse of the RGC dendrites and the bipolar cell axons. The inner nuclear layer contains the nuclei of bipolar,

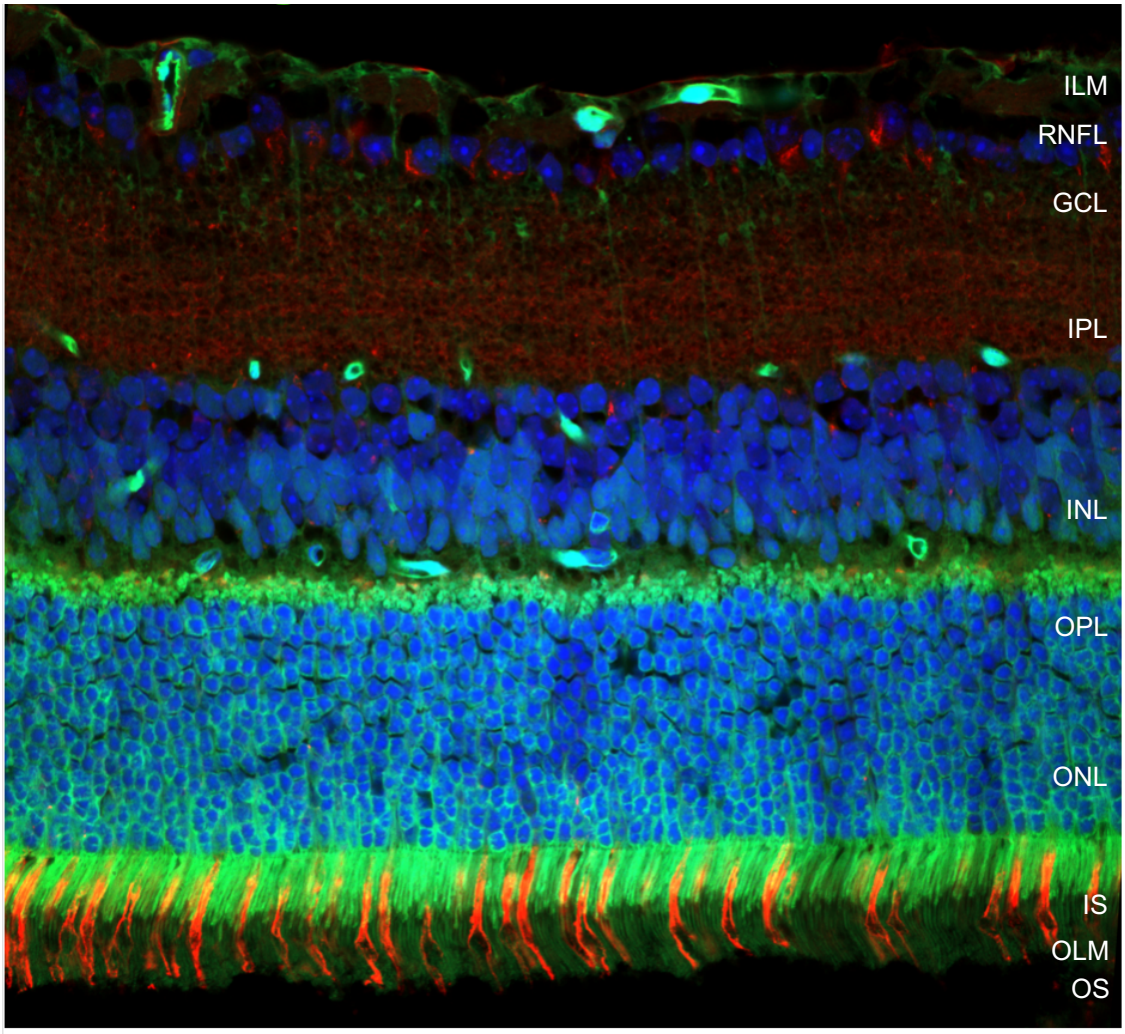


Figure 1. Organization of the Neurosensory Retina

ILM = inner limiting membrane; RNFL = retinal nerve fiber layer; GCL = ganglion cell layer; IPL = inner plexiform layer; INL = inner nuclear layer; OPL = outer plexiform layer; ONL = outer nuclear layer; OLM = outer limiting membrane; IS = photoreceptor inner segments; OS = photoreceptors outer segments.

amacrine, and horizontal cells (HC).[3] Bipolar cells (BC) can be classified according to which photoreceptor relays the signal to their dendrites. Rod BC (rBC) synapse with rods photoreceptors and cone BC (cBC) synapse with cone photoreceptors.[3] Bipolar cells can be further classified by their de/hyperpolarization pattern. The ON BC depolarize to increments in light intensity, and the OFF BC hyperpolarize in contrast.[3] More specific,

rBC are only ON BC, whereas cBC can be either ON or OFF BC.[3]

The outer plexiform layer contains the synapse of rods (spherule) and cones (pedicle) with their corresponding BC and HC. The outer nuclear layer contains the cell bodies of rod and cone photoreceptors. The outer limiting membrane separates the photoreceptor cell nuclei from their inner segments. The inner segment/outer segment layer contains rod and cone photoreceptors segments necessary for the sensory phototransduction of the visual system.

Finally, the retinal pigment epithelium is adjacent to the photoreceptor outer segments. It is composed by a monolayer of cuboidal pigmented cells which provide structural and functional support to photoreceptors. [7]

1.2. Molecular Events of Rod Phototransduction

Phototransduction is the transduction of the light stimulus into an electrical signal and occurs at the outer segments of rod photoreceptors.

At baseline scotopic conditions without any light stimuli, the rod remains in an inactive depolarized state. This is sustained by the influx of Na+ into the rod, though a cGMP-activated ion channel. Subsequently to the ion influx and at this baseline inactive depolarized state, the rod releases glutamate which subsequently binds to the mGluR6 receptor, hyperpolarizing and inhibiting bipolar cells.[8]

Under scotopic conditions, as Lamb and Pugh Jr describe in their Proctor Lecture, the phototransduction process can be understood as a 5 step cascade.[9]

Activation of Rhodopsin: Rhodopsin in a G-coupled receptor protein which has bound a light absorbing retinaldehyde. When a photon of light reaches the rods, it is absorbed at the outer segments though various rhodopsin molecules. The rhodopsin molecule is activated by isomerization of the 11-*cis* retinal to 11-all-*trans* configuration.

Activation of Transducin: The activated rhodopsin molecule will catalytically activate the heterotrimeric G protein (G) called Transducin, which exchanges GDP for GTP, and ultimately activating the G-Protein.

Activation of Phosphodiesterase (PDE): The activated transducin diffuses and contacts the regulatory PDE γ-subunit, activating PDE. The process is repeated in the second PDE γ-subunit, resulting in full activation of PDE

Hydrolysis of cGMP: The activated PDE hydrolyses cGMP, resulting in a dramatic drop in this secondary messenger intracellular concentration.

Closure of Cyclic Nucleotide-gated Channels (CGNC): When the cytosolic concentration of cGMP is drastically reduced, the CNGC-bound cGMP unbinds leading to CNGC closure. As a result, the influx of Na^{+} and Ca^{+2} ions is blocked, and the rod becomes hyperpolarized. Consequently, to the rod hyperpolarization, the release of inhibitory Glutamate is decreased.

1.3. Molecular Events of Rod Retinal Circuitry

The light s three principal rod circuit that connects the rods to the ganglion cells is comprised by different types of neurons, including: i) rod photoreceptor; ii) rod bipolar cell; iii) cone bipolar cell, iv) AII amacrine cell; v) ganglion cell.[2]

There are three well known circuits or pathways that relay the phototransduction signal generated at the rod photoreceptors to the RGC.[10,11,20–22,12–19]

The primary pathway is active when the rod signal is relayed through a synapse with rBC, which subsequently and exclusively synapse with AII amacrine cells (AII) at the inner plexiform layer. The AII cells will relay the signal to: i) ON-center cone cBC with sign-conserving electrical synapses through gap junctions; and ii) OFF-center cBC with sign-inverting chemical synapses thorough inhibitory glycinergic synapsis. The signal is then

relayed from the ON cBC to the corresponding ON RGC, and from the OFF cBC to the corresponding OFF RGC.[21,22]

In the secondary pathway, the rod signal is passed to cones via interconnecting gap junctions. Cones subsequently relay the signal to ON and OFF cBC and further on to ON and OFF RGCs.[21,22]

In the tertiary pathway, the rod signal is passed through a synapse with a unique ON cBC and presumably to an OFF cBC. Subsequently these project to ON and OFF RGC respectively.[21,22]

As different light intensities reach the rods, different pathways get activate depending on the light intensity, in a non-overlapping range.[22] The primary rod pathway is the most sensitive pathway carrying signals with the lowest threshold. Following, the secondary rod pathway is less sensitive and the tertiary pathway is the least sensitive of all three.[21,22]

1.4. Photoreceptor Cell Death

The key feature of retinal degeneration is photoreceptor cell death. Although in most cases, rods harbor the defective mutated genes which conditions their survival, cone photoreceptors ultimately die as well.[23] Therefore, it is reasonable to rescue this rod population, to prevent cone cell death and vision loss.

1.4.1. Photoreceptor Cell Death: Quantitation Methods

To effectively address retinal degenerations, an accurate estimate of cell death must be achieved first. From the many tools used to characterize and quantify retinal cell death, retinal thickness and the terminal deoxynucleotidyl transferase (dUTP) nick end labeling (TUNEL) assay are likely the most widespread methods.[24] Total retinal and nuclear layer thickness measurement can give an overall estimate of the retinal architecture, as the layer thickness is determined by the number of cells comprised in each layer. In parallel,

TUNEL assay detects fragmented or nicked DNA by means of a terminal deoxynucleotidyl transferase enzyme, which incorporates fluorescent-labeled dUTPs in damaged nucleic acid regions. In essence, TUNEL assay identifies cells that are undergoing cell death. To date, in order to quantitate retinal thickness and TUNEL assay results in tissue sections, manual thickness measurement and counting of TUNEL-positive (TUNEL+) cells, ideally by two masked observers is performed. However, this process is time consuming, prone to measurement errors, and not entirely reproducible. Therefore, an automated approach is of interest to address these difficulties, and to carry forward a more robust analysis of retinal degeneration.

1.4.2. Photoreceptor Cell Death: Replacement Therapies

Given reliable methods to quantitate cell death and retinal degeneration, we expect that consistent and effective translational approaches will follow. In that avenue, several emerging strategies including drug delivery, gene therapy, and photoreceptor transplant are being investigated to tackle photoreceptor cell death, with limited clinical success. Gene therapy has shown the most promise, but this approach requires targeted cells to still be viable in the native retina in order to transduce the gene of interest. This results in a narrow therapeutic window for retinal degenerative diseases. To overcome this limitation, photoreceptor transplant has been put forward as a repair strategy to tackle degenerating or already degenerated retinas. Unfortunately, efforts to achieve successful photoreceptor transplantation strategies have been limited by extremely limited efficiency.[25] Thus, it seems crucial to reflect and re-investigate the conditions necessary to fine tune photoreceptor transplantation.

Chapter II

Photoreceptor Cell Death: Hypotheses and Objectives

Chapter II. Hypotheses and Objectives

2.1. Project Goal

The objective of this thesis is to: i) develop consistent automated quantitative methods necessary for accurate photoreceptor cell death assessment; and ii) improve the outcome of photoreceptor transplantation cell-replacement therapies in retinal degeneration. We expect to address this technical bottleneck and subsequently advance the knowledge of photoreceptor transplantation therapies for retinal degeneration. For this purpose, we propose the following aims and hypothesis.

2.2. Aim 1: Photoreceptor Cell Death: Automated Cell Death

2.2.1. Objectives

The objective of this aim is to develop a novel ImageJ software tool for the automated detection of cell death in retinal cross-sections and to validate its performance with manual cell counting. We seek to address the pitfalls of manual quantitation in retinal degeneration by developing and validating an automated tool which performs in a precise, accurate, and fast manner.

2.2.2. Working Hypothesis

We defined the following alternative hypothesis for this aim:

H_1 = Automated cell death quantitation performance is superior to manual quantitation

We will test this hypothesis in the setting of cell death assays in an animal model of acquired retinal degeneration. We will evaluate the performance of ImageJ automated quantitation compared to manual observers. Superiority will be evaluated by agreement analysis, Bland-Altman plots, experimental validation, and time burden. Statistical

analyses are described in the *materials and methods* section for this aim.

2.3. Aim 2: Photoreceptor Cell Death: Automated Retinal Layer Thickness

2.3.1. Objectives

The objective of this aim is to develop a new ImageJ software to quantitate retinal layer thickness in retinal cross-sections, and similarly, to validate its performance with manual observers.

2.3.2. Working Hypothesis

We defined the following alternative hypothesis for this aim:

H_1 = *Automated retinal layer thickness performance is superior to manual quantitation*

We will evaluate this hypothesis in the setting of retinal thickness thinning in an animal model of retinal detachment. We will compare the performance of automated quantitation to manual observers. Similarly, superiority will be evaluated by agreement analysis, Bland-Altman plots, experimental validation, and time burden. Statistical analyses are described in the *materials and methods* section for this aim.

2.4. Aim 3: Photoreceptor Cell Death: Replacement Therapy

2.4.1. Objectives

Following the development of novel and automated quantitative methods in retinal degenerations, the objective of this aim is to carry forward a translational replacement therapy approach. We aim to optimize the survival of photoreceptor transplantation in an animal model of inherited retinal degeneration. We will target the receptor-interacting protein kinase 3 (RIP3/RIPK3), a protein deeply involved in necrotic cell death and inflammation, in both hosts and donors.

2.4.2. Working Hypothesis

We defined the following alternative hypothesis for this aim:

H_1 = RIP3K modulation in hosts and/or donors improves the survival of grafted photoreceptors

We will evaluate this hypothesis in an animal model of retinal degeneration subject to subretinal photoreceptor transplantation. We will compare photoreceptor survival in donors and hosts of wild type and transgenic animals. Similarly, superiority will be evaluated by number of grafted photoreceptors in the subretinal space at endpoint. Statistical analyses are described in the *materials and methods* section for this aim.

Chapter III

Photoreceptor Cell Death: Automated Cell Death Analysis

A Novel ImageJ Macro for Automated Cell Death Quantitation in the Retina

Daniel E. Maidana, Pavlina Tsoka, Bo Tian, Bernard Dib, Hidetaka Matsumoto, Keiko Kataoka, Haijiang Lin, Joan W. Miller, Demetrios G. Vavvas

Paper published in *Investigative Ophthalmology & Visual Science* (IOVS)

Chapter III. Automated Cell Death Analysis

3.1. Abstract

Purpose: TUNEL assay is widely used to evaluate cell death. Quantification of TUNEL positive (TUNEL+) cells in tissue sections is usually performed manually, ideally by two masked observers. This process is time consuming, prone to measurement errors, and not entirely reproducible. We present here an automated quantification approach to address these difficulties.

Materials and Methods: We developed an ImageJ macro to quantitate cell death by TUNEL assay in retinal cross-section images. The algorithm was coded using IJ1 scripting language on Fiji (http://imagej.net/TUNEL_Cell_Counter). To validate this tool, we selected a dataset of TUNEL assay digital images, calculated layer area and cell count manually by two observers and compared measurements between observers and macro.

Results: The automated macro segmented outer (ONL) and inner nuclear layers (INL) successfully. Automated $TUNEL^{+}$ cell counts were in-between counts of inexperienced and experienced observers. The intra-observer coefficient of variation (COV) ranged from 13.09% to 25.20%. The COV between both observers was 51.11 ± 25.83% for the ONL, and 56.07 ± 24.03% for the INL. When comparing observers vs. macro, COV was 23.37 ± 15.97% for the ONL, and 23.44 ± 18.56% for the INL.

Conclusions: We developed and validated an ImageJ macro that can be used as an accurate and precise quantitative tool for retina researchers, to achieve repeatable, unbiased, fast, and accurate cell death quantitation. We believe that this standardized measurement tool could be advantageous to compare results across different research groups, as it is freely available as open source.

2.2. Introduction

In recent years, research for degenerative retinal diseases has been focused on understanding cell death and ways to interfere with them. Several experimental animal models, such as diabetic retinopathy, retinitis pigmentosa, retinal detachment, or glaucoma, have revealed that cell death in the retina can occur through different modalities, including apoptosis and necroptosis among others, encompassing multiple inducers and pathways and displaying different morphological characteristics.[1-5] From the many tools to characterize and quantify cell death, terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) assay, is likely the most widespread method. This assay detects fragmented or nicked-DNA by means of a terminal deoxynucleotidyl transferase enzyme, which incorporates fluorescent-labeled dUTPs to damaged nucleic acid regions. First described by Gavrieli et al[6] and though not entirely specific, this assay has been proven to be the most sensitive method in detecting apoptosis.[7] However, since this assay detects DNA fragments regardless of the induced cell death pathway, it cannot distinguish between apoptotic or necroptotic forms of programmed cell death.[3,8-10] Despite its limitations, TUNEL assay remains the most widespread method used to screen for any form of programmed cell death to date.

In order to quantitate TUNEL assay results in tissue sections, manual counting of TUNEL positive ($TUNEL^+$) cells, ideally by two masked observers is performed.[11] Observers count $TUNEL^+$ cells and measure the respective retinal area, and results are expressed as either $TUNEL^+$ cells/area or $TUNEL^+$ cells/total cells. Evidently, this process is time consuming, prone to measurement errors, and not entirely reproducible. Therefore, an automated approach is of interest to address these difficulties. To the best of our knowledge, there is no automated cell death quantitation method available for this platform, capable of segmenting retinal layers and counting individual $TUNEL^+$ cells in the retina without performance variability.

The purpose of this study was to introduce a novel ImageJ macro for the automated detection of cell death in retinal cross-sections. In addition, we studied the differences in

grading between an experienced and inexperienced observer. Since the introduction of a new measurement method requires to be paralleled against the gold standard or commonly accepted method, we validated this macro by comparing its performance against two observers. We found that the ImageJ macro can achieve a fast and accurate retinal cell death quantitation.

3.3. Materials and Methods

3.3.1. Fluorescent Microscope Image Database

Digital images of TUNEL-stained retinal cross-sections were obtained from the Angiogenesis Laboratory fluorescent microscope database. In order to quantitate $TUNEL^+$ cells on the ONL, we randomly selected 90 images from an experimental murine retinal detachment model which induces photoreceptor cell death by subretinal injection of sodium hyaluronate.[12] For experimental comparison between groups, we selected 30 images from an experimental group (mammalian sterile 20-like kinase 2 (MST2) knock-out mice) compared to a control group (C57BL/6 mice), as previously published.[13] To assess INL $TUNEL^+$ cells, we randomly selected 90 images from an experimental N-Methyl-D-aspartic acid (NMDA)-induced excitotoxicity murine model using two different drug doses (10 and 100 nM). Eyes from these experimental models were enucleated and 7-10 μm thin sections were cut using a Leica CM1850 cryostat (Leica Microsystems, Bannockburn, IL). TUNEL staining was performed with ApopTag fluorescein direct in situ apoptosis detection kit (Millipore, MA), according to instructions of the manufacturer. Prior to mounting, cell nuclei were stained with TO-PRO 3 (Life Technologies, NY).

Since we sought to build an assorted and representative sample, we included images from the aforesaid models acquired by multiple independent researchers, which used different acquisition protocols. These images were acquired with Axio Imager M2 fluorescence microscope (Carl Zeiss Inc., NY) using a 20x/0.8 air objective, as we considered this

magnification was the most suitable to assess a reasonably large area without missing out on details in cell morphology. Image exclusion criteria were poor TUNEL staining quality, uneven image focus, ONL or INL layers not distinguishable, retina cross-section not centered on image frame, or significant shadowing. Finally, selected digital images were exported as 24-bit RGB image in an uncompressed tagged image file format (TIFF) for further analysis.

3.3.2. ImageJ Macro Script Programming

To develop an automated method to quantitate cell death by TUNEL assay in retinal cross-section images, we designed a macro for the ImageJ platform (Bethesda, MD) and coded the algorithm using IJ1 scripting language on Fiji (http://fiji.sc/Fiji). After TIFF file import, the most relevant image processing steps are detailed as follows:

Image scale setup. The image spatial scale (pixel to mm^2) was acquired from the microscope bundled software Zen version 1.1 (Carl Zeiss Inc., NY), in order to express results in mm^2. This scale value can be easily obtained from image metadata.

ONL and INL segmentation. From the native RGB digital image (Figure 1A), an 8-bit blue channel was extracted and a Gaussian blur filter was applied (Figure 1B-C), in order to reduce noise and detail.[14]

After contrast enhancement, we used Tsai moment-preserving thresholding method to obtain a binary image.[15] Following several erosion iterations to achieve a fitted segmentation, outlier pixels were removed to a specific radius. The largest area in the image was presumed to be the ONL area. Following, the segmented ONL region of interest (ROI) was subtracted from the processed blue channel. The INL area was segmented following the same approach as the ONL area (Figure 1D).

ONL and INL total cells. The ONL and INL ROIs were used to determine specific areas where total cells were counted. By determining the local maxima corresponding to cell nuclei, individual cells at these layers were identified and recorded (Figure 1E).

ONL and INL TUNEL^{+} cells. From the native RGB digital image, an 8-bit green channel was extracted, and background noise was subtracted, in order to achieve a cleaner image (Figure 1F). Following, we applied Tsai moment-preserving thresholding method to obtain a binary image and used binary watershed segmentation to separate any contiguous cells (Figure 1G). Individual cells were counted according to specific size and circularity.

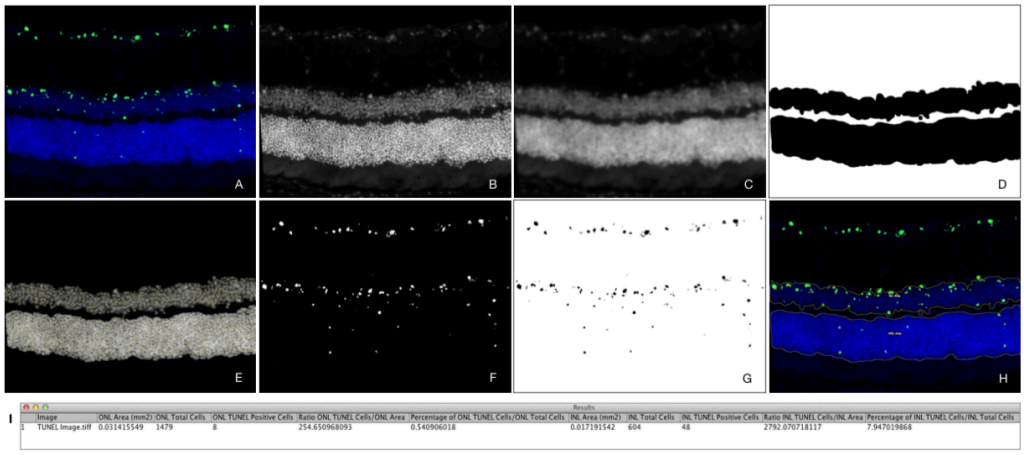


Figure 1. Segmentation of outer nuclear layer (ONL), inner nuclear layer (INL), and quantitation of TUNEL positive cells by ImageJ Macro.

From the native RGB image (A), an 8-bit blue channel was extracted (B) and a Gaussian blur filter was applied (C). Using Tsai moment-preserving thresholding method, we segmented ONL and INL layers (D), and by determining the local maxima corresponding to cell nuclei, individual cells at these layers were identified (E). The 8-bit green channel was extracted, and background noise was subtracted (F). Using Tsai moment-preserving thresholding method and binary watershed segmentation, we counted TUNEL+ cells (G). A JPEG image overlay was automatically created and exported for visual assessment of the quantitation (H). Results from the macro were reported (I) as area (mm2), total cells (count), TUNEL+ cells (count), ratio of TUNEL+ cells/area (count/mm2), and percentage of TUNEL+ cells over total cells (%), for both ONL and INL.

After image processing, the macro generated a report with ten quantitative outcome variables (Figure 1I), which included: i) ONL area (mm2); ii) ONL total cells (count); iii) ONL TUNEL+ cells (count); iv) ratio of ONL TUNEL+ cells/ONL area (count/mm2);

v) percentage of ONL TUNEL+ cells/ONL total cells (%); vi) INL area (mm2); vii) INL total cells (count); viii) INL TUNEL+ cells (count); ix) ratio of INL TUNEL+ cells/INL area (count/mm2); x) percentage of INL TUNEL+ cells/INL total cells (%). Additionally, for a qualitative analysis of the algorithm performance, a JPEG image overlay was automatically created and exported for visual assessment of segmented areas and quantitated TUNEL+ cells (Figure 1H).

3.3.3. Algorithm Training, Validation, and Statistical Analysis

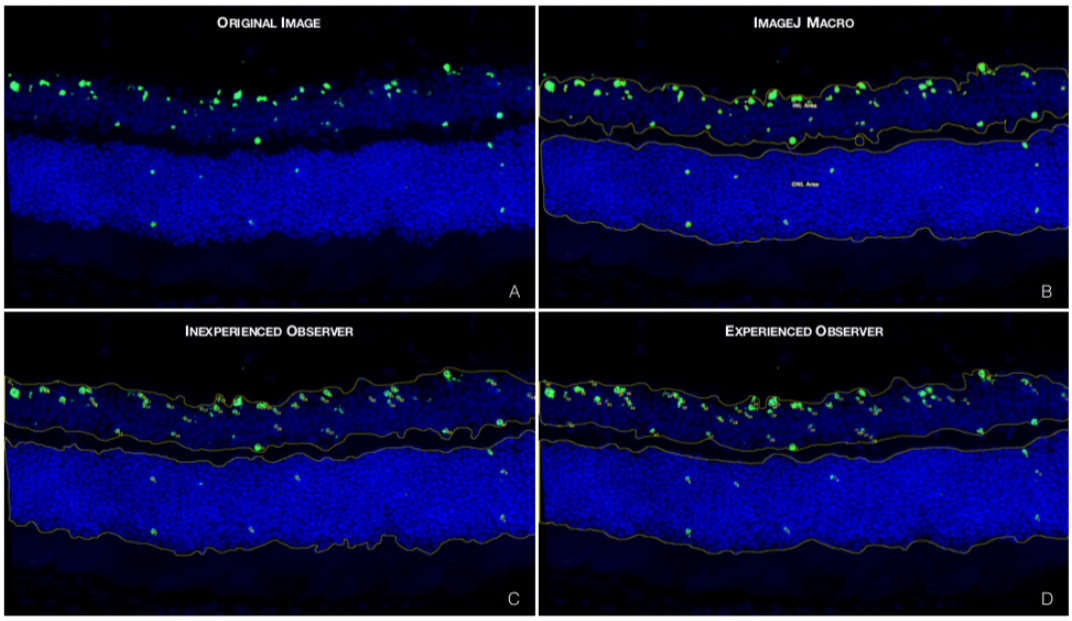
To calibrate and train the macro to experimental observer values, we first defined specific criteria to count $TUNEL^+$ cells, including fluorescent signal co-localization between green and blue channel, nuclei-like size and shape, and if $TUNEL^+$ cells appeared clumped, cell number was counted according to clump size. Most important, we encouraged observers to keep the same criteria for all images. Initially, images were assessed by an independent observer (B.D.), who manually segmented ONL and INL areas, and counted $TUNEL^+$ cells using ImageJ (version 1.48). Algorithm parameters were fine-tuned to these values. Following, two masked, one experienced (P.T.) and one inexperienced observer (B.T.), assessed the image dataset using the same approach. Observers were given duplicate images within the set without their knowledge in order to assess measurement reproducibility in a blind manner. Outcome variables were recorded as well as time required quantitating each image.

Statistical analysis was performed with Prism version 6 (GraphPad Software, CA) and JMP Pro software version 11.2.0 (SAS, NC). Normality was assessed with Shapiro-Wilk test. Statistical significance for differences between groups was determined with Wilcoxon signed-rank test for matched pairs, and one-way ANOVA with Tukey post-hoc correction for multiple comparisons. For correlation analysis, values were fitted to linear regression, and results were expressed as Spearman rho correlation coefficient and R^2 goodness of fit for the model. Intra- and inter-observer agreement analysis was achieved by plotting the mean difference between measurements against their mean using Bland-Altman plot method.[16] Coefficient of variation (COV) was calculated as the difference between measurements over their mean $[(A - B)/(A + B)/2]*100$.[17] Results are expressed

as mean ± standard deviation (SD). A p value of $< .05$ was considered statistically significant.

3.4. Results

All images that matched our selection criteria and included in the dataset were processed by the ImageJ macro, with accurate segmentation of ONL and INL areas and TUNEL$^+$ cell count, as seen in the exported JPEG image overlay (Figure S1). Additionally, all defined variables were quantitated and reported.



Supplementary Figure 1. Image analysis by ImageJ macro, inexperienced, and experienced observer.

A representative TUNEL assay digital image, centered in-frame, focused, with retinal layers clearly identifiable (A). Manual retinal layer segmentation and cell counting from the ImageJ macro (B), inexperienced (C), and experienced observer (D).

3.4.1. ImageJ macro and observers' performance for area and TUNEL+ cells

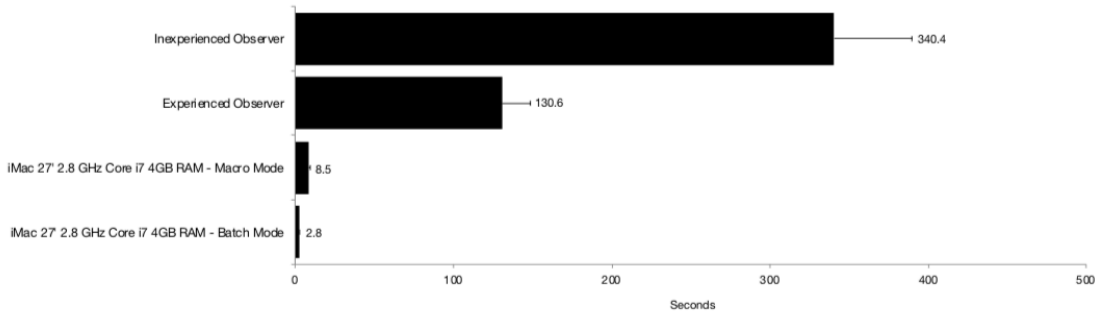
Results from observers and macro measurements can be seen in Table 1. Outer nuclear layer area measured by the automated macro (0.028 mm^2) was similar to the inexperienced observer (0.028 mm^2, $p = .958$) and 10% less than the experienced observer (0.030 mm^2, $p = .019$). No statistically significant differences were observed in $TUNEL^+$ cell counts and ratio of $TUNEL^+$ cells/area between the observers and macro. Inner nuclear layer area measured by the macro (0.016 mm^2) was similar to the inexperienced observer (0.018 mm^2, $p = .286$) and 10% less than the experienced observer (0.018 mm^2, $p = .009$). No statistically significant differences were observed in $TUNEL^+$ cell counts and ratio of $TUNEL^+$ cells/area between the experienced observer and macro. In contrast, significant differences were observed in these variables between the inexperienced observer and macro ($p < .001$).

	ImageJ Macro		Experienced Observer			Inexperienced Observer		
	Mean ± SD	Range	Mean ± SD	Range	*P* value*	Mean ± SD	Range	*P* value*
Outer Nuclear Layer	0.028 ± 0.004	0.020 - 0.037	0.030 ± 0.004	0.022 - 0.040	**0.019**	0.028 ± 0.004	0.021 - 0.037	0.958
Area (mm2)	32.4 ± 19.6	3 - 97	40.5 ± 27.0	3.5 - 138	0.121	25.0 ± 19.0	2 - 89.5	0.194
TUNEL+ Cells (count)	1197 ± 833	108 - 4742	1403 ± 1045	121 - 5938	0.552	915 ± 772	66 - 4019	0.230
Ratio TUNEL+ Cells/Area (count/mm2)								
Inner Nuclear Layer	0.016 ± 0.003	0.011 - 0.025	0.018 ± 0.003	0.012 - 0.027	0.009	0.018 ± 0.003	0.011 - 0.026	0.286
Area (mm2)	30.6 ± 15.3	5 - 74	35.9 ± 19.6	6 - 98	0.204	19.9 ± 12.4	2 - 58.5	< .001
TUNEL+ Cells (count)	1805 ± 797	357 - 3551	1922 ± 986	319 - 4332	0.878	1106 ± 616	97 - 2854	< .001
Ratio TUNEL+ Cells/Area (count/mm2)	2.82 ± 0.04	2.77 - 2.88	130.60 ± 17.96	95 - 196.50	< .0001	340.40 ± 49.07	195.50 - 450.50	< .0001
Total Burden Time per Image**								

$TUNEL^+$, TUNEL positive cells; SD, standard deviation; Range = minimum - maximum value; * = one-way analysis of variance (ANOVA) with post-hoc Tukey honest significant differences (HSD) for comparison between groups. A p value of < .05 was considered statistically significant;** = seconds

Table 1. Outcome variables for observers and ImageJ macro.
Values for the experienced and inexperienced observers represent the average between first and second measurements.

We calculated observers total burden time per image by adding the time needed for ONL and INL segmentation and TUNEL$^+$ cell counting. On average, an experienced observer needed 2.16 minutes per image whereas the inexperienced observer needed more than 5.66 minutes. Macro needed less than 3 seconds per image with batch processing on an iMac 2.8 GHz Core i7 computer, a 60- and 160-fold faster than the experienced and inexperienced observers, respectively (Figure S4).



Supplementary Figure 4. Burden Time per Image Comparison Between Observers and Macro. Total burden time per image represents the required time for ONL and INL segmentation and TUNEL+ cells counting.

Next, ONL and INL TUNEL$^+$ cells values were plotted for each image, observers, and automated macro. TUNEL$^+$ cell values for the macro were located between the lower values of inexperienced observer and the higher counts of the experienced observer, with only few exceptions corresponding to images with significant background noise, due to lack of post-processing in the native image (Figure 2A-B).

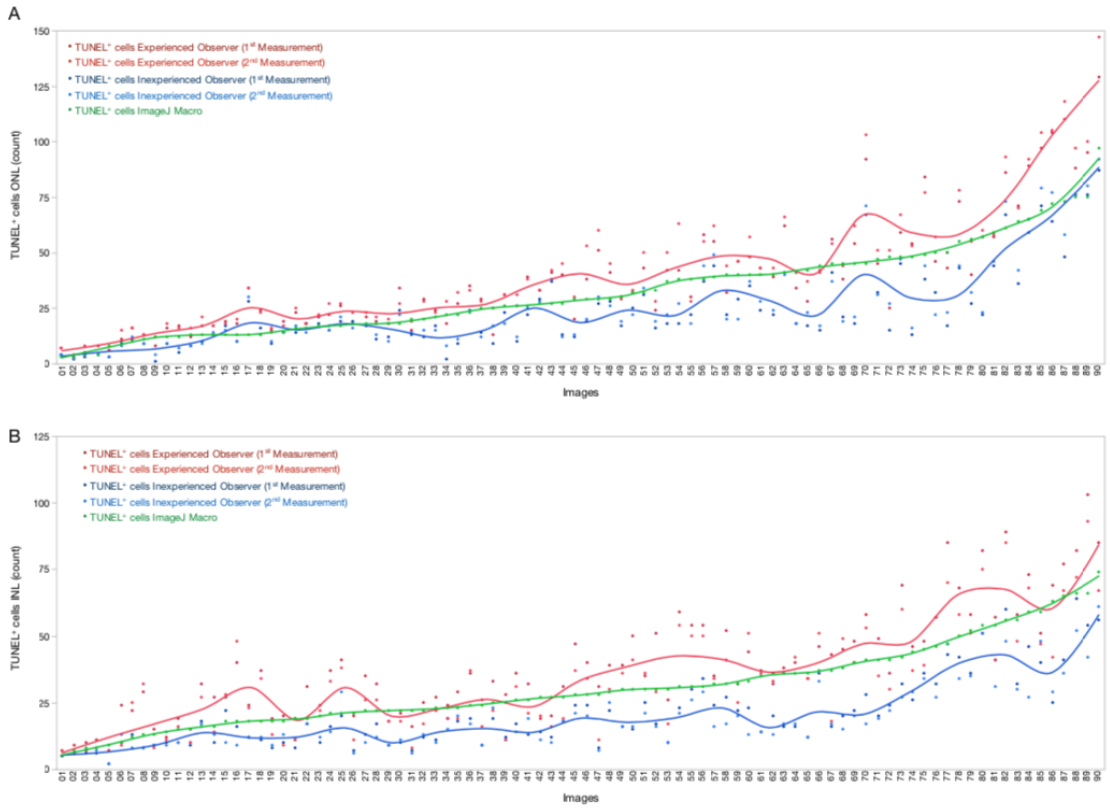


Figure 2. TUNEL positive cell counts for macro and mean values of inexperienced and experienced observers.

Cell counts for ImageJ macro and first and second measurements from observers plotted for every image, for ONL (A) and INL (B) datasets. A colored trendline represents mean variation for each observer. Macro TUNEL+ counts are located between observers' measurements.

3.4.2. Intra-observer Agreement and Correlation

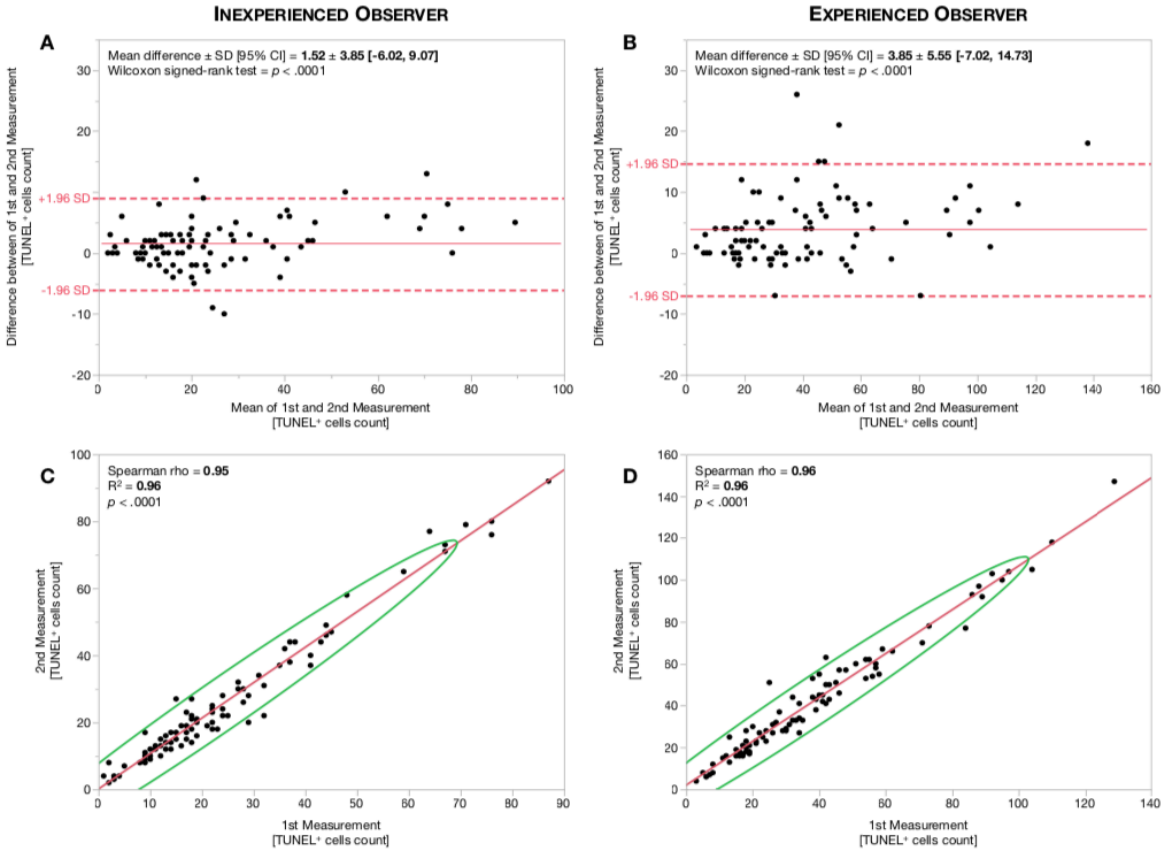
Since repeatability is a crucial element for any measurement method, we first sought to determine intra-observer variability. To answer this question, observers quantitated 180 images within the dataset, twice, in a masked fashion. Values from the first and second measurement were compared for each image. Following, we plotted their differences against the mean using Bland-Altman plots. The coefficient of variation of $TUNEL^+$ cell counts between these two measurements for the ONL dataset was 15.21 ± 19.69% for the inexperienced observer, and 13.09 ± 13.84% for the experienced observer (Table S1 and

Figure S2A-B). For the INL dataset, the COV was 16.41 ± 15.39% for the inexperienced observer, and 25.20 ± 21.82% for the experienced observer (Table S1 and Figure S3A-B). Spearman rho correlation coefficient between first and second measurement from the inexperienced observer was 0.95 in ONL images and 0.92 in INL images (Table S1 and Figure S2C-S3C). This coefficient was 0.96 in ONL images and 0.93 in INL images for the experienced observer (Table S1 and Figure S2D-S3D).

	Intra-Observer Agreement and Correlation		**Inter-Observer Agreement and Correlation**	
	Inexperienced Observer 1st vs. 2nd Measurement	Experienced Observer 1st vs. 2nd Measurement	Inexperienced vs. Experienced Observer	Average of Observers vs. ImageJ Macro
Outer Nuclear Layer	**< .001**	**< .0001**	**< .0001**	0.826
Wilcoxon signed rank test	0.95	0.96	0.90	0.90
Spearman correlation	1.52 ± 3.85	3.85 ± 5.55	15.56 ± 11.51	0.37 ± 9.78
Mean Difference or Bias (count)	15.21 ± 19.69	13.09 ± 13.84	51.11 ± 25.83	23.37 ± 15.97
Coefficient of Variation (%)				
Inner Nuclear Layer	**< .001**	**< .0001**	**< .0001**	**< .001**
Wilcoxon signed rank test	0.92	0.93	0.89	0.86
Spearman correlation	1.76 ± 4.69	6.30 ± 6.47	16.00 ± 10.03	2.72 ± 7.26
Mean Difference or Bias (count)	16.41 ± 15.39	25.20 ± 21.82	56.07 ± 24.03	23.44 ± 18.56
Coefficient of Variation (%)				

A *p* value of < .05 was considered statistically significant

Supplementary Table 1. Intra- and inter-observer agreement and correlation.

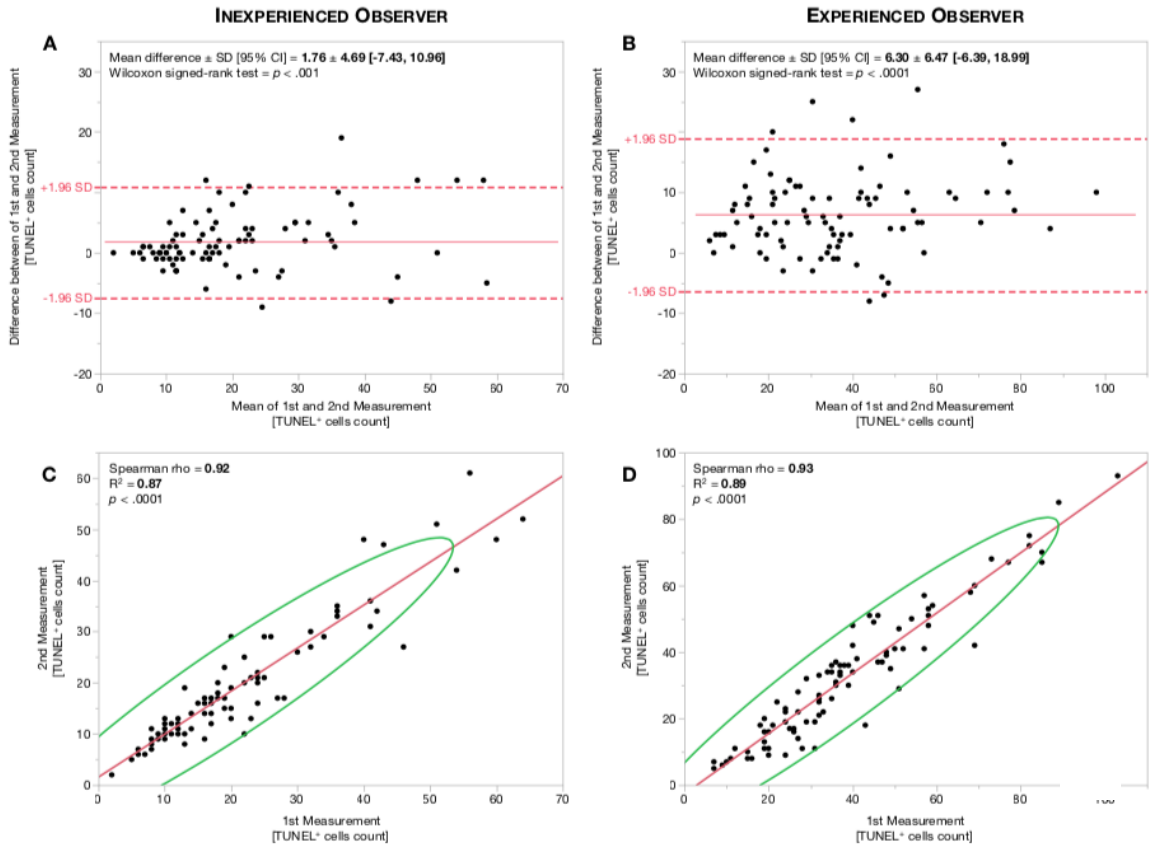


Supplementary Figure 2. Outer Nuclear Layer (ONL) Intra-observer Agreement and Correlation. Bland-Altman plots for agreement analysis between the first and second ONL measurements of the inexperienced (A) and experienced observer (B). Red lines represent the mean difference or bias. Red dashed-lines represent the 95% limit of agreement (±1.96 standard deviation (SD)). Correlation analysis between first and second measurements of inexperienced (C) and experienced observer (D). Red lines represent the fitted linear regression trendline. Green lines represent the 95% confidence interval area.

3.4.3. Inter-observer Agreement and Correlation

To determine the inter-observer variability, we averaged the first and second measurements of each observer and obtained their respective mean TUNEL$^+$ cell count. The coefficient of variation of TUNEL$^+$ cell counts between the two observers was 51.11 ± 25.83% for the ONL dataset (Table S1 and Figure 3A), and 56.07 ± 24.03% for the INL dataset (Table S1 and Figure 4A). Following, we sought to determine whether the ImageJ

macro would agree with the observers. We used the limits of agreement between these two observers to define the range to which the macro TUNEL$^+$ cell counts should fit. Of note, these ranges are estimated for the whole dataset, and should not be extrapolated or predicted as per image. When compared to observers, the COV for the macro was 23.37 ± 15.97% for the ONL (Table S1 and Figure 3B), and 23.44 ± 18.56% for the INL (Table S1 and Figure 4B). These results can be visualized in Bland-Altman plots, as the values for the macro against observers presented a smaller bias value and a more grouped and narrow confidence interval than the existing between observers. In addition, Spearman rho coefficient between observers was 0.90 for the ONL and 0.89 for the INL, and 0.90 for the ONL and 0.86 for the INL between observers and macro (Table S1, Figure 3C-D and Figure 4C-D).



Supplementary Figure 3. Inner Nuclear Layer (INL) Intra-observer Agreement and Correlation. Bland-Altman plots for agreement analysis between the first and second INL measurements of the inexperienced (A) and experienced observer (B). Red lines represent the mean difference or bias.

Red dashed-lines represent the 95% limit of agreement (±1.96 standard deviation (SD)). Correlation analysis between first and second measurements of inexperienced (C) and experienced observer (D). Red lines represent the fitted linear regression trendline. Green lines represent the 95% confidence interval area.

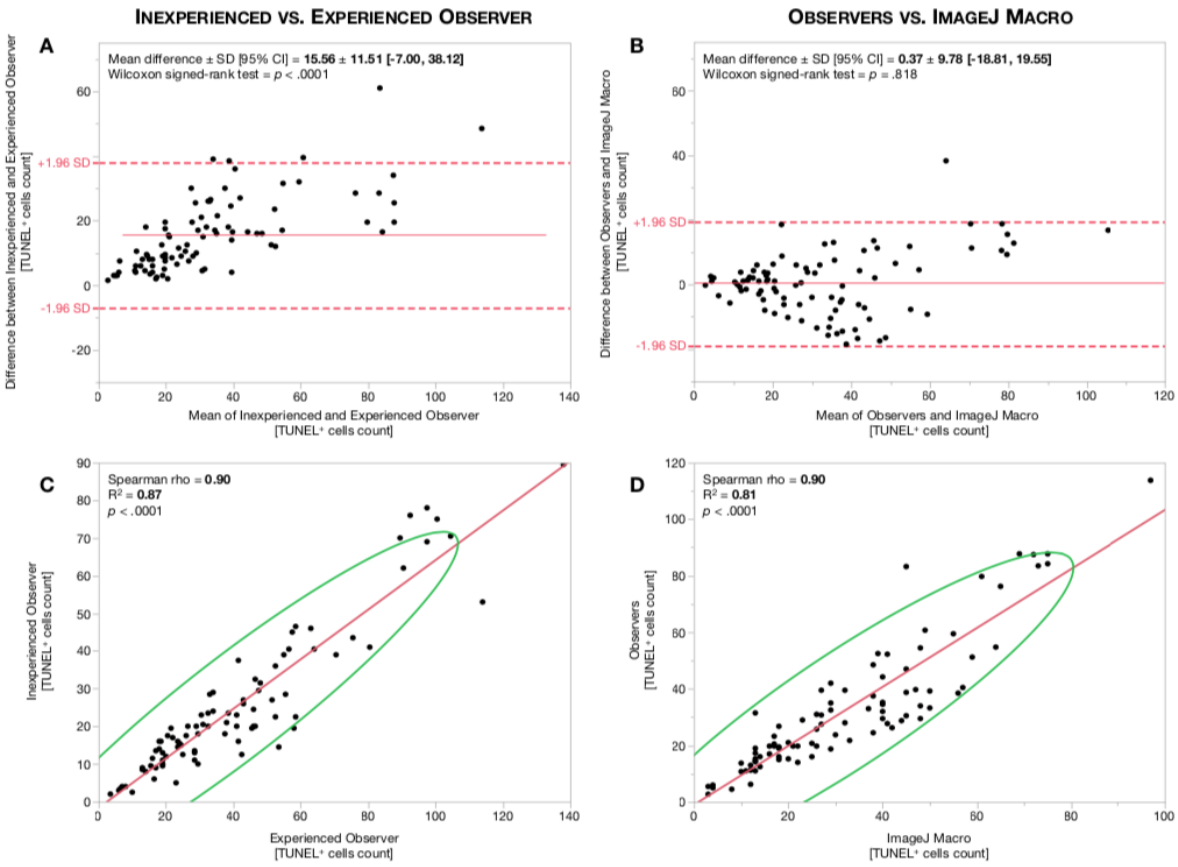


Figure 3. Outer Nuclear Layer (ONL) Inter-observer Agreement and Correlation.
Bland-Altman plots for ONL measurement agreement between observers (A) and between observers and ImageJ macro (B). Red lines represent the mean difference or bias. Red dashed-lines represent the 95% limit of agreement (±1.96 standard deviation (SD)). Correlation analysis between observers (C) and between observers and ImageJ macro (D). Red lines represent the fitted linear regression trendline. Green lines represent the 95% confidence interval area.

3.4.4. Experimental Testing

To confirm its validation and to further test performance, we sought to determine whether the macro was able to detect statistically significant differences, if present, between experimental groups. We randomly selected 30 images from a control group and 30 from

an experimental group and compared results between observers and macro.

In the ONL dataset (Figure 5A), all observers and macro were able to detect a statistically significant difference between the experimental group and their respective control ($p <$ .0001). No statistical differences were observed in control and experimental groups between macro and average all observers. However, there were small but statistically significant differences in the control values obtained by the different observers ($p = .004$), but no statistically significant differences obtain by the different observers in the experimental group ($p = .943$).

In the INL dataset (Figure 5B), all observers and macro were able to detect a statistically significant difference between the two experimental groups ($p < .0001$). No statistical differences were observed for both 10 and 100 nM groups between the experienced observer and macro. There was though a statistically significant difference in the absolute values obtained for the 10 nM group ($p = .018$) and 100 nM group ($p < .0001$) between the two observers.

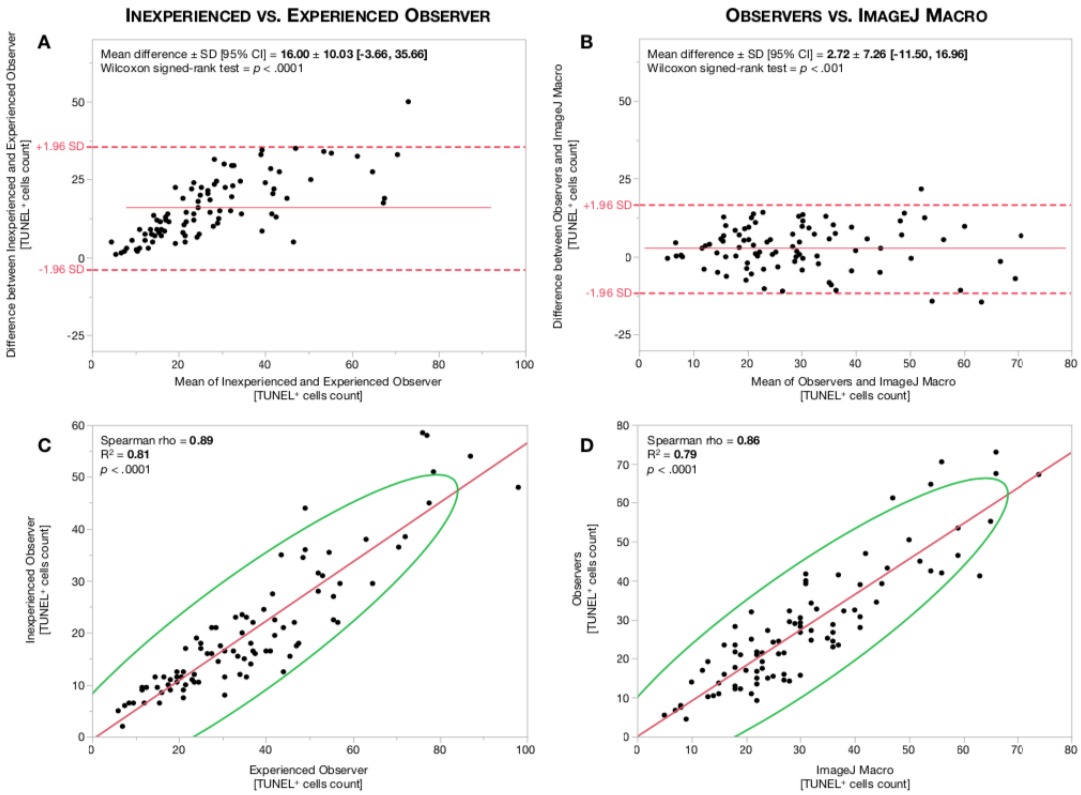
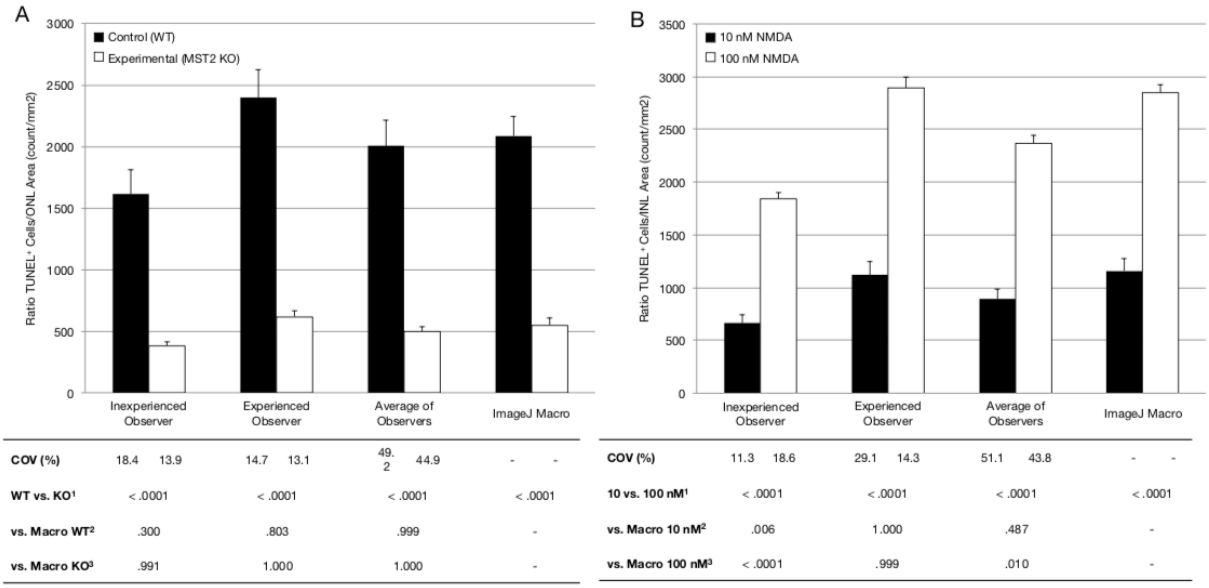


Figure 4. Inner Nuclear Layer (INL) Inter-observer Agreement and Correlation.
Bland-Altman plots for INL measurement agreement between observers (A) and between observers and ImageJ macro (B). Red lines represent the mean difference or bias. Red dashed-lines represent the 95% limit of agreement (±1.96 standard deviation (SD)). Correlation analysis between observers (C) and between observers and ImageJ macro (D). Red lines represent the fitted linear regression trendline. Green lines represent the 95% confidence interval area.



COV (%)	18.4	13.9	14.7	13.1	49.2	44.9	-	-
WT vs. KO[1]	< .0001		< .0001		< .0001		< .0001	
vs. Macro WT[2]	.300		.803		.999		-	
vs. Macro KO[3]	.991		1.000		1.000		-	

COV (%)	11.3	18.6	29.1	14.3	51.1	43.8	-	-
10 vs. 100 nM[1]	< .0001		< .0001		< .0001		< .0001	
vs. Macro 10 nM[2]	.006		1.000		.487		-	
vs. Macro 100 nM[3]	< .0001		.999		.010		-	

Figure 5. Experimental Testing of ImageJ Macro
From the ONL dataset, 30 images from a wild-type (WT) control group and 30 from an experimental (MST2 knock-out) group were selected (A). From the INL dataset, 30 images from a 10 nM NMDA group and 30 from a 100 nM NMDA group were selected (B). Columns represent the mean values and standard error of the mean for the ratio of TUNEL+ cells/area per observer. [1]Wilcoxon signed rank test between groups within the same observer. [2]One-way ANOVA with Tukey post-hoc correction for comparison between all control groups or 10 nM against the macro. [3]One-way ANOVA with Tukey post-hoc correction for comparison between all experimental groups or 100 nM against the macro.

3.5. Discussion

In this study, we developed an ImageJ macro that was able to segment ONL and INL retinal layers and count TUNEL$^+$ cells accurately. The reported parameters were able to fit within the established limits of agreement and were highly correlated with the observers' measurements. In addition, the macro was able to detect statistically significant differences between experimental groups. Collectively, these results validate the macro as a measurement tool for automated cell death quantitation in the retina. We speculate that the advent of this tool can aid to achieve accurate, repeatable, fast, and unbiased measurements.

A moderate degree of variability was observed between repeated measurements in both inexperienced and experienced observers, as expected. Collectively, both observers presented good reproducibility evidenced by bias value and range of limits of agreement. Bland-Altman plots for the inexperienced observer displayed a lower bias value and narrower limits of agreement as compared to the experienced observer, most likely due to the time invested in the analysis and higher TUNEL$^+$ cell counts by the experienced observer. Still, both observers were able to detect statistically significant differences among experimental groups, and the magnitude of this difference was higher in the experienced observer. In contrast, Spearman correlation coefficient for both observers showed similar values between groups. As Bland and Altman previously described, statistical hypothesis tests and correlation coefficients are not suitable statistical tests to compare the agreement of two measurement methods. The former only determines whether the means of groups are equal or not. The latter measures the linear dependence between to variables, hence a paralleled change in the measurement magnitude will not affect the correlation coefficient, but will certainly affect agreement.[16,18] We suggest that for the purpose of validating new measurement techniques, Bland-Altman plots and coefficient of variation should be calculated, and a limit of agreement should be determined beforehand.

The ImageJ macro was able to detect layers and count TUNEL$^+$ cells within the limits of

agreement determined by both observers. The coefficients of variation between macro and average of observers were approximately half the value of those between observers, for both ONL and INL dataset. These values are supported by the narrower limits of agreement between macro and average of observers than between both observers, as seen in Bland-Altman plots. These premises argue in favor of the high level of accuracy of this measurement tool.[16] On a side note, $TUNEL^+$ cell count variability was increased in INL images with more $TUNEL^+$ cells, probably owing to the presence of cells of different sizes[19] and staining patterns. Nonetheless, since there is no absolute reference value in $TUNEL^+$ cell counting, we believe it is crucial to follow consistent criteria when counting cells, in order to detect consistent differences between groups.

The burden time to analyze these images is significant. On average, an experienced observer needed over two minutes per image, whereas an inexperienced observer required over five minutes. In contrast, the macro required less than three seconds for each image analysis, when executed in batch mode. This advantage opens up the possibility to assess a larger area of retina. We calculated that the time required to manually count two images per retina would still be higher than analyzing the entire retina with 8-9 images using ImageJ macro, even considering increased acquisition time. The ability to examine larger retina area in the same amount of time will result in higher experimental objectivity.

While estimating expected sample size for each experimental group, the variability both of the biological model and the measurement tool should be taken into account. If the measurement tool used to quantitate outcome variables has significant variability, it increases the overall variability as bias and limits of agreement are also subject to sampling variation.[20] In essence, higher variability reduces the ability to detect statistical significance, increasing required sample size. Therefore, using an automated measurement tool can reduce bias and sample size.

3.5.1 Image acquisition, post processing recommendations and limitations

To challenge the macro versatility, we included images taken by multiple observers using different acquisition parameters. However, we recommend a standardized acquisition and

post processing approach, customized for each researcher and equipment used, and to test the algorithm prior to experimental testing. Images should ideally be centered, with ONL and INL layers not merged, without shadows or uneven focus, to ensure an appropriate macro performance. Image background noise should be reduced as much as possible.

Still, several limitations should be considered. First, macro results were compared against two observers. However, these observers performed with good reproducibility. Second, the macro assumes the largest nuclei area in the image to be the ONL area. This should be considered when assessing images from experimental neurodegenerative models with significant ONL/INL thinning. Finally, digital images should be acquired with a 20x/0.8 objective, as this is the chosen working magnification of the algorithm, for the aforementioned reasons. We are currently working on new algorithm versions to broaden the type of images suitable for automated analysis.

In summary, the developed and validated ImageJ macro can be an accurate and precise quantitative tool for retina researchers, to achieve repeatable, unbiased, fast, and above all, accurate cell death quantitation. We believe this standardized measurement tool could be advantageous to compare results across different research groups. The macro will be freely available to all as open source.

3.6. Acknowledgements

Authors contributions: D.E.M. and D.G.V. designed research. D.E.M. coded the macro, performed statistical analysis, and drafted the manuscript. P.T., B.T., and B.D. segmented retinal layers and counted TUNEL+ cells from the digital image dataset and contributed to the manuscript. P.T., H.M., and K.K. performed TUNEL assay image acquisition. H.L., J.W.M, and D.G.V. reviewed the data and made substantial intellectual contributions to the research design and manuscript.

Please visit http://imagej.net/TUNEL_Cell_Counter to download the macro and original

image dataset. We added a detailed tutorial, setup instructions, and suggestions to improve performance. For further information, please contact Demetrios G. Vavvas at Demetrios_Vavvas@meei.harvard.edu. Please cite this article as follows: Maidana DE, Tsoka P, Tian B, Dib B, Matsumoto H, Kataoka K, Lin H, Miller JW, Vavvas DG (2015) A Novel ImageJ Macro for Automated Cell Death Quantitation in the Retina. Invest Ophthalmol Vis Sci.

3.7. Funding Support

This work was supported by Bayer® Healthcare GOAP Ophthalmology Award to Daniel E. Maidana. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

3.8. Financial Disclosures

The authors have no relevant financial relationships or interests to disclose.

3.9. Chapter Summary in Spanish

3.9.1. Introducción

En los últimos años, la investigación en el campo de las enfermedades degenerativas de la retina se ha centrado en comprender la muerte celular y las formas de interferir con ella. Varios modelos animales experimentales, como retinopatía diabética, retinitis pigmentosa, desprendimiento de retina, o glaucoma, han revelado que la muerte celular en la retina puede ocurrir a través de diferentes modalidades, incluidas la apoptosis y la necroptosis, entre otras. Estas vias abarcan múltiples inductores y a su vez muestran diferentes características morfológicas.1-5 De las muchas herramientas para caracterizar y cuantificar la muerte celular, el ensayo de marcado de final de corte de deoxouridina trifosfato (dUTP) terminal deoxinucleotidil transferasa (terminal deoxynucleotidyl transferase dUTP nick end labeling, TUNEL) es probablemente el método más extendido. Este ensayo detecta ácido desoxirribonucleico (ADN) fragmentado o mellado por medio de una enzima desoxinucleotidil transferasa terminal, que incorpora dUTP marcados con fluorescencia a las regiones dañadas de ácido nucleico. Descrito por primera vez por Gavrieli et al6 y, aunque no es completamente específico, se ha demostrado que este ensayo es el método más sensible para detectar apoptosis.7 Sin embargo, dado que este ensayo detecta fragmentos de ADN independientemente de la vía de muerte celular inducida, no puede distinguir entre apoptósis o formas necroptóticas de muerte celular programada.3,8-10 A pesar de sus limitaciones, el ensayo TUNEL sigue siendo el método más extendido utilizado para detectar cualquier forma de muerte celular programada hasta la fecha.

Para cuantificar los resultados del ensayo TUNEL en secciones de tejido, normalmente se realiza un conteo manual de células TUNEL positivas (TUNEL+), idealmente por dos observadores enmascarados.11 Los observadores cuentan células TUNEL+ y miden el área retiniana correspondiente, y los resultados se expresan como células TUNEL+ /área o TUNEL+ células/células totales. Evidentemente, este proceso lleva mucho tiempo, es propenso a errores de medición, y no es completamente reproducible. Por lo tanto, un

enfoque automatizado es de interés para abordar estas dificultades. Actualmente, no existe método de cuantificación de muerte celular automatizado disponible para la plataforma ImageJ, capaz de segmentar capas retinianas y contar células TUNEL+ individuales en la retina sin variabilidad.

El propósito de este estudio fue introducir una nueva macro ImageJ para la detección automática de muerte celular en secciones transversales de la retina. Además, estudiamos las diferencias en el conteo entre un observador experto y otro inexperto. Dado que la introducción de un nuevo método de medición requiere validación con el estándar o método comúnmente aceptado, validamos esta macro comparando su desempeño frente a dos observadores. Descubrimos que la macro ImageJ puede lograr una cuantificación de muerte celular retiniana rápida y precisa.

3.9.2. Resultados

Todas las imágenes que cumplieron con nuestros criterios de selección fueron procesadas por la macro ImageJ, con una segmentación precisa de las áreas ONL e INL y posterior contaje de células TUNEL+, como se ve en la Figura S1B. Además, todas las variables definidas fueron cuantificadas y reportadas.

Los resultados de los observadores y las mediciones de la macro se pueden ver en la Tabla 1. El área de la capa nuclear externa medida por la macro automatizada (0.028 mm 2) fue similar a la del observador inexperto (0.028 mm^2, $p = 0.958$) y 10% menos que la medida por un observador experto (0.030 mm 2 , p = 0.019). No se observaron diferencias estadísticamente significativas en los recuentos de células TUNEL+ y la relación de células TUNEL+ /área entre los observadores y la macro. El área de la capa nuclear interna medida por la macro (0.016 mm^2) fue similar al observador inexperto (0.018 mm^2, p = 0.286) y 10% menos que el observador experto (0.018 mm^2, $p = 0.009$). No se observaron diferencias estadísticamente significativas en los recuentos de células TUNEL+ y la proporción de células TUNEL+/área entre el observador experto y la macro. En contraste, se observaron diferencias significativas en estas variables entre el observador inexperto y la macro ($p < 0.001$).

Calculamos el tiempo de trabajo total de los observadores por imagen sumando el tiempo necesario para la segmentación de la ONL e INL y el recuento de células TUNEL+. En promedio, un observador experto necesitó 2.16 minutos por imagen, mientras que el observador inexperto necesitó más de 5.66 minutos. La ImageJ macro requirió menos de 3 segundos por imagen en modo procesamiento por lotes en un equipo i7 iMac 2,8 GHz Core. Esto representa 60 y 160 veces más rápido que los observadores experto e inexperto, respectivamente (Figure S4).

A continuación, se trazaron los valores de las células TUNEL + en la ONL e INL para cada imagen, observadores y macro automatizado. Los valores de células TUNEL + para la macro se ubicaron entre los valores más bajos del observador inexperto y los recuentos más altos del observador experto, con solo algunas excepciones correspondientes a imágenes con significativo ruido, debido a la falta de pos-procesamiento en la imagen nativa (Figure 2A-B).

Dado que la repetitividad es un elemento crucial para cualquier método de medición, primero determinamos la variabilidad intra-observador. Para este fin, los observadores cuantificaron 180 imágenes, dos veces, de forma enmascarada. Los valores de la primera y segunda medición se compararon para cada imagen. A continuación, trazamos sus diferencias con respecto a la media usando las gráficas de Bland-Altman. El coeficiente de variación (COV) de los recuentos de células TUNEL+ entre estas dos mediciones para la ONL fue 15.21 ± 19.69% para el observador inexperto, y 13.0 9 ± 13.84% para el observador experimentado (Tabla S1 y Figura S2A-B). Para la INL, el COV fue 16.41 ± 15.39% para el observador inexperto, y 25.20 ± 21.82% para el observador experimentado (Tabla S1 y Figura S3A-B). El coeficiente de correlación de Spearman *rho* entre la primera y la segunda medición del observador inexperto fue de 0.95 para la ONL y 0.92 para la INL (Tabla S1 y Figura S2C-S3C). Este coeficiente fue 0.96 para la ONL y 0.93 para la INL para el observador experimentado (Tabla S1 y Figura S2D-S3D).

Para determinar la variabilidad entre observadores, promediamos la primera y segunda medición de cada observador y obtuvimos su respectivo recuento medio de células TUNEL+. El coeficiente de variación de los recuentos de células TUNEL+ entre

los dos observadores fue 51.11 ± 25.83% para la ONL (Tabla S1 y Figura 3A), y 56.07 ± 24.03% para la INL (Tabla S1 y Figura 4A). A continuación, buscamos determinar si la performance de la macro ImageJ se encuentra en acuerdo con los observadores. Utilizamos los límites de acuerdo entre estos dos observadores para definir el rango al que deben ajustarse los recuentos de células macro TUNEL+ . Cabe destacar que estos rangos se estiman para todas las imágenes de forma global. En relación con los observadores, el COV para la macro fue 23.37 ± 15.97% para la ONL (Tabla S1 y Figura 3B), y 23.44 ± 18.56% para la INL (Tabla S1 y Figura 4B). Estos resultados pueden visualizarse en gráficos de Bland-Altman, ya que los valores para la macro contra los observadores presentaron un valor de sesgo menor y un intervalo de confianza más agrupado y estrecho que el existente entre los observadores. Además, el coeficiente de Spearman rho entre los observadores fue 0.90 para la ONL y 0.89 para la INL; y 0.90 para la ONL y 0.86 para la INL entre observadores y macro (Tabla S1, Figura 3C-D y Figura 4C-D).

Para confirmar su validación y para mejorar el rendimiento de la prueba, buscamos determinar si la macro es capaz de detectar diferencias estadísticamente significativas, si están presentes, entre los grupos experimentales. Seleccionamos al azar 30 imágenes de un grupo de control y 30 de un grupo experimental, y comparamos los resultados entre observadores y macro.

Para evaluar la performance en la ONL (Figura 5A), se utilizaron imágenes de un desprendimiento de retina experimental. Todos los observadores y macro pudieron detectar una diferencia estadísticamente significativa entre el grupo experimental y su control respectivo ($p < 0.0001$). No se observaron diferencias estadísticas en los grupos de control y experimentales entre todos los observadores macro y promedio. Sin embargo, hubo diferencias pequeñas, pero estadísticamente significativas en los valores de control obtenidos por los diferentes observadores ($p = 0.004$). No se encontraron diferencias estadísticamente significativas entre los diferentes observadores en el grupo experimental ($p = 0.943$).

Para evaluar la performance en la INL (Figura 5B), se utilizaron imágenes de un modelo

de excito-toxicidad mediante la inyección intravítrea de N-metil-D-aspartato (NMDA). Todos los observadores y macro pudieron detectar una diferencia estadísticamente significativa entre los dos grupos experimentales ($p < 0.0001$). No se observaron diferencias estadísticas para los grupos de 10 y 100 nM entre el observador experimentado y la macro. Sin embargo, hubo una diferencia estadísticamente significativa en los valores absolutos obtenidos para el grupo 10 nM ($p = 0.018$) y el grupo 100 nM ($p < 0.0001$) entre los dos observadores.

3.9.3. Discusión

En este trabajo, desarrollamos una macro ImageJ que fue capaz de segmentar capas retinianas ONL e INL y contar células TUNEL+ con precisión. Los resultados se situaron dentro de los límites de acuerdo establecidos y se correlacionaron positivamente con las mediciones de los observadores. Además, la macro pudo detectar diferencias estadísticamente significativas entre los grupos experimentales. En conjunto, estos resultados validan la macro como una herramienta de medición para la cuantificación automatizada de la muerte celular en la retina. Especulamos que el uso de esta herramienta pueda ayudar a lograr mediciones precisas, repetibles y rápidas.

Se observó un grado moderado de variabilidad entre mediciones repetidas en observadores inexpertos y experimentados, como se esperaba. Colectivamente, ambos observadores presentaron buena reproducibilidad, dada por el valor de sesgo y el rango de límites de acuerdo. Las gráficas de Bland-Altman para el observador inexperto mostraron un valor de sesgo más bajo y límites de acuerdo más estrechos en comparación con el observador experimentado, probablemente debido al tiempo invertido en el análisis y a los mayores recuentos de células TUNEL + por parte del observador experimentado. Aún así, ambos observadores pudieron detectar diferencias estadísticamente significativas entre los grupos experimentales, y la magnitud de esta diferencia fue mayor en el observador experimentado. En contraste, el coeficiente de correlación de Spearman para ambos observadores mostró valores similares entre los grupos. Como Bland y Altman describieron anteriormente, las pruebas de hipótesis estadísticas y los coeficientes de correlación no son pruebas estadísticas adecuadas para

comparar la concordancia de dos métodos de medición. El primero solo determina si los medios de los grupos son iguales o no. Este último mide la dependencia lineal entre las variables, por lo tanto, un cambio paralelo en la magnitud de la medición no afectará el coeficiente de correlación, pero ciertamente afectará la concordancia.[16,18] Sugerimos con el propósito de validar nuevas técnicas de medición, el calculo y visualización mediante gráficas de Bland-Altman y el coeficiente de variación, y la determinación *a priori* de un límite de acuerdo.

La macro ImageJ pudo detectar capas y contar células TUNEL+ dentro de los límites de acuerdo determinados por ambos observadores. Los coeficientes de variación entre macro y el promedio de observadores fueron aproximadamente la mitad del valor de aquellos entre observadores, tanto para la ONL como para la INL. Estos valores están respaldados por los límites de acuerdo más estrechos entre macro y el promedio de observadores que entre ambos observadores, como se ve en las gráficas de Bland-Altman. Estas premisas argumentan a favor del alto nivel de precisión de esta herramienta de medición.[16] Al margen, la variabilidad del recuento de células TUNEL+ aumentó en la INL, con más células TUNEL+, probablemente debido a la presencia de células de diferente tamaño[19] y patrones de tinción. No obstante, dado que no hay un valor de referencia absoluto en el recuento de células TUNEL+, creemos que es crucial seguir criterios rigurosos al contar las células, para detectar diferencias consistentes entre los grupos.

El tiempo de trabajo para analizar estas imágenes es significativo. En promedio, un observador experimentado necesita más de dos minutos por imagen, mientras que un observador sin experiencia requería más de cinco minutos. En contraste, la macro requirió menos de tres segundos para cada análisis de imagen, cuando se ejecutó en modo por lotes. Esta ventaja abre la posibilidad de evaluar un área retiniana más extensa. Calculamos que el tiempo requerido para contar manualmente dos imágenes por retina aún sería mayor que el análisis de la retina completa con 8-9 imágenes usando ImageJ macro, incluso considerando un mayor tiempo de adquisición. La capacidad de examinar un área más extensa en la misma cantidad de tiempo dará como resultado una mayor objetividad experimental.

Mientras se estima el tamaño de muestra esperado para cada grupo experimental, se debe tener en cuenta la variabilidad tanto del modelo biológico como de la herramienta de medición. Si la herramienta de medición utilizada para cuantificar las variables de resultado tiene una variabilidad significativa, aumenta la variabilidad general ya que el sesgo y los límites de acuerdo también están sujetos a la variación del muestreo.[20] Es decir, una mayor variabilidad en la medición reduce la capacidad de detectar diferencias estadísticamente significativas, aumentando el tamaño de muestra requerido. Por lo tanto, el uso de una herramienta de medición automatizada puede reducir el sesgo y el tamaño de la muestra.

Para desafiar la macro versatilidad, incluimos imágenes tomadas por múltiples observadores usando diferentes parámetros de adquisición. Sin embargo, recomendamos un enfoque estandarizado de adquisición y pos-procesamiento, personalizado para cada investigador y equipo utilizado, y evaluar el algoritmo antes del análisis experimental. Idealmente, las imágenes deben estar centradas, con capas ONL e INL no fusionadas, sin sombras o enfoque irregular, para garantizar un rendimiento adecuado de la macro. El ruido de fondo de la imagen debe reducirse tanto como sea posible.

Aún así, se deben considerar varias limitaciones. Primero, los resultados macro se compararon con dos observadores. Sin embargo, estos observadores realizaron con buena reproducibilidad. En segundo lugar, la macro supone que el área mas grande de la imagen es el área de la ONL. Esto debe tenerse en cuenta al evaluar imágenes de modelos neurodegenerativos experimentales con adelgazamiento significativo de ONL/INL. Finalmente, las imágenes digitales deben adquirirse con un objetivo 20x/0.8, ya que esta es la ampliación elegida del algoritmo, por las razones antes mencionadas. Actualmente estamos trabajando en nuevas versiones de algoritmos para ampliar el tipo de imágenes para el análisis automatizado.

En resumen, la macro ImageJ desarrollada y validada puede ser una herramienta cuantitativa precisa para la investigación, a fin de lograr una cuantificación de muerte celular precisa, repetible, rápida y sobre todo confiable. Creemos que esta herramienta de medición estandarizada podría ser ventajosa para comparar resultados entre diferentes

grupos de investigación. La macro se encuentra disponible gratuitamente para su descarga y uso.

Chapter IV

Photoreceptor Cell Death: Automated Retinal Layer Thickness Measurement

ThicknessTool: Automated ImageJ Retinal Layer Thickness and Profile in Digital Images.

Daniel E. Maidana, Shoji Notomi, Takashi Ueta, Tianna Zhou, Danica Joseph, Cassandra Kosmidou, Josep Maria Caminal Mitjana, Joan W. Miller, and Demetrios G. Vavvas

Paper accepted for publication at *Scientific Reports*

Chapter IV. Automated Retinal Layer Thickness Measurement

4.1. Abstract

Purpose: To develop an automated layer thickness measurement tool for the ImageJ platform, to quantitate nuclear layers following the retina contour.

Materials and Methods: We developed the ThicknessTool (TT), an automated thickness measurement plugin for the ImageJ platform. To calibrate TT, we created a calibration dataset of mock images with increasing thickness masks and different rotations. Following, we created a training dataset and performed an agreement analysis of thickness measurements between TT and two masked manual observers. Finally, we tested the performance of TT measurements in a validation dataset of retinal detachment images.

Results: In the calibration dataset, there were no differences in layer thickness between measured and known thickness masks, with an overall coefficient of variation of 0.00%. Training dataset measurements of immunofluorescence retina nuclear layers disclosed no significant differences between TT and any observer's average outer nuclear layer (ONL) ($p = 0.998$), inner nuclear layer (INL) ($p = 0.807$), and ONL/INL ratio ($p = 0.944$) measurements. Agreement analysis showed that bias between TT vs. observers' mean was lower than between any observers' mean against each other in the ONL (0.77 ± 0.34 µm vs 3.25 ± 0.33 µm) and INL (1.59 ± 0.28 µm vs 2.82 ± 0.36 µm). Validation dataset showed that TT can detect significant and true ONL thinning ($p = 0.006$), more sensitive than manual measurement capabilities ($p = 0.069$).

Conclusions: ThicknessTool can measure thickness in a fast, accurate, and precise manner with multi-platform capabilities. In addition, the TT can be customized to user preferences and is freely available to download.

4.2. Introduction

Progressive photoreceptor cell death is a significant culprit in retinal degenerative diseases. Inasmuch, quantitation of this cell loss has been addressed by a myriad of approaches, predominantly in experimental models of retinal diseases, such as retinal detachment.[1,2] Among these methods, photoreceptor cell death assays[3,4], outer nuclear layer cell counting[5], and outer nuclear layer thickness[6,7], have been used in animal models to quantitate photoreceptor degeneration. Given the technical and time constraints related to cell death assays and manual counting; to date, outer nuclear layer (ONL) thickness has been widely used as a practical proxy method to estimate the depth of photoreceptor cell death.[6–8]

The utility of retina ONL thickness quantitation to infer photoreceptor degeneration relies on fast caliper measurements. However, the advantage of this approach can be compromised as calipers are manually drawn and measured by an observer. In addition, given the radius of curvature of the retina, such calipers should be perpendicular to the layer contour. Moreover, given the high magnification of microscope images, a small area of interest is often analyzed to expedite the analysis, which can lead to bias. Furthermore, given the new imaging modalities that appear, a versatile method that can adapt to multiple imaging modalities is ideal. Therefore, a tool that can adapt to layer architecture and contour to measure thickness in a broad segment, in multiple layers, in either single images or large tiles, and in multiple platforms, is very compelling. Despite pioneering work has been done in this area[9], to the best of our knowledge, there is currently no freely available tool for the ImageJ platform to automatedly quantitate multiple layer thicknesses in large images and in different imaging modalities.

The purpose of this work was to develop an automated retinal layer thickness measurement tool for the ImageJ platform, which can quantitate nuclear layers following the retina contour, with calipers as close to 1-pixel to each other. For this purpose, we developed the ThicknessTool (TT) and validated its accuracy by objective calibration and agreement analysis with two masked observers. We found that this measurement tool can

provide accurate and precise thickness measurements. In addition, TT can process images from multiple imaging modalities.

4.3. Materials and Methods

All animals used in experiments and breeding adhered to the statement of the Association for Research in Vision and Ophthalmology (ARVO). Animal protocols were reviewed and approved by the Animal Care Committee of the Massachusetts Eye and Ear Infirmary. To validate the ThicknessTool in experimental model, a retinal detachment was induced in eight-week-old C57BL/6J mice, as previously described.[10] C57BL/6J mice were purchased from The Jackson Laboratories and maintained in a standard 12-hour light/dark cycle.

4.3.1. Digital Image Dataset

Calibration Dataset. We created an image dataset of fifty 8-bit mock images of 1344 x 1024 pixels containing masks of increasing thickness, from 10 to 500-pixel thickness with 10-pixel increments. Following, we created a second dataset of images with a known mask area of 1344 x 200 pixels with 10º rotation intervals, plus 45º and 135º rotations for a total of 21 images. Finally, to emulate layer thinning and thickening, we created a single mock image containing a mask with abrupt thickness changes, with segments of 200, 150, and 250 pixels. All images were stored in TIFF format, and the measured thickness was compared to the known standard.

Calibration Criteria. We defined the overall mean thickness limit of agreement to 1 pixel, and the following criteria for the ThicknessTool script calibration: i) overall mean thickness equal to 200 ± 1 pixel; ii) index of dispersion equal to 0, as the given area is a rectangle and all calipers should be parallel to each other with equal thickness; iii) single caliper measurement at 0º, 45º, 90º, 135º, and 180º rotations equal to 200 pixels; and iv) single caliper measurement equal to 200 pixels ± 2 pixels at the remaining rotations. This difference corresponds to the aliasing effect on area edges at rotations other than the specified on iii). The occurrence of uneven jagged edges requires a broader margin of

error (2 pixels) than images with straight (0º, 90º, 180º) or even (45º and 135º) edges. For comparisons, the measured mean was tested against a hypothesized mean of 200 pixels. An equivalence test with a two one-sided tests (TOST) approach was used to assess minimum and maximum single caliper measurements, and the threshold difference considered equivalent to no difference was defined as 1 for images with even edges and 2 for images with uneven edges.

Training Dataset. Immunofluorescence digital images of 4′,6-diamidino-2-phenylindole (DAPI)-stained retinal cross-sections were obtained from the Angiogenesis Laboratory fluorescence microscopy database. We selected images from a murine retinal detachment model in order to evaluate TT performance with different ONL thicknesses and various retinal distortions. A total of 64 randomly selected images were included in this dataset and stored in TIFF format for further processing. Training dataset images were measured by an inexperienced observer who never performed the given task before, and by an experienced observer who had performed these measurements before. Images of murine retina displaying the outer (ONL) and inner nuclear layer (INL) were masked and given to observers. Each image was measured twice in a masked manner. Observers measured each layer with six individual calipers, placed across the retinal layer at their discretion. Single calipers were used to calculate mean, minimum, and maximum layer thickness.

Validation Dataset. We randomly selected 16 images from 8 eyes of an experimental murine retinal detachment model. The layer thickness was manually assessed by an observer, who measured ONL and INL in the detached and attached retina, with six individual calipers for each layer.

4.3.2. Thickness Script

We developed a script for the ImageJ platform (version 1.52p, http://imagej.nih. gov/ij/; provided in the public domain by the National Institutes of Health, Bethesda, MD, USA. To facilitate the use of the script, we designed a graphical user interface dialog to tailor performance parameters to users' preferences, including image scale, number of callipers, skeleton width, among others (Supplementary Figure 1).

Briefly, retinal layers were segmented using a previously validated image processing algorithm (Figure 1A-D).[4] From segmented layers, a skeletonized mask was created with AnalyzeSkeleton function, as previously described by Arganda-Carreras and coauthors.[11] Following, thickness callipers were constructed as follows (Figure 1D): i) The segmented mask skeleton was fragmented in segments by user-defined n number of callipers; ii) Centroids from these segmented skeletons were isolated as single points and defined by the coordinates (X0, Y0); iii) The skeleton segment angle was calculated by an user-defined grid which evaluated the angle of the segmented mask skeleton in a given area; iv) From each centroid (X0, Y0), a positive final vector and a negative initial vector were created perpendicularly to the segmented skeleton. The initial and final vectors were composed by a corresponding horizontal component vector (Xstep), and a vertical component vector (Ystep); v) The Xstep and Ystep components were created for the initial and final resolved vector at a 1-pixel distance from the centroid in horizontal and vertical directions, respectively. The preliminary initial (Xi, Yi) and final (Xf, Yf) coordinates were created from the centroid as described in Equation 1; vi) Pixel values were evaluated in the retinal layer mask at initial (Xi, Yi) and final (Xf, Yf) coordinates. If the pixel value returned zero, this was interpreted as the retinal layer mask, and the Xstep and Ystep components were increased by 1-pixel, to analyse the contiguous pixel; vii) This process was iterated until the pixel value was > 0, and this point was assumed as the initial (Xi, Yi) and final (Xf, Yf) vector endpoint; viii) Thickness was calculated as the rounded Euclidean distance between initial (Xi, Yi) and final (Xf, Yf), as seen in Equation 2.

(1)
$$X_i = round\left(\sin\left(\alpha \times \frac{\pi}{180}\right) \times X_{step} + X_0\right)$$

$$Y_i = round\left(Y_0 - \cos\left(\alpha \times \frac{\pi}{180}\right) \times Y_{step}\right)$$

$$X_f = round\left(X_0 - \sin\left(\alpha \times \frac{\pi}{180}\right) \times (X_{step} - \sqrt{1 + {Y_{step}}^2})\right)$$

$$Y_f = round\left(\cos\left(\alpha \times \frac{\pi}{180}\right) \times (Y_{step} - \sqrt{1 + {Y_{step}}^2}) + Y_0\right)$$

(2)
$$Thickness\ Caliper = \sqrt{(X_f - X_i)^2 + (Y_f - Y_i)^2}$$

Where *i* = initial; *f* = final; α = skeleton supplementary angle; X_0 = x-axis centroid coordinate; Y_0 = y-axis centroid coordinate; X_i = *initial* horizontal vector component endpoint x-axis coordinate; Y_i = *initial* vertical vector component endpoint y-axis coordinate; X_f = *final* horizontal vector component endpoint x-axis coordinate; Y_f = *final* vertical vector component endpoint y-axis coordinate; X_{step} = 1-pixel increment in the x-axis; Y_{step} = 1-pixel increment in the y-axis.

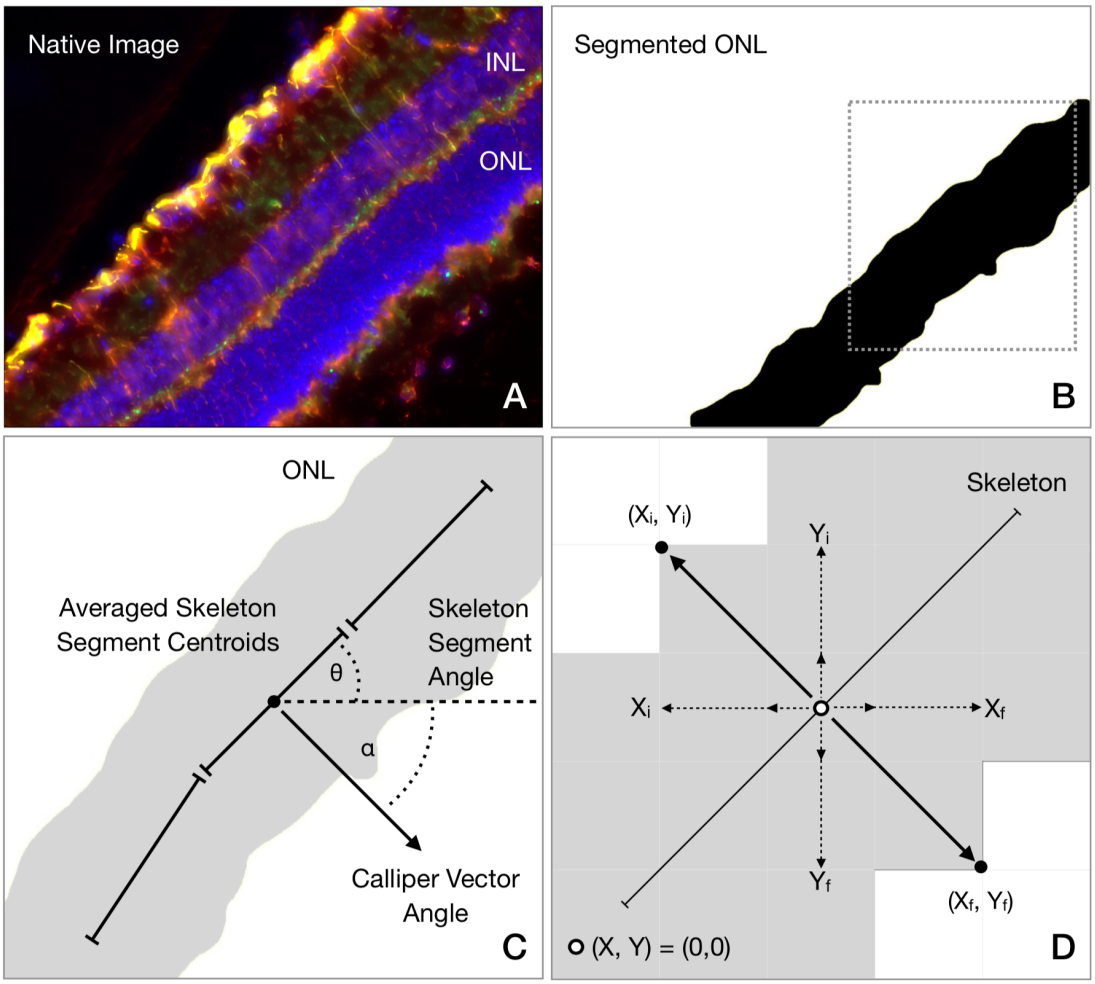


Figure 1. Outline of automated segmentation, thickness caliper constructs, and thickness measurements by ThicknessTool.

A) Representative native image acquired from mouse retina section and immunostaining. (B) Segmented outer nuclear layer (ONL) and mask. (C) Cropped area of the outer nuclear layer (ONL) mask. The angle of each skeleton segment (θ) was calculated respective to the zero-plane. The calliper angle (α), and resulting vector, were calculated as the supplementary angle of (θ). (D) Digital image representation displaying a pixel-grid and averaged skeleton segment calculated for the ONL mask, resulting vectors, and calliper endpoints.

4.3.3. Statistical Analyses

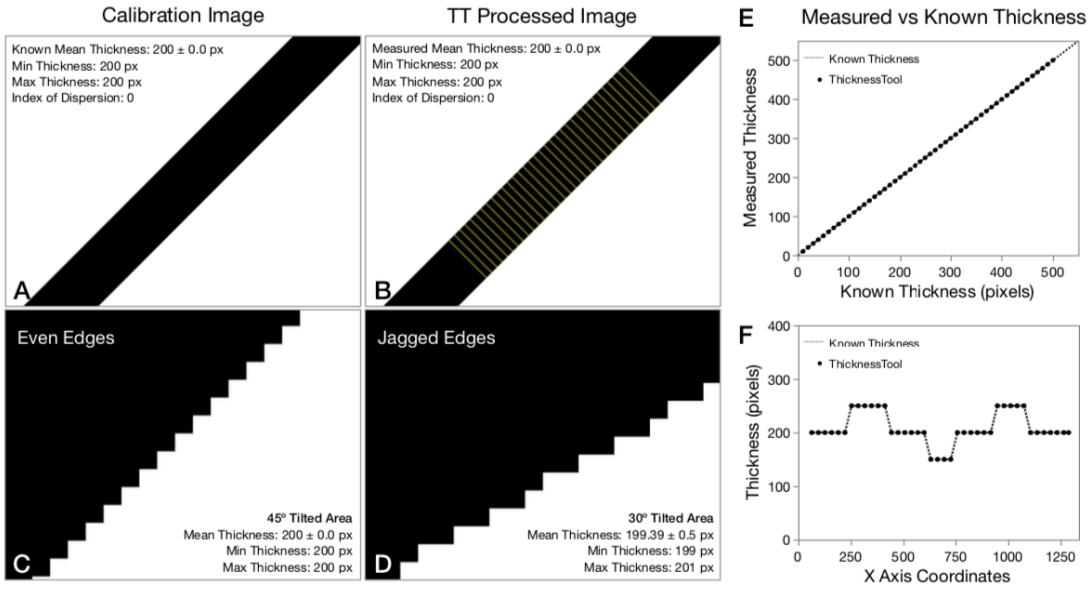
Statistical analysis was performed with SAS software (2019 SAS, NC). Normality was assessed with Shapiro-Wilk test. Statistical significance for differences between groups was determined with t-test for matched pairs or Wilcoxon signed-rank test for matched pairs, t-test for independent samples, and one-way ANOVA with Tukey post hoc correction for multiple comparisons. For correlation analysis, values were expressed as Pearson correlation coefficient (r) and statistical significance. Parallelism plots were constructed fitting a cubic spline with a default lambda of 0.05. Agreement analysis was performed by Bland-Altman plot method[11,12], and as previously described.4 Coefficient of variation (CoV) was calculated as the difference between measurements over their mean [(A – B)/(A + B)/2] and CoV% as CoV*100. The index of dispersion was calculated as variance-to-mean ratio. Results are expressed as mean ± standard deviation (SD). A p value of < 0.050 was considered statistically significant.

4.4. Results

4.4.1. ThicknessTool Calibration

To calibrate the ThicknessTool algorithm, we created three mock image datasets with masks of known area thickness and rotation. First, the TT was able to accurately measure areas of increasing thickness, from 10 to 500 pixels, with a bias of 0.00 ± 0.00 (Supplementary Figure 1E). The correlation coefficient for the known vs. measured thickness was 1.00 ($p < 0.001$). Following, we evaluated the performance of the algorithm at different layer rotations with 200-pixel masks at different angles. We predefined the known or theoretical mean, minimum, and maximum thickness to 200 pixels, and a coefficient of variation (CoV) of 0.00, and a margin of error of 1-pixel. As seen in Supplementary Table 1, the mean thickness measured by Thickness was 199.88 ± 0.25 pixels, with a CoV of 0.00 for the overall sample. There were no statistically significant differences between the measured overall mean thickness and hypothesized mean (p =

0.06). Mean thickness for images with even and jagged edges fulfilled the calibration criteria as previously defined, with CoV of 0.00 (Supplementary Figure 1A-D). Equivalence Test of mean, minimum and maximum caliper thickness rejected both null hypotheses (both $p < 0.001$), indicating that the observed difference does not exceed the defined threshold. Finally, we tested TT capability to detect true thinning and thickening in a mock image with decreasing and increasing thickness, respectively. TT accurately measured varying thickness ($r = 1.00$, $p < 0.001$), with a bias of 0.00 ± 0.00 and no significant difference between known vs. measured markers ($p = 1.000$) (Supplementary Figure 1F). These results indicate that the ThicknessTool can automatedly measure the different thicknesses and at different rotations in an accurate and precise manner.



Supplementary Figure 1. Outline of ThicknessTool (TT) Calibration.

(A) Representative calibration image of a known thickness (200 pixels). (B) Representative calibration image after processing with thickness calipers (200 pixels). (C) Cropped area in a 45º tilted image showing even edges. (D) Cropped area in a 30º tilted image showing jagged edges. (E) Known vs. measured thickness in images of increasing mask size. (F) Known vs. measured thickness in a mock image with abrupt thinning and thickening.

4.4.2. Training Dataset Mean Layer Thickness

To evaluate the performance of observers and TT across different images and thickness, we first constructed a parallelism plot of mean ONL and INL thickness. As seen in Figure 2, the inexperienced observer's mean measurements were consistently higher than the experienced and TT, which in return were similar and parallel to each other in both ONL and INL layers, suggesting a proportionate performance. Correlation coefficient (r) analysis between TT and observers' measurements were > 0.88 ($p < 0.001$) for the ONL and > 0.84 ($p < 0.001$) for the INL (Supplementary Table 2 and 3).

The mean ONL thickness was 65.26 ± 7.78 μm, 62.00 ± 7.14 μm, and 62.85 ± 7.33 μm for the inexperienced, experienced, and TT, respectively (Table 1). There were no significant differences between TT and any observer or average ONL measurements. The mean INL thickness was 39.13 ± 6.22 μm, 36.31 ± 5.17 μm, 36.12 ± 4.72 μm for the inexperienced, experienced and TT, respectively. In this group, there was only a significant difference between the inexperienced first measurement and TT (39.23 ± 6.30 μm vs. 36.12 ± 4.72 μm, $p = 0.044$), and no significant differences between TT and any observer average INL measurements. The average for the inexperienced vs. experienced was not significantly different, in either the ONL ($p = 0.219$) and INL ($p = 0.097$), with overlapping confidence intervals. Moreover, there were no significant differences between TT and any observer or average ONL/INL ratio measurements, with overlapping 95% confidence intervals. Altogether, these results suggest that no significant difference was evidenced regarding the observers' prior experience. However, TT ONL and INL measurements stand more closely to experienced observer's results. Moreover, TT measurements are not significantly different from any observers' ONL and INL average.

	Theoretical	Overall (n = 21)	Straight/Even Edges (n = 5)	Jagged Edges (n = 16)
Hypothesized Mean	-	200.00 ± 1	200.00 ± 1	200.00 ± 2
Mean Thickness	200.00 ± 0.00	199.88 ± 0.25	200.00 ± 0.00	199.85 ± 0.28
Minimum Thickness	200.00 ± 0.00	199.04 ± 0.74	200.00 ± 0.00	198.75 ± 0.74
Maximum Thickness	200.00 ± 0.00	200.57 ± 0.59	200.00 ± 0.00	200.75 ± 0.57
CoV Thickness	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
CoV Thickness (%)	0.00 ± 0.00	0.10 ± 0.09	0.00 ± 0.00	0.13 ± 0.07

CoV = coefficient of variation
Values shown are means ± standard deviations

Supplementary Table 1. Calibration dataset theoretical and ThicknessTool measurements.

4.4.3. Training Dataset Minimum and Maximum Layer Thickness

Qualitative assessment of observers' thickness calipers showed no overlapping in repeated measurements, suggesting that measurements are not likely reproducible, as seen in Supplementary Figure 2. Moreover, caliper over and undershoot was observed, which can lead to over and under measurement, respectively. In addition, the thickness vectors were not always perpendicular to the layer angle, hence prone to spurious results. Regardless of the lack of statistical differences in mean ONL and INL thickness between observers' average and TT measurements, we analyzed the minimum and maximum caliper values for each image.

First, we constructed a parallelism plot of the minimum and maximum ONL and INL thickness caliper per image and observer. As seen in Figure 3, the experienced observer had lower maximum and higher minimum ONL and INL values, without curve overlapping. Similarly, the inexperienced observer had higher minimum ONL and INL values without curve overlapping. These results suggest that experienced observers' measurements are not reproducible. In addition, there were no significant differences between TT and any observers' overall ONL minimum and maximum measurements (Table 2).

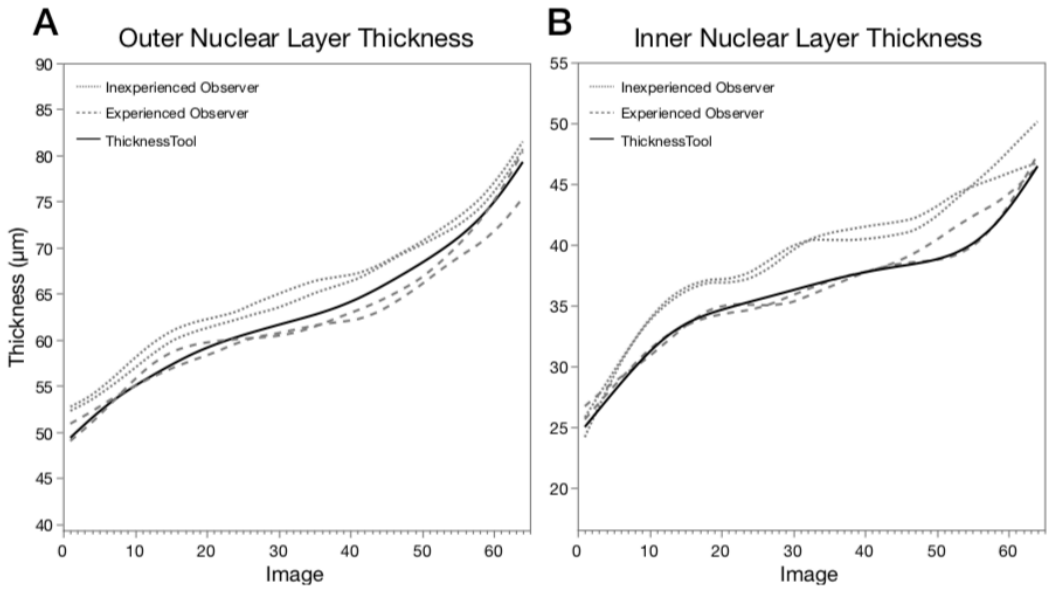


Figure 2. Parallelism plot of mean retinal thickness measurements of observers and ThicknessTool. Results were fitted to a cubic spline with a default lambda of 0.05. (A) Outer nuclear layer. (B) Inner nuclear layer. Dotted lines represent the first and second measurements by an inexperienced observer. Dashed lines represent the first and second measurements by an experienced observer. Solid lines represent measurements by ThicknessTool. Images are sorted by ThicknessTool values in ascending order. (n = 64)

	Inexperienced 1st	Inexperienced 2nd	Inexperienced Mean	Experienced 1st	Experienced 2nd	Experienced Mean	Observers Mean	ThicknessTool
Inexperienced 1st	-	0.96	0.99	0.90	0.91	0.92	0.97	0.88
Inexperienced 2nd	0.96	-	0.99	0.92	0.93	0.94	0.98	0.90
Inexperienced Mean	0.99	0.99	-	0.92	0.93	0.94	0.99	0.90
Experienced 1st	0.90	0.92	0.92	-	0.93	0.98	0.96	0.92
Experienced 2nd	0.91	0.93	0.93	0.93	-	0.98	0.97	0.90
Experienced Mean	0.92	0.94	0.94	0.98	0.98	-	0.98	0.93
Observers Mean	0.97	0.98	0.99	0.96	0.97	0.98	-	0.93
ThicknessTool	0.88	0.90	0.90	0.92	0.90	0.93	0.93	-

1st = first measurement; 2nd = second measurement
Values shown are Pearson correlation (*r*) coefficients (all coefficients $p < 0.001$)

Supplementary Table 2. Training dataset correlation analyses of outer nuclear layer thickness measurements.

	Inexperienced 1st	Inexperienced 2nd	Inexperienced Mean	Experienced 1st	Experienced 2nd	Experienced Mean	Observers Mean	ThicknessTool
Inexperienced 1st	-	0.94	0.98	0.86	0.80	0.88	0.97	0.86
Inexperienced 2nd	0.94	-	0.98	0.84	0.78	0.86	0.96	0.84
Inexperienced Mean	0.98	0.98	-	0.86	0.80	0.89	0.98	0.86
Experienced 1st	0.86	0.84	0.86	-	0.77	0.95	0.93	0.87
Experienced 2nd	0.80	0.78	0.80	0.77	-	0.94	0.89	0.87
Experienced Mean	0.88	0.86	0.89	0.95	0.94	-	0.97	0.92
Observers Mean	0.97	0.96	0.98	0.93	0.89	0.97	-	0.92
ThicknessTool	0.86	0.84	0.86	0.87	0.87	0.92	0.92	-

1st = first measurement; 2nd = second measurement
Values shown are Pearson correlation (*r*) coefficients (all coefficients $p < 0.001$)

Supplementary Table 3. Training dataset correlation analyses of inner nuclear layer thickness measurements.

	Outer Nuclear Layer			Inner Nuclear Layer			ONL/INL Ratio	
	Mean ± SD	95% CI	*P* value*	Mean ± SD	95% CI	*P* value*	Mean ± SD	*P* value*
Inexperienced 1st	65.69 ± 7.94	63.70 - 67.67	0.394	39.23 ± 6.30	37.66 - 40.80	**0.044**	1.70 ± 0.26	0.910
Inexperienced 2nd	64.83 ± 7.78	62.88 - 66.77	0.814	39.03 ± 6.33	37.45 - 40.61	0.078	1.69 ± 0.26	0.768
Inexperienced Average	65.26 ± 7.78	63.31 - 67.20	0.613	39.13 ± 6.22	37.57 - 40.68	0.059	1.70 ± 0.26	0.832
Experienced 1st	62.44 ± 7.72	60.51 - 64.37	1.000	36.50 ± 5.70	35.07 - 37.92	1.000	1.74 ± 0.26	0.999
Experienced 2nd	61.56 ± 6.83	59.84 - 63.26	0.978	36.12 ± 5.29	34.79 - 37.44	1.000	1.73 ± 0.21	0.994
Experienced Average	62.00 ± 7.14	60.21 - 63.78	0.998	36.31 ± 5.17	35.01 - 37.60	1.000	1.73 ± 0.21	0.996
Observers Average	63.47 ± 7.60	61.79 - 65.46	0.998	37.64 ± 5.68	36.34 - 39.10	0.807	1.71 ± 0.22	0.944
ThicknessTool	62.85 ± 7.33	61.01 - 64.68	-	36.12 ± 4.72	34.94 - 37.30	-	1.76 ± 0.22	-

ONL = outer nuclear layer; INL = inner nuclear layer; SD = standard deviation; CI = confidence interval; 1st = first measurement; 2nd = second measurement; * = comparisons for all pairs vs ThicknessTool using Tukey-Kramer HSD test
Values shown are means ± standard deviations

Table 1. Training dataset mean thickness measured by observers and ThicknessTool.

However, we found significant differences across individual observer's minimum and maximum measurements and between the inexperienced overall minimum INL and TT measurements ($p < 0.001$). In addition, the coefficient of variability between measurements was significantly different between inexperienced and observers' average measurements for both ONL and INL ($p < 0.001$) and between the experienced observer and observers' average measurements for the ONL ($p < 0.009$) (Supplementary Table 4). In summary, given the qualitative assessment of observers' calipers' lack of repeatability

together with the quantitative analyses between manual calipers and a calibrated automated algorithm, results indicate that manual calipers measurements are not accurate nor reproducible.

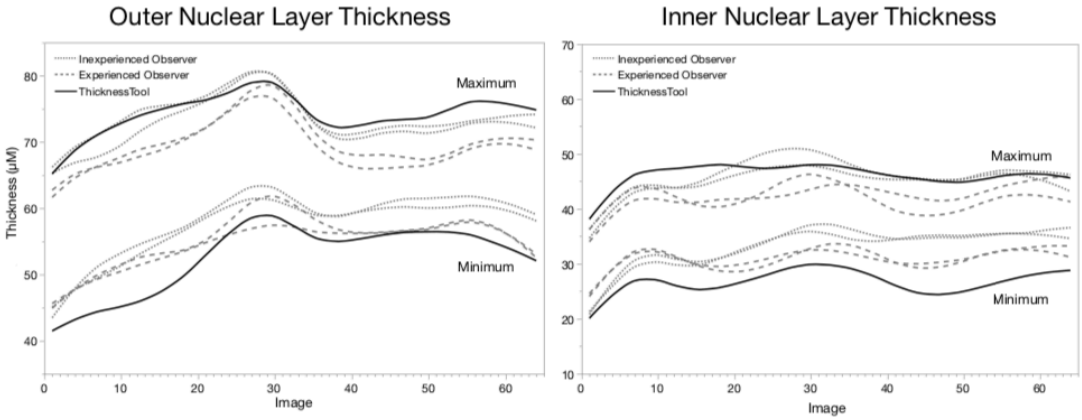
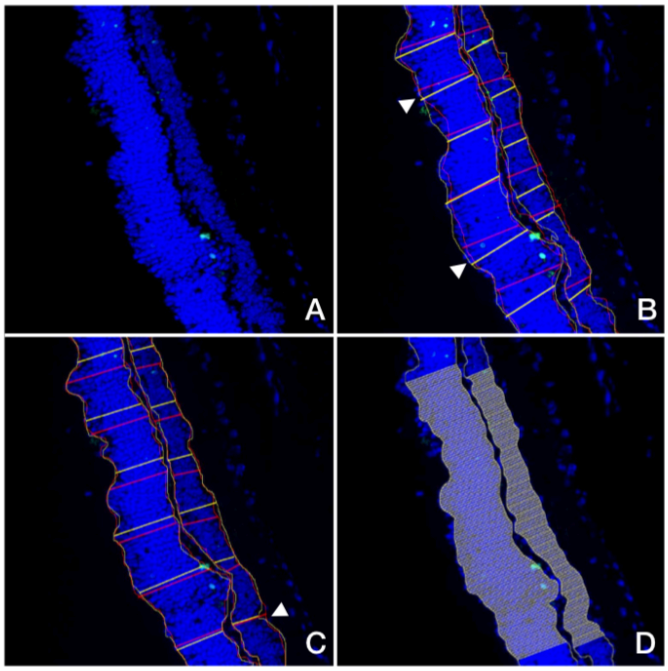


Figure 3. Parallelism plot of minimum and maximum retinal thickness measurements of observers and ThicknessTool.

Results were fitted to a cubic spline with a default lambda of 0.05. (A) Outer nuclear layer. (B) Inner nuclear layer. Dotted lines represent the first and second measurements by an inexperienced observer. Dashed lines represent the first and second measurements by an experienced observer. Solid lines represent measurements by ThicknessTool. (n = 64)



Supplementary Figure 2. Qualitative assessment of manual and automated thickness measurements of outer (ONL) and inner nuclear (INL) retina layers.

(A) Representative native image of mouse retinal section. (B) Inexperienced observer's measurements, displaying a lack of reproducibility and oblique caliper vector (arrowheads). (C) Experienced observer's measurements, displaying a lack of reproducibility and overshooting caliper vector (arrowhead). (D) ThicknessTool measurement of the INL and ONL at 1-pixel interval.

	Outer Nuclear Layer				**Inner Nuclear Layer**			
	Minimum ± SD	*P* value*	Maximum ± SD	*P* value*	Minimum ± SD	*P* value*	Maximum ± SD	*P* value*
Inexperienced 1st	58.31 ± 8.76	**0.003**	73.94 ± 10.09	1.000	33.12 ± 7.30	**< 0.001**	45.44 ± 6.90	0.997
Inexperienced 2nd	57.35 ± 8.14	**0.031**	72.78 ± 10.06	0.987	33.39 ± 6.97	**< 0.001**	45.86 ± 6.91	1.000
Inexperienced Overall	56.28 ± 8.54	0.218	75.00 ± 10.33	0.999	31.73 ± 7.12	**< 0.001**	46.96 ± 6.97	0.999
Experienced 1st	55.51 ± 8.19	0.534	70.03 ± 9.67	0.198	31.01 ± 5.88	**0.002**	42.45 ± 7.09	**0.048**
Experienced 2nd	54.60 ± 7.67	0.899	69.09 ± 8.53	**0.049**	30.95 ± 4.92	**0.002**	41.91 ± 7.35	**0.012**
Experienced Overall	53.32 ± 7.83	0.999	71.34 ± 9.43	0.667	29.41 ± 4.89	0.207	44.53 ± 7.83	0.858
Observers Overall	52.44 ± 8.07	1.000	75.74 ± 10.29	0.990	28.36 ± 5.52	0.812	47.92 ± 7.78	0.894
ThicknessTool	52.73 ± 8.20	-	74.30 ± 8.88	-	26.76 ± 5.23	-	46.28 ± 5.72	-

SD = standard deviation; 1st = first measurement; 2nd = second measurement; * = comparisons for all pairs vs ThicknessTool using Tukey-Kramer HSD test
Values shown are means ± standard deviations

Table 2. Training dataset minimum and maximum thickness measured by observers and ThicknessTool.

	Outer Nuclear Layer		Inner Nuclear Layer	
	Mean ± SD	*P* value*	Mean ± SD	*P* value*
Inexperienced	2.77 ± 2.52	**< 0.001**	4.82 ± 3.45	**< 0.001**
Experienced	3.81 ± 2.83	**0.009**	7.80 ± 6.42	0.602
Observers Average	5.39 ± 3.55	-	8.70 ± 5.55	-

SD = standard deviation; * = comparison for pair vs observers average using Tukey-Kramer HSD test
Values shown are means ± standard deviations

Supplementary Table 4. Training dataset thickness measurement coefficient of variation between observers.

4.4.4. Agreement Analysis between Manual and ThicknessTool Measurements

We performed a bias analysis using Bland and Altman method.12 For the ONL group, as seen in Table 3, we found statistically significant differences when comparing paired measurements, with a greater bias in inexperienced mean vs. experienced mean measurements (3.25 ± 0.33 μm, $p < 0.001$). In contrast, we found the lowest bias when comparing TT vs. observers' means (0.77 ± 0.34 μm, $p = 0.028$). For the INL group, we found statistically significant differences when comparing inexperienced mean vs. experienced mean ($p < 0.001$), TT vs. inexperienced mean ($p < 0.001$), and TT vs. observers' mean ($p < 0.001$). Most importantly, the bias between TT vs. observers' INL mean (1.59 ± 0.28 μm) was lower than the those between any observers mean against each other (2.82 ± 0.36 μm). In addition, Bland-Altman plots displayed random variability within measurements (Figure 4). In conclusion, these results indicate that Thickness can measure ONL and INL with lower bias than observers' average against each other for both ONL and INL.

	Bias ± SE	Lower 95% LoA	Upper 95% LoA	LoA Interval	*P* value*
Outer Nuclear Layer					
Inexperienced 1st vs. Inexperienced 2nd	0.86 ± 0.27	-3.37	5.09	8.47	**0.002**
Experienced 1st vs. Experienced 2nd	0.880 ± 0.36	-4.76	6.52	11.29	**0.018**
Experienced Mean vs. Inexperienced Mean	3.25 ± 0.33	-1.92	8.42	10.35	**< 0.001**
ThicknessTool vs. Inexperienced Mean	2.40 ± 0.42	-4.19	8.99	13.17	**< 0.001**
ThicknessTool vs. Experienced Mean	0.85 ± 0.34	-4.48	6.18	10.66	**0.015**
ThicknessTool vs. Observers Mean	0.77 ± 0.34	-4.56	6.10	10.66	**0.028**
Inner Nuclear Layer					
Inexperienced 1st vs. Inexperienced 2nd	0.20 ± 0.27	-4.03	4.43	8.47	0.465
Experienced 1st vs. Experienced 2nd	0.38 ± 0.46	-6.83	7.59	14.43	0.412
Experienced Mean vs. Inexperienced Mean	2.82 ± 0.36	-2.82	8.46	11.29	**< 0.001**
ThicknessTool vs. Inexperienced Mean	3.00 ± 0.40	-3.27	9.27	12.54	**< 0.001**
ThicknessTool vs. Experienced Mean	0.18 ± 0.25	-3.74	4.10	7.84	0.462
ThicknessTool vs. Observers Mean	1.59 ± 0.28	-2.80	5.98	8.78	**< 0.001**

SE = standard error; LoA = limit of agreement; 1st = first measurement; 2nd = second measurement; * = T-Test for matched pairs
Values shown are means ± standard errors

Table 3. Training dataset agreement analysis of thickness measurements between observers and ThicknessTool.

ıre 4

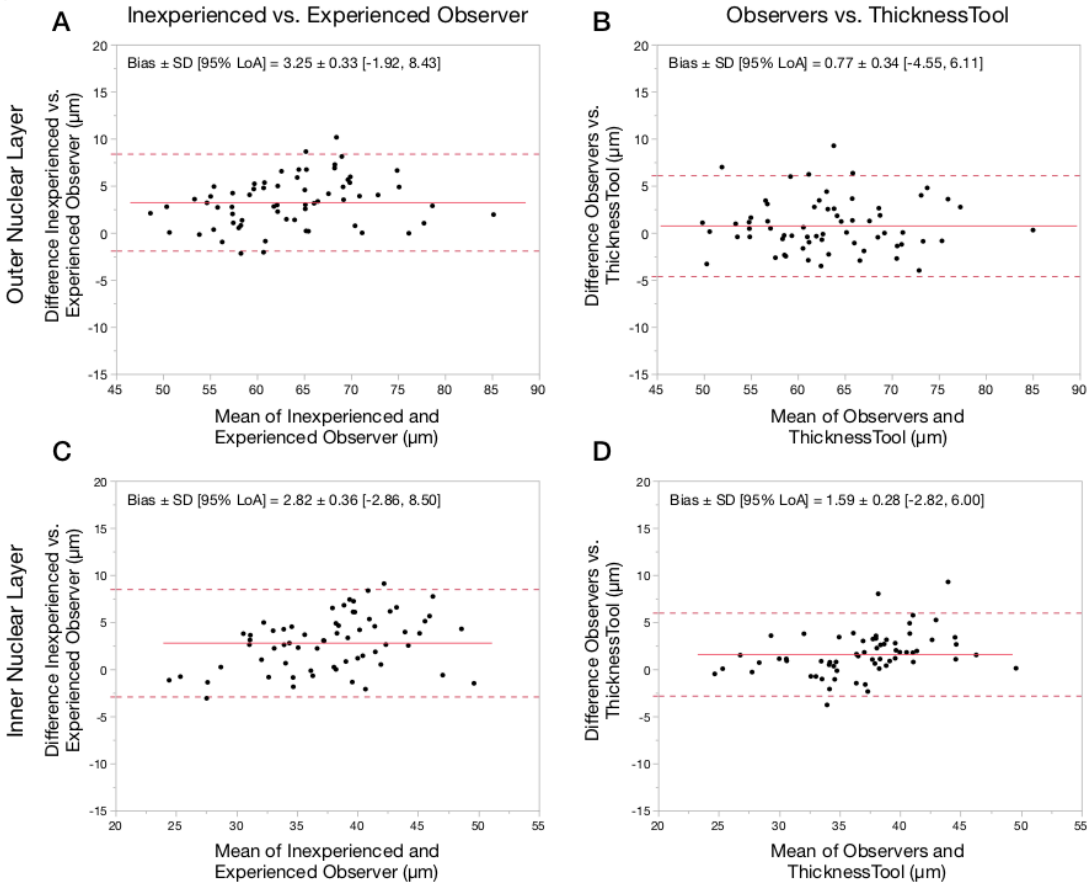


Figure 4. Agreement analysis of observers and ThicknessTool by Bland-Altman plots. Agreement between inexperienced and experienced average mean (A) ONL and (C) INL thickness. Agreement between observers and ThicknessTool (B) ONL and (D) INL thickness. (n = 64) SD = standard deviation; LoA = limit of agreement.

4.4.5. Validation of ThicknessTool in a Retinal Detachment Model

To definitively validate the TT, we tested the tool in an experimental setting using a retinal detachment model, which causes photoreceptor cell death and subsequent ONL thinning and distortion (Figure 5). We quantitated the ONL and INL thickness manually and using TT (Table 4). Interestingly, manual measurements failed to show significant ONL thinning in the detached retina (p = 0.069), in comparison to TT (p = 0.006). No differences were observed in the INL between the attached and detached retina by both methods as expected. Both methods showed significant differences in the ONL/INL ratio in the detached retina. These results suggest that TT can detect significant true thinning beyond manual measurement capabilities. Collectively, these results suggest that TT is a sensitive, precise, and reliable tool to measure retinal nuclear layers in experimental models of disease.

	Outer Nuclear Layer			**Inner Nuclear Layer**			**ONL/ONL + INL Ratio**		
	Attached	Detached	*P* value*	Attached	Detached	*P* value*	Attached	Detached	*P* value*
Manual	61.24 ± 8.39	50.47 ± 12.80	0.069	31.01 ± 2.58	33.43 ± 6.40	0.347	1.97 ± 0.22	1.51 ± 0.28	**0.003**
ThicknessTool	70.82 ± 10.04	56.14 ± 8.25	**0.006**	38.58 ± 4.34	39.79 ± 5.24	0.624	1.83 ± 0.16	1.41 ± 0.12	**< 0.001**

ONL = outer nuclear layer; INL = inner nuclear layer; * = T-Test for independent samples
Values shown are means ± standard deviations

Table 4. Validation dataset manual and ThicknessTool measurements in a retinal detachment model.

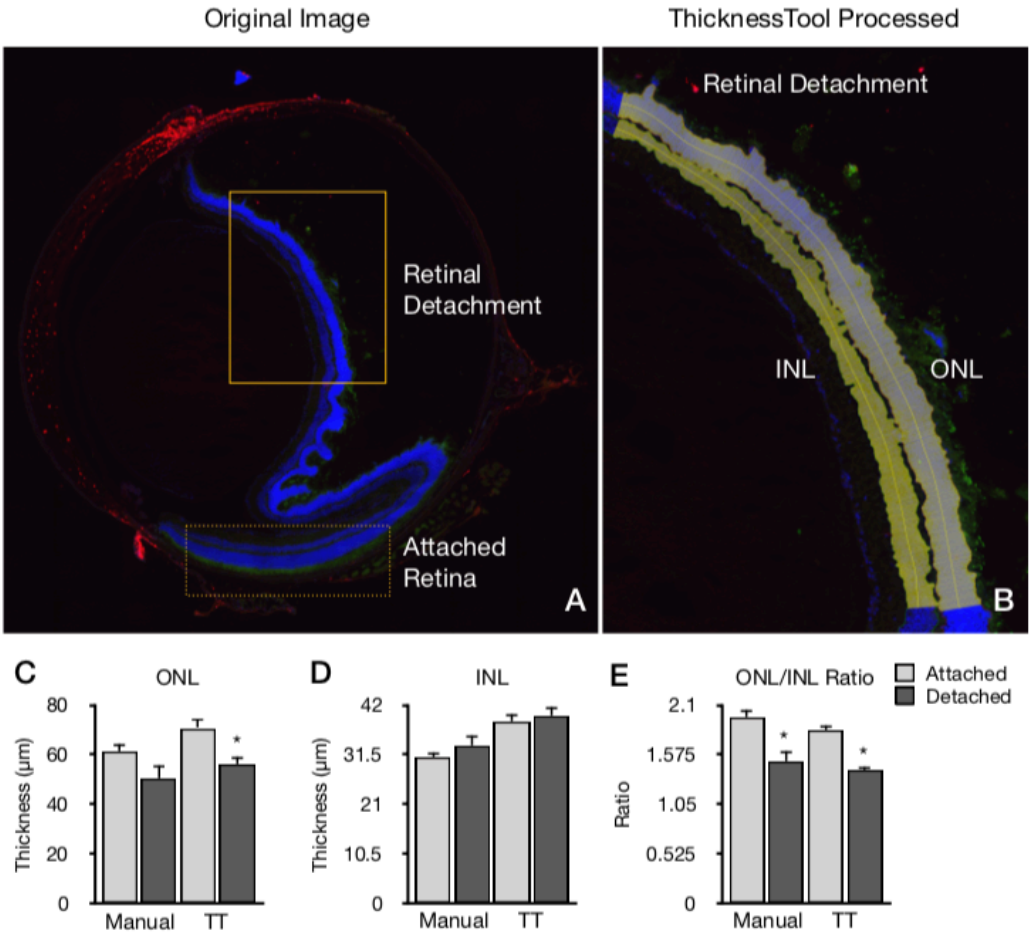


Figure 5. Validation of ThicknessTool in a retinal detachment model.

(A) Representative image of mouse retina section of a retinal detachment displaying the attached (dotted lines box) and detached retina (solid lines box). (B) Representative image of ThicknessTool measurement of detached retina inner nuclear layer (INL) and outer nuclear layer (ONL) at 1-pixel interval calipers. (C) Outer nuclear layer (ONL) and (D) Inner nuclear layer (INL) thickness measurements from observer manual calipers and ThicknessTool in the attached and detached retina. (E) ONL/INL ratio measurements of observer manual calipers and ThicknessTool. (n = 8 per group, *p ≤ .05)

4.4.6. Application of ThicknessTool in Digital Images

In order to examine the applicability of TT in different image modalities, we quantitated the thickness in various digital images. As seen in Figure 6, TT was capable of

quantitating the ONL thickness and corresponding profile in histology sections. Moreover, we challenged this tool in digital images of macular spectral-domain optical coherence tomography, with satisfactory results. Finally, we tested the performance in the quantitative assessment of retinal vessels in fluorescein angiogram images. The TT was able to draw accurate calipers with the correct vector in all cases. Collectively, these results indicate that ThicknessTool is a versatile tool with multi-platform capabilities.

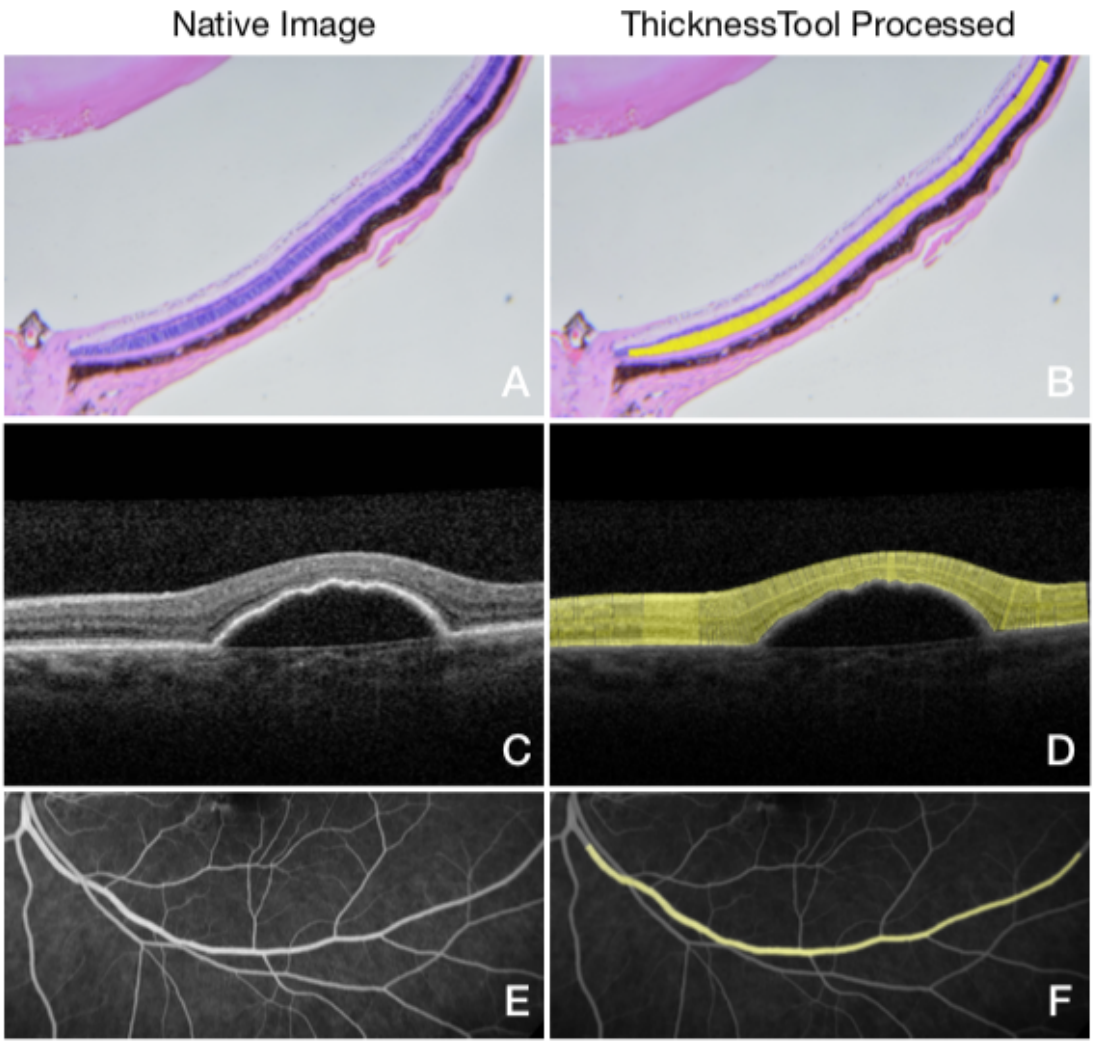


Figure 6. ThicknessTool application in digital images.
(A-B) Hematoxylin-eosin staining of retinal cryosection. (C-D) Macular spectral-domain optical coherence tomography scan showing a pigment epithelial detachment in age-related macular degeneration. (E-F) Fluorescein angiogram of retinal vessels.

4.5. Discussion

In this study, we developed a sensitive, precise, and reliable tool to measure retinal layers. The performance of this tool was tested in calibration, training, and validation datasets with adequate and reliable performance. Collectively, results validate ThicknessTool for automated retinal thickness measurement. We speculate that this tool can aid in achieving precise, accurate, and fast measurements in multi-platform digital images.

In the development of new measurement tools, calibration images are mandatory to benchmark and finetune the performance of image processing algorithms to a standard.[13] We designed mock images with various known thickness ranging from 10 pixels (given a single nuclei size in our images) to 500 pixels (twice the outer nuclear layer measured in our dataset). In addition, we created a dataset with multiple mask rotations to mimic the random mask orientations. This step is necessary as TT calipers are constructed in a vectorial manner, in number and angles as the user or region of interest requires. Moreover, we tested TT capability to detect true thinning and thickening in a mock image with decreasing and increasing thickness, respectively. We found that the TT can automatedly and accurately measure thickness when challenged with different thicknesses and rotations.

Accurate assessment of photoreceptor cell death currently relies on manual counting by masked observers. Ideally, two masked two observers quantitate retina thickness, and their average is used to determine the overall outcome. We found that single observers' measurements were prone to over and underestimation, regardless of their previous experience. In addition, observers draw incorrectly angled calipers, which can indirectly contribute to the abovementioned bias. Moreover, calipers were not located in corresponding locations in repeated same-image analysis, which dampens the reproducibility of the measurements. Altogether, this leads to a lack of accurate or reproducible retina thickness quantitation. This matter is of critical importance, as this bias can lead to α or β errors. For example, finding no difference in retinal thickness in cases where a neuroprotective treatment is effective, or even worse, finding a significant

difference in retinal thickness in an ineffective treatment.[14,15] Such errors are not only misleading to the scientific community but also high priced.

Regardless of the performance of manual measurements to date, it remains the gold standard. For purposes of method agreement, we refrained from relying on mean comparison or correlation, as Bland and Altman have previously reported in their seminal work.[11,16] As authors describe, such agreement must be quantified, and cannot be assessed by mean comparison or hypothesis testing, as it does not exist in a present or absent binary form.12 In addition, high correlation does not necessarily imply a high agreement between two methods, as it quantifies the degree to which two methods are related.[17] Agreement analysis showed lower bias of TT against observers than observers' average against each other for both ONL and INL. Taken into consideration ITT capability to detect true mask thinning and thickening, collectively, these results confer to TT a higher measurement agreement.

Moreover, the automated thickness analysis by TT yields a series of convenient metrics that remain hidden otherwise in manual measurements. On this note, Byun et al. have developed pioneering methods to estimate the ONL thickness exclusively.[9] However, their approach was not validated by thickness agreement analysis with manual observers. In their work, authors also described a method to quantify structural distortions in retinal tissue before and after injury, also known as distortion index. In essence, the authors stated that structural changes in the ONL are directly proportional to the distortion index. In our work, we estimated overall distortion in a similar approach using the index of dispersion (ID), calculated as the variance-to-mean ratio, as previously defined by Cox and Lewis.[18] Similarly, the index of dispersion allows the assessment of the overall caliper variation, which can provide detailed layer architecture data. Briefly, an ID of 0 represents no dispersion; therefore, we used the ID for calibration purposes, as mock images of known and constant thickness were uniform rectangles (ID = 0). We speculate that the index of dispersion can be equally used for further morphometric assessment.

Several limitations to this work should be considered. First, despite the calibration analysis performed beforehand, the validation of TT can only be performed against

manual measurements. This process is cumbersome, as observers' measurements carry an intrinsic bias. However, observers showed comparable performance, with overlapping confidence intervals. Second, TT analysis should be performed in areas distant from image edges, as skeleton are trimmed to avoid caliper-to-end image measurements. Despite the fact that TT allows skeleton and edge trimming, users should evaluate processed images for qualitative control. Third, TT measures thickness in binary segmented masks. Therefore, its performance is as strong as the segmentation algorithm used. In our case, we build this tool based on a previously validated algorithm for retina layer segmentation.4 We recommend users to first tailor and validate their segmentation algorithm for optimal performance. Finally, regardless of the potential image modality application, this tool has been validated only for ONL and INL measurements in DAPI nuclei-stained murine retina cryosections. We believe future work can help validate TT for more imaging platforms.

In summary, we developed a fast, accurate, and precise retinal thickness measurement tool with multi-platform capabilities. In addition, the TT can be customized to user-defined preferences. This includes caliper number with a 1-pixel minimum distance between calipers, skeleton and edge trimming, and automated or manual layer segmentation. In addition, given the potential use with multi-platform capabilities, the source code can be easily modified to fit different applications. Finally, this tool is freely available to download and use. We expect this measurement tool could improve the outcomes and reduce bias in retina research.

4.6. Acknowledgements

Authors contributions: D.E.M and D.G.V. designed research. D.E.M. coded the ThicknessTool, performed the statistical analysis, and drafted the manuscript. S.N., T.U., T.Z., D.J., and C.K. measured retinal thickness from the digital image dataset and contributed to the manuscript. J.W.M., and D.G.V. reviewed the data and made substantial intellectual contributions to the research design and manuscript. Please visit

https://imagej.net/RETINA_Analysis_Toolkit to download the ThicknessTool.

4.7. Funding Support

This work was supported by Bayer Healthcare Global Ophthalmology Awards Program (DEM); the Yeatts Family Foundation (DGV, JWM); Loefflers family foundation (JWM, DGV); a Macula Society Research Grant award (DGV); a Physician Scientist Award (DGV) and an unrestricted grant (JWM) from the Research to Prevent Blindness Foundation; NEI Grant R21EY023079-01A1 (DGV); and NEI Grant EY014104 (MEEI Core Grant). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

4.8. Financial Disclosures

The authors have no relevant financial relationships or interests to disclose.

4.9. Chapter Summary in Spanish

4.9.1. Introducción

La muerte celular progresiva de fotorreceptores es un factor significativo en las enfermedades degenerativas de la retina. Esta pérdida celular ha sido abordada por una miríada de enfoques, predominantemente en modelos experimentales de enfermedades de la retina, como el desprendimiento de retina.[1,2] Entre estos métodos, los ensayos de muerte celular de fotorreceptores,[3,4] el contaje de células de la capa nuclear externa[5] y el espesor de la capa nuclear externa,[6,7] se han utilizado en modelos animales para cuantificar la degeneración de fotorreceptores. Dadas las limitaciones de técnica y tiempo

relacionadas con los ensayos de muerte celular y el conteo manual; hasta la fecha, el grosor de la capa nuclear externa (*outer nuclear layer,* ONL) se ha utilizado ampliamente como un método práctico para estimar la profundidad de la muerte celular de los fotorreceptores.[6–8]

La utilidad de la cuantificación del espesor de la capa ONL de la retina para inferir la degeneración de fotorreceptores se basa en su versatilidad y rápida medición. Sin embargo, la ventaja de este método puede verse comprometida ya que los calibres son dibujados manualmente y medidos por un observador. Además, dado el radio de curvatura de la retina, dichos calibres deben ser perpendiculares al contorno de la capa nuclear. Además, dada la gran ampliación de las imágenes de microscopio, a menudo se analiza un área pequeña para acelerar el análisis, lo que puede conducir a un sesgo. Además, dadas las nuevas modalidades de imagen, un método versátil que pueda adaptarse a dichas plataformas es ideal. Por lo tanto, es de gran necesidad una herramienta que puede adaptarse a la arquitectura de capa y al contorno para medir el grosor en un segmento amplio, en múltiples capas, en imágenes individuales o en mosaicos, y en múltiples plataformas. A pesar de un trabajo pionero en esta área,[9] actualmente no existe una herramienta disponible gratuita para la plataforma ImageJ, que cuantifique de manera automática el espesor de múltiples capas de la retina en imágenes grandes y en diferentes modalidades.

El propósito de este trabajo fue desarrollar una herramienta de medición automatizada del espesor de las capas nucleares retinianas para la plataforma ImageJ, que puede cuantificar dichas capas siguiendo el contorno de la retina, con calibres tan próximos como 1 píxel entre sí. Para este propósito, desarrollamos ThicknessTool (TT) y validamos su precisión mediante calibración objetiva y análisis de acuerdo con dos observadores enmascarados. Descubrimos que esta herramienta de medición puede proporcionar mediciones de espesor precisas y confiables. Además, TT puede procesar imágenes de múltiples modalidades de imágenes.

4.9.2. Resultados

Para calibrar el algoritmo ThicknessTool, creamos tres grupos de imágenes creados con máscaras de espesor y rotación conocidas. Primero, el TT pudo medir con precisión áreas de grosor creciente, de 10 a 500 píxeles, con un sesgo de 0.00 ± 0.00 (Figura complementaria 1E). El coeficiente de correlación para el espesor conocido frente al medido fue de 1.00 (p <0.001). A continuación, evaluamos el rendimiento del algoritmo en diferentes rotaciones con máscaras de 200 píxeles en diferentes ángulos. Predefinimos la media conocida o teórica, el espesor mínimo y máximo a 200 píxeles, y un coeficiente de variación (CV) de 0.00, y un margen de error de 1 píxel. Como se ve en la Tabla 1 suplementaria, el grosor medio medido fue de 199.88 ± 0.25 píxeles, con un CV de 0.00 para la muestra general. No hubo diferencias estadísticamente significativas entre el grosor medio general medido y la media hipotética (p = 0.06). El grosor medio de las imágenes con bordes uniformes e irregulares cumplió los criterios de calibración definidos anteriormente, con un CV de 0.00 (Figura complementaria 1A-D). La prueba de equivalencia del grosor medio, mínimo, y máximo del calibrador rechazó ambas hipótesis nulas (ambas p <0.001), lo que indica que la diferencia observada no excede el umbral definido. Finalmente, probamos la capacidad del TT para detectar el verdadero adelgazamiento y engrosamiento en una imagen simulada con grosores decrecientes y crecientes, respectivamente. El TT midió con precisión el grosor variable (r = 1.00, p <0.001), con un sesgo de 0.00 ± 0.00 y ninguna diferencia significativa entre los marcadores conocidos y medidos (p = 1.000) (Figura 1F complementaria). Estos resultados indican que el ThicknessTool puede medir automáticamente los diferentes espesores y en diferentes rotaciones de manera precisa y precisa.

Para evaluar el rendimiento de los observadores y el TT en diferentes imágenes y grosores, primero construimos un diagrama de paralelismo del grosor de la ONL y INL medios. Como se ve en la Figura 2, las mediciones medias del observador inexperto fueron consistentemente más altas que las del observador experto y TT, que a cambio fueron similares y paralelas entre sí en las capas ONL e INL, lo que sugiere un rendimiento similar. El análisis del coeficiente de correlación (r) entre el TT y las

mediciones de los observadores fueron > 0.88 (p <0.001) para la ONL y > 0.84 (p <0,001) para el INL (tablas complementarias 2 y 3).

El grosor medio de ONL fue de 65.26 ± 7.78 µm, 62.00 ± 7.14 µm y 62.85 ± 7.33 µm para el observador inexperto, experimentado, y TT, respectivamente (Tabla 1). No hubo diferencias significativas entre TT y cualquier observador o mediciones promedio de ONL. El grosor medio de INL fue de 39.13 ± 6.22 µm, 36.31 ± 5.17 µm, 36.12 ± 4.72 µm para el observador inexperto, experimentado, y TT, respectivamente. En este grupo, solo hubo una diferencia significativa entre la primera medición del observador inexperto y TT (39.23 ± 6.30 µm vs. 36.12 ± 4.72 µm, p = 0.044), y no hubo diferencias significativas entre TT y cualquier medida promedio de INL de los observadores. El promedio para el observador inexperto frente a el observador experimentado no fue significativamente diferente, ya sea en el ONL (p = 0.219) y el INL (p = 0.097), con intervalos de confianza del 95% superpuestos. Además, no hubo diferencias significativas entre TT y cualquier observador o promedio en la proporción ONL/INL, con intervalos de confianza del 95% superpuestos. En conjunto, estos resultados sugieren que no existen diferencias significativas con respecto a la experiencia previa de los observadores. Sin embargo, las mediciones TT para la ONL e INL están más cerca de los resultados del observador experimentado. Además, las mediciones de TT no son significativamente diferentes del promedio ONL e INL de ningún observador.

La evaluación cualitativa de los calibres de espesor en los observadores no mostró solapamiento en las mediciones repetidas, lo que sugiere que las mediciones no son reproducibles, como se ve en la Figura complementaria 2. Además, se observaron calibres por encima y por debajo del área, que puede conducir a una medición excesiva o insuficiente, respectivamente. Además, los vectores de grosor no siempre fueron perpendiculares al ángulo de la capa, por lo tanto, propensos a resultados espurios. Independientemente de la falta de diferencias estadísticamente significativas en el grosor medio de ONL e INL entre las mediciones promedio de los observadores y el TT, analizamos los valores de calibre mínimo y máximo para cada imagen.

Primero, construimos una gráfica de paralelismo del calibre de grosor mínimo y máximo de ONL e INL por imagen y observador. Como se ve en la Figura 3, el observador experimentado tiene valores ONL e INL mínimos máximos y mínimos más bajos, sin superposición de las curvas. Del mismo modo, el observador inexperto tiene valores mínimos de ONL e INL mínimos más altos sin superposición de curvas. Estos resultados sugieren que las mediciones de los observadores no son reproducibles. Además, no hubo diferencias significativas entre TT y las mediciones mínimas y máximas generales de ONL de los observadores (Tabla 2). Además, el coeficiente de variabilidad entre las mediciones fue significativamente diferente entre las mediciones promedio de los observadores para ONL e INL ($p < 0.001$) (Tabla complementaria 4) En resumen, dada la evaluación cualitativa de la falta de repetitividad de los calibres de los observadores, junto con los análisis cuantitativos entre calibres manuales y un algoritmo automatizado calibrado, los resultados indican que las mediciones manuales no son precisas ni reproducibles.

Realizamos un análisis de sesgo utilizando el método de Bland y Altman.[12] Para la ONL, como se ve en la Tabla 3, encontramos diferencias estadísticamente significativas al comparar mediciones pareadas, con un mayor sesgo en la media del observador inexperto frente a las mediciones medias del observador experimentado (3.25 ± 0.33 µm, $p < 0.001$). Por el contrario, encontramos el sesgo más bajo al comparar TT versus medias de los observadores (0.77 ± 0.34 µm, $p = 0.028$). Para la INL, encontramos diferencias estadísticamente significativas al comparar la media del observador inexperto versus la media del observador experimentado ($p < 0.001$), TT vs media observador inexperto ($p < 0.001$), y TT vs media de los observadores ($p < 0.001$). Lo que es más importante, el sesgo entre la media de INL de los TT frente a los observadores (1.59 ± 0.28 µm) fue menor que el promedio entre los observadores entre sí (2.82 ± 0.36 µm). Además, las gráficas de Bland-Altman mostraron una variabilidad aleatoria dentro de las mediciones (Figura 4). En conclusión, estos resultados indican que el TT puede medir las capas nucleares de la retina con un sesgo menor que el promedio de los observadores entre sí, tanto para ONL como para INL.

Para validar definitivamente el TT, probamos la herramienta en un diseño experimental utilizando un modelo de desprendimiento de retina, que causa la muerte de las fotorreceptores y el consiguiente adelgazamiento y distorsión de la ONL (Figura 5). Cuantificamos el grosor de la ONL e INL manualmente y utilizando TT (Tabla 4). Curiosamente, las mediciones manuales no pudieron mostrar un adelgazamiento significativo de ONL en la retina desprendida ($p = 0.069$), en comparación con TT ($p = 0.006$). No se observaron diferencias en la INL entre la retina no deprendida y desprendida por ambos métodos, como es de esperar. Ambos métodos mostraron diferencias significativas en la proporción ONL/INL en la retina desprendida. Estos resultados sugieren que TT puede detectar adelgazamiento verdadero y significativo más allá de las capacidades de medición manual. En conjunto, estos resultados sugieren que TT es una herramienta sensible, precisa y confiable para medir las capas nucleares de la retina en modelos experimentales.

Para examinar la aplicabilidad de TT en diferentes modalidades de imagen, cuantificamos el grosor en varias imágenes digitales. Como se ve en la Figura 6, TT fue capaz de cuantificar el grosor de la ONL y el perfil correspondiente en secciones de histología teñidas con hematoxilina y eosina. Además, probamos esta herramienta en imágenes digitales de tomografía de coherencia óptica de dominio espectral de la mácula, con resultados satisfactorios. Finalmente, evaluamos el rendimiento en la evaluación cuantitativa de los vasos retinianos en imágenes de angiografía con fluoresceína. El TT fue capaz de dibujar calibres precisos con un vector correcto en todos los casos. En conjunto, estos resultados indican que ThicknessTool es una herramienta versátil con capacidades multi-plataforma.

4.9.3. Discusión

En este estudio, desarrollamos una herramienta sensible, precisa y confiable para medir las capas retinianas. El rendimiento de esta herramienta se probó en conjuntos de datos de calibración, capacitación y validación con un rendimiento adecuado y confiable. En conjunto, los resultados validan ThicknessTool para la medición automatizada del espesor de la retina. Especulamos que esta herramienta puede ayudar a lograr mediciones

precisas, precisas y rápidas en imágenes digitales multi-plataforma.

En el desarrollo de nuevas herramientas de medición, las imágenes de calibración son necesarias para ajustar el rendimiento de los algoritmos de procesamiento de imágenes a un estándar.[13] Diseñamos imágenes representativas con varios espesores conocidos desde 10 píxeles (dado el tamaño de un núcleo celular en nuestras imágenes) hasta 500 píxeles (el doble de la ONL medida en nuestra base de datos). Además, creamos un conjunto de imágenes con múltiples rotaciones para imitar las orientaciones de la ONL. Este paso es necesario ya que los calibres de TT se construyen de manera vectorial, en número y ángulo según lo requiera el usuario o la región de interés, respectivamente. Además, evaluamos la capacidad del TT para detectar el verdadero adelgazamiento y engrosamiento en una imagen simulada con grosores decrecientes y crecientes, respectivamente. Descubrimos que el TT puede medir el espesor de manera automática y precisa cuando se lo desafía con diferentes espesores y rotaciones.

La evaluación precisa de la muerte de fotorreceptores depende actualmente del conteo manual por parte de observadores enmascarados. Idealmente, dos observadores enmascarados cuantifican el grosor de la retina, y su promedio se usa para determinar el resultado general. Descubrimos que las mediciones de los observadores individuales son propensas a la sobre y subestimación, independientemente de su experiencia previa. Además, los observadores dibujan calibres con un ángulo incorrecto, lo cual puede contribuir indirectamente al sesgo mencionado anteriormente. Además, los calibres no se ubicaron en ubicaciones correspondientes en los análisis repetidos de la misma imagen, lo que reduce la reproducibilidad de las mediciones. En conjunto, esto conduce a una falta de cuantificación precisa o reproducible del grosor de la retina. Este es de gran importancia, ya que este sesgo puede conducir a errores α o β. Por ejemplo, no encontrar diferencias en el grosor de la retina en los casos en que un tratamiento neuroprotector es efectivo, o incluso peor, encontrar una diferencia significativa en el grosor de la retina en un tratamiento ineficaz.[14,15] Tales errores tienen un elevado costo para la comunidad científica.

Independientemente del rendimiento de las mediciones manuales, hasta la fecha, sigue

siendo el estándar de referencia. A fin de determinar el acuerdo entre los observadores y TT, nos abstuvimos de utilizar exclusivamente la comparación o correlación, como Bland y Altman han descrito previamente en sus trabajos.[11,16] Como describen los autores, dicho acuerdo debe cuantificarse y no puede evaluarse mediante comparación de medias o pruebas de hipótesis, ya que el acuerdo no existe en forma binaria presente o ausente.[12] Además, una alta correlación no implica necesariamente un alto acuerdo entre dos métodos, ya este método cuantifica el grado en que dos métodos están relacionados.[17] El análisis del acuerdo mostró un sesgo menor de TT contra los observadores que el promedio de los observadores entre sí, tanto para ONL como para INL. Teniendo en cuenta la capacidad de TT para detectar el verdadero adelgazamiento y engrosamiento de las capas, colectivamente, estos resultados confieren a TT un mayor acuerdo en la medición.

Además, el análisis de espesor automatizado por TT produce una serie de parámetros que, de lo contrario, permanecen ocultos en las mediciones manuales. En esta nota, Byun et al. han desarrollado métodos pioneros para estimar el grosor exclusivamente de la ONL.[9] Sin embargo, su enfoque no fue validado por análisis de acuerdo de espesor con observadores manuales. En su trabajo, los autores también describieron un método para cuantificar las distorsiones estructurales en el tejido retiniano antes y después del desprendimiento de retina, también conocido como índice de distorsión. En esencia, los autores describen que los cambios estructurales en la ONL son directamente proporcionales al índice de distorsión. En nuestro trabajo, estimamos la distorsión general en un enfoque similar utilizando el índice de dispersión (ID), calculado como la relación varianza-media, como definieron previamente Cox y Lewis.[18] De manera similar, el índice de dispersión permite la evaluación de la varianza del calibre, y puede proporcionar datos de la arquitectura de capa. Brevemente, un ID de 0 representa ausencia de dispersión; por lo tanto, utilizamos la ID para fines de calibración, ya que las imágenes simuladas de espesor conocido y constante eran rectángulos uniformes, ergo ID = 0. Especulamos que el índice de dispersión se puede utilizar de forma similar para una evaluación morfométrica adicional.

Se deben considerar varias limitaciones para este trabajo. Primero, a pesar del análisis de calibración realizado, la validación de TT solo se puede realizar con mediciones manuales. Este proceso es difícil, ya que las mediciones de los observadores conllevan un sesgo intrínseco. Sin embargo, los observadores mostraron un rendimiento comparable, con intervalos de confianza superpuestos. En segundo lugar, el análisis TT debe realizarse en áreas distantes de los bordes de la imagen, ya que el esqueleto se recorta para evitar mediciones al extremo de la imagen. A pesar de que TT permite el corte de esqueletos y bordes, los usuarios deben evaluar las imágenes procesadas para un control cualitativo. En tercer lugar, TT mide el grosor en máscaras segmentadas binarias. Por lo tanto, su rendimiento es tan preciso como el algoritmo de segmentación utilizado. En nuestro caso, creamos esta herramienta basada en un algoritmo previamente validado para la segmentación de la capa de retina.[4] Recomendamos a los usuarios que primero validen su algoritmo de segmentación para un rendimiento óptimo. Finalmente, independientemente de la posible aplicación en las distintas modalidades de imagen, esta herramienta se ha validado solamente para mediciones de ONL e INL en crio-secciones de retina de ratones teñidas con 4 ',6-diamidino-2-fenilindol (DAPI). Confiamos que en el futuro se pueda validar a TT en plataformas de imágenes.

En resumen, desarrollamos una herramienta rápida, precisa y confiable para medir el grosor de la retina con capacidades multi-plataforma. Además, el TT se puede personalizar según las preferencias definidas por el usuario. Esto incluye el número de calibres con una distancia mínima de 1 píxel entre si, recorte de esqueleto y bordes, y segmentación de capa automática o manual. Además, dado el uso potencial con capacidades multi-plataforma, el código fuente puede modificarse fácilmente para adaptarse a diferentes aplicaciones. Finalmente, esta herramienta está disponible gratuitamente para su descarga y uso. Esperamos que esta herramienta de medición pueda mejorar los resultados y reducir el sesgo en la investigación de las enfermedades degenerativas de la retina.

Chapter V

Photoreceptor Cell Death: Replacement Therapy

Necrotic Cell Death Pathway in Both Donor Photoreceptor and Host Immune Cells Synergize To Affect Photoreceptor Precursors Transplant Survival

Daniel E. Maidana, Lucia Gonzalez-Buendia, Josep Maria Caminal Mitjana, Joan W. Miller, Demetrios G. Vavvas*

Paper to be submitted for publication at *Proceedings of the National Academy of Sciences (PNAS)*

Chapter V. Photoreceptor Replacement Therapy

5.1. Abstract

Photoreceptor transplant has been put forward as a repair strategy to tackle degenerating or already degenerated retinas. Nonetheless, cell death and immune rejection seriously limit the success of this strategy with only a small fraction of transplanted cells surviving. Improving survival of transplanted cells is of significant importance. Recent evidence has identified receptor-interacting protein kinase 3 (RIPK3) as a molecular trigger controlling necroptotic cell death and inflammation. However, its role in photoreceptor transplantation and regenerative medicine has not been studied. We hypothesized that modulation of RIPK3 to address both cell death and immunity could have advantageous effects on photoreceptor survival. In a model of inherited retinal degeneration, deletion of RIPK3 in donor photoreceptor precursors significantly increases survival of transplanted cells. Simultaneous RIPK3 deletion in donor photoreceptors and recipients maximizes graft survival. Lastly, to discern the role of RIPK3 in the host immune response, bone marrow transplant experiments demonstrated that peripheral immune cell RIPK3 deficiency is protective for both donor and host photoreceptor survival. Interestingly, this finding is independent of photoreceptor transplantation, as the peripheral protective effect is also observed in a retinal detachment model. Altogether, these results indicate that immunomodulatory and neuroprotective strategies targeting the RIPK3 pathway can aid regenerative therapies of photoreceptor transplantation.

5.2. Introduction

Retinal degenerations are a heterogeneous group of diseases, including age-related macular degeneration (AMD) and retinitis pigmentosa (RP), which are characterized by photoreceptor cell death and severe vision loss affecting millions of people without effective therapy.(1–4) Several emerging therapeutic strategies including drug delivery,

gene therapy, and photoreceptor transplant are being investigated with limited clinical success. Gene therapy has shown the most promise, but this approach requires targeted cells to still be viable in the native retina in order to transduce the gene of interest.(5) This results in a narrow therapeutic window for retinal degenerative diseases. To overcome this limitation, photoreceptor transplant has been put forward as a repair strategy to tackle degenerating or already degenerated retinas. Unfortunately, efforts to achieve successful photoreceptor transplantation strategies have been limited by extremely limited efficiency.(6–8) Thus, it seems crucial to reflect and re-investigate the conditions necessary to fine tune photoreceptor transplantation.

Cell death and immune rejection have been a major bottle-neck in organ and cellular transplantation.(9, 10) Therefore, it is reasonable to assume that targeting these two entangled variables could improve overall outcomes through promoting graft survival and reducing immune rejection. Recent work has highlighted the important role receptor-interacting protein kinase 3 (RIP3/RIPK3) in necrotic cell death (necroptosis) and inflammation.(11–17) This isoform is a member of the RIP Kinases family which drives the necroptosis pathway triggered by TNF primarily thorough receptors TLR3 and TLR4, among others.(18) RIPK3 not only mediates programmed necrotic cell death, but has also shown necroptosis-independent functions such as NLRP3 inflammasome activation.(19–21) Furthermore, solid organ transplantation studies have shown that allograft rejection can be mediated by RIPK3 necroptosis and inflammation.(22, 23) Despite this encouraging evidence, it remains unknown whether selective targeting of the RIPK3 pathway could enhance the outcome of single-cell photoreceptor transplant to degenerated retinas in the immune privilege site of the eye.

In this work, we investigated the role of RIPK3 pathway in the survival of photoreceptor transplants in a model of inherited retinal degeneration. Using bone marrow transplant chimeras, we dissected the role of RIP Kinases in the immune cells. We observed that combination of RIPK3 deficiency in donor photoreceptors and recipients maximizes graft survival. Moreover, we confirmed with bone marrow transplant experiments that peripheral immune cell RIPK3 deficiency is protective for photoreceptor survival.

Furthermore, the neuroprotective effects of RIPK3 can be replicated in part by pharmacologic inhibition of RIP kinase. Altogether, these results indicate that neuroprotective strategies in photoreceptor transplantation can aid in restoration therapies.

5.3. Materials and Methods

All reagents used in this work are listed in Supplementary Table 1 with their respective manufacturer and catalogue number.

Reagent	Manufacturer	Catalog #	Notes
Papain Dissociation System	Worthington	LK003150	-
MACS Buffer	Miltenyi Biotec	130-091-221	-
MACS BSA Stock Solution	Miltenyi Biotec	130-091-376	-
Rat Anti-Mouse CD73 antibody	BD Biosciences	550738	-
Goat Anti-Rat IgG MicroBeads	Miltenyi Biotec	130-048-502	-
LS Columns	Miltenyi Biotec	130-042-401	-
MACS Pre-separation Filter	Miltenyi Biotec	130-041-407	-
Midi MACS Manual Separator	Miltenyi Biotec	130-042-301	-
DMEM/F-12, HEPES, no phenol red	ThermoFisher	11039021	-
Penicillin-Streptomycin	ThermoFisher	15140122	-
N-2 Supplement	ThermoFisher	17502048	-
Fetal Bovine Serum	ThermoFisher	16140071	-
D-(+)-Glucose	Sigma-Aldrich	G7021	-
CytoTox 96® Cytotoxicity Assay	Promega	G1780	-
2,2,2-tribromoethanol	Sigma-Aldrich	152463	-
2- methyl-2-butanol	Sigma-Aldrich	T48402	-
5% Phenylephrine, 0.5% Tropicamide	MEEI Pharmacy	-	-
Proparacaine hydrochloride 0.5%	Akorn	NDC 17478-263-12	-
Hamilton Syringe 701 RN SYR	Hamilton	7635-01	-
Bacitracin ointment	Perrigo	NDC 0574-4022-35	-
Paraformaldehyde solution 4%	Santa Cruz	30525-89-4	-
Phosphate Buffered Saline (1X)	Lonza	17-516Q	-
Sucrose	Sigma-Aldrich	S8501	-
Tissue-Tek* O.C.T. Compound	Sakura	25608-930	-
Triton™ X-100	Sigma-Aldrich	X100	-
Bovine Serum Albumin	BIOWorld	9048-46-8	-
Anti-Ribeye	BD Biosciences	612044	Dilution 1:200
PNA-Lectin	ThermoFisher	L-32460	Dilution 1:50
Anti-RPE65	Novus Biologicals	NB100-355SS	Dilution 1:50
Anti-Rhodopsin	Milipore	MAB5356	Dilution 1:100
Anti-Brn3a	Santa Cruz	sc-31984	Dilution 1:100
Anti-PKCα	Cell Signaling	2056	Dilution 1:50
Anti-GFAP	DAKO	Z0334	Dilution 1:200
Isolectin B4	ThermoFisher	I32450	Dilution 1:100
Alexa Fluor 647 Goat anti-Rat	ThermoFisher	A-21247	Dilution 1:500
Alexa Fluor 647 Donkey anti-Mouse	ThermoFisher	A-31571	Dilution 1:500
Alexa Fluor 647 Donkey anti-Rabbit	ThermoFisher	A-31573	Dilution 1:500
Alexa Fluor 647 Donkey anti-Goat	ThermoFisher	A-21447	Dilution 1:500
DAPI	Roche	10236276001	-
Fluoromount-G®	SouthernBiotech	0100-01	-

Reagent	Manufacturer	Catalog #	Notes
T-PER™ Tissue Protein Extraction Reagent	ThermoFisher	78510	-
Halt™ Protease and Phosphatase Inhibitor	ThermoFisher	78440	-
TissueRuptor II	QIAGEN	9002755	-
TissueRuptor Disposable Probes (25)	QIAGEN	990890	-
Pierce™ Coomassie Protein Assay Kit	ThermoFisher	23200	-
Proteome Profiler Mouse Cytokine Array	R&D Systems	ARY006	-
RBC Lysing Buffer	Leinco	R1364	-
EDTA	ThermoFisher	AM9261	-
Sodium Azide	Sigma-Aldrich	71289	-
Osmotic Pumps	Alzet	1004	-
Necrostatin-1	Sigma-Aldrich	N9037	-
DMSO	ATCC	ATCC® 4-X™	-
(2-Hydroxypropyl)-β-cyclodextrin	Sigma-Aldrich	C0926	-
Polyethylene Glycol 400	Sigma-Aldrich	91893	-

Supplementary Table 1. Reagents, manufacturers, and corresponding catalog numbers used.

5.3.1. Animals and Breeding

All animal used in experiments and breeding adhered to the statement of the Association for Research in Vision and Ophthalmology (ARVO). Animal protocols were reviewed and approved by the Animal Care Committee of the Massachusetts Eye and Ear Infirmary. Nrl-EGFP mice were kindly provided by Anand Swaroop, Ph.D. (National Eye Institute, Bethesda, MD). Rip3−/− (RIP3KO) mice were provided by Vishva M. Dixit (Genentech, San Francisco, CA). We backcrossed the RIP3KO mice to C57BL/6J. We confirmed Rip3−/− and Crb1+/+ in the offspring through a commercial vendor (Transnetyx, Cordova, TN). Results were confirmed in-house with PCR. We subsequently bred the RIP3KO.6J for at least 10 generations before experiments. C57BL/6J, C57BL/6-Tg(CAG-EGFP)131Osb/LeySopJ (EGFP), and Pde6βrd10/rd10 (Pde6brd10) were purchased from The Jackson Laboratories. Pde6brd10 mice were crossed with RIP3KO mice to generate the Pde6brd10/rd10 Rip3 -/- (Pde6brd10

RIP3KO). EGFP mice were crossed with RIP3KO mice to generate the C57BL/6-Tg(CAG-EGFP)131Osb/LeySopJ Rip3 -/- (RIP3KO EGFP). All strains mice were genotyped at weaning by a commercial vendor (Transnetyx, Cordova, TN) to confirm strains and absence of Crb1rd8 mutation.

5.3.2. Photoreceptor Isolation and Sorting

Murine photoreceptors were isolated from selected donor strains using the Papain Dissociation System following manufacturer's instructions and published literature.(24) Briefly, pups (P4-7) were euthanized by decapitation and eyes were enucleated and placed in in ice-cold phosphate-buffered saline (PBS). Retina was dissected and placed in Eppendorf tubes containing 1 mL Papain previously warmed at 37º. Retinas were incubated for 45-60 minutes in a shaker at 37ºC at 700 rpm. After incubation, partially digested retinas were resuspended with dissolved DNase solution, Earle's Balanced Salt Solution (EBSS), and Ovomucoid mixture. Finally, the dissociated retina was filtered using a 40 µm strainer, and elution was spun at 1100 rpm for 4 minutes at room temperature. For MACS, procedure was performed following manufacturer's instructions and published literature.(25) Pellet was resuspended in 500 µL of MACS Buffer containing 0.5% bovine serum albumin (BSA). Up to 5 x 107 cells were incubated for 5 minutes on ice with 10 µL of Rat Anti-Mouse CD73 antibody at a final concentration of 10 µg/mL. Following, MACS buffer was added to the suspension and spun at 1100 rpm for 4 minutes at room temperature. Pellets were then incubated for 15 minutes on ice with Goat Anti-Rat IgG MicroBeads. MACS buffer was added to the suspension and spun at 1100 rpm for 4 minutes at room temperature. Suspension was loaded to a LS Column with a Pre-separation Filter on a Midi MACS Manual Separator. The initial flow-through (CD73-negative cells) was discarded. CD73-positive cells were collected after removal of the LS Column from the Midi MACS Manual Separator. Sorted cells were centrifuge tubes at 1100 rpm for 4 minutes at room temperature and resuspended in retina media (see below) at a concentration of 2 x 105 cells/µL and kept on ice.

5.3.3. Photoreceptor Culture and LDH Assay

CD73-positive photoreceptors were cultured in retina media containing phenol-free DMEM:F12, 1% FBS, N2 Supplement, HEPES 5 mM, 6 mg/mL Glucose, 100 IU/mL Penicillin, and 100 µg/mL Streptomycin, as previously described.(26) After preparation, media was filtered using a 0.2 µm vacuum filter and stored at -20ºC until used. A total of 6 x 105 CD73-positive photoreceptors was plated in 6-well plates and cultured at 37ºC with 5% CO2. Images were obtained at different timepoints with a ZEISS Primovert Cell Microscope (Thornwood, NY). Cells were counted automatedly using ImageJ (version 1.49, National Institutes of Health, Bethesda, MD, USA) from the acquired images.

LDH Assay was performed using the CytoTox 96® Non-Radioactive Cytotoxicity Assay according to manufacturer's instructions. Briefly, 50 µL aliquots from wells were transferred to an empty 96-well clear bottom plate. A volume of 50 µL of the CytoTox 96® Reagent was added to each sample aliquot. Plate was covered and incubated for 30 minutes at RT. A volume of 50 µL of Stop Solution was added to each well and absorbance was recorded at 490 nm within 1 hour. Culture media background was subtracted from the obtained absorbance values.

5.3.4. Photoreceptor Subretinal Transplantation

Recipients (P30) were anesthetized with an intraperitoneal injection of a mixture of 2,2,2-tribromoethanol and 2- methyl-2-butanol at a dose of 125 mg/kg. Pupils were dilated with 5% phenylephrine and 0.5% tropicamide, and proparacaine hydrochloride 0.5% ophthalmic solution was used as topical ocular anesthesia. Briefly, animals were placed on a heat pad under the surgical microscope. A nasal sclerotomy was performed to gain access to the vitreous cavity. A manual Hamilton microliter injector with a 34-gauge needle was inserted into the vitreous. Subretinal injection was performed under direct visualization in the temporal retina. A total of 400,000 CD73-postive EGFP-positive photoreceptors were transplanted, as previously described.(25) Eyes with hemorrhage, leakage, or cataract were excluded from further analysis. A topical ointment containing bacitracin was applied after the procedure.

5.3.5. Immunohistochemistry and Imaging

At given timepoints, animals were euthanized, and eyes were enucleated and placed in ice-cold PBS. Eyes were fixed with 4% PFA and the eyecup was dissected from the anterior segment and lens. Eyes were subsequently washed in PBS and cryoprotected with 30% sucrose prior to embedding in Tissue-Tek Optimal Cutting Temperature compound. The entire eyecups were sectioned to 25 μm cryosections and placed on 4 intercalated microscope slides and stored at -80ºC until used.

For immunofluorescence staining, slides were blocked for 30 minutes at 4ºC with PBS containing 2% BSA and 0.05% Triton-X. Slides were washed with blocking buffer diluted 1:1 with PBS. Primary antibodies were diluted in blocking buffer and incubated overnight at 4ºC. Slides were subsequently washed with PBS containing 1% BSA and 0.1% Tween-20 (PBST). Secondary antibodies diluted in blocking buffer were incubated for 2 hours at room temperature. Slides were washed with PBST, stained with DAPI, and mounted with Fluoromount-G mounting medium.

To assess photoreceptor survival, EGFP-positive transplanted photoreceptors were visualized in the subretinal space under a ZEISS Axio (Thornwood, NY) epifluorescence microscope and counted by a masked observer. Total integrated cells per eye were calculated as the product of the counted EGFP-positive cells per slide and the number of slides sectioned per eye (4). Percentage was calculated over a total 400,000 transplanted CD73-postive EGFP-positive photoreceptors. Laser confocal images were acquired with a ZEISS LSM 800 (Thornwood, NY) microscope.

5.3.6. Cytokine Expression

To quantitate cytokine expression, lysates were obtained from dissected retinas and RPE/Choroid. Briefly, tissues were lysed with T T-PER™ Tissue Protein Extraction reagent containing Halt Protease Inhibitor and EDTA. Tissues were homogenized for 1 minute in a 15 mL Falcon tube with TissueRuptor II using disposable probes. Samples were subsequently placed on ice and sonicated twice for 5 seconds at 20% ultrasound power. Tubes were centrifuged at 14,000 × g for 15 minutes at 4ºC. Supernatants were

collected and stored for further analysis. Protein concentration was assessed with Coomassie Protein Assay kit and equal total protein concentrations were used for comparisons. Cytokine proteome profile was assessed with Proteome Profiler Mouse Cytokine Array Kit, Panel A according to manufacturer's instructions. Background signal was subtracted from each cytokine and fold change ratios were calculated from wild-type expression.

5.3.7. Bone Marrow Isolation and Transplantation

Recipient animals (P21-28) were conditioned with Busulfan as previously described. (27, 28)Animals received an intraperitoneal injection of 35 mg/kg busulfan 7, 5, and 3 days before bone marrow transplantation (BMT). On the day of BMT, bone marrow EGFP-positive donors were euthanized and their bone marrow was isolated as previously described.(29) A total 2 x 107 bone marrow progenitors were injected to recipients via tail vein in 100 µL PBS with a 27G needle. Four weeks after BMT, mice were assessed for immune reconstitution. For this purpose, blood was collected through facial vein venipuncture and lysed with RBC Lysis Buffer. Samples were spun at 1,300 rpm for 5 minutes at room temperature. Following washes with PBS, 2% BSA, 1 mM EDTA, and 0.05% sodium azide, samples were incubated with 2% PFA for 10 minutes on ice. Samples were washed and kept for same-day FACS analysis with a BD LSR II analyzer.

5.3.8. Pharmacologic RIPK Inhibition

Receptor-interacting protein kinase inhibition was performed pharmacologically using Necrostatin-1, a small molecule inhibitor of RIP1 Kinase. For intravitreal delivery, Necrostatin-1 was dissolved in 15% Cyclodextrin-ß and DMSO and added to the CD73-positive EGFP-positive photoreceptors prior to subretinal transplantation, at a final concentration of 300 µM. Control animals received the same vehicle. For systemic delivery, osmotic pumps were loaded with Necrostatin-1 dissolved in DMSO:PEG400 with a target dose of 15 mg/kg/day, as previously described.(12) Pumps were surgically placed in the subcutaneous interscapular region of recipients. Control animals received the same vehicle. The presence of Necrostatin-1 in the systemic circulation at 4-week

endpoint was confirmed through facial venipuncture and subsequent HPLC analysis.

5.3.9. Statistical Analyses

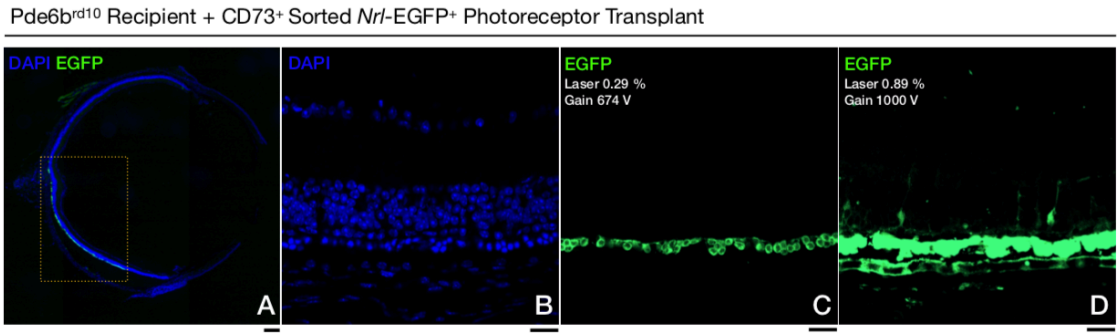
Statistical analyses were performed with SAS Software (2016, SAS Institute Inc., Cary, NC). Group comparison and statistical tests used are described accordingly. A p value of ≤ 0.05 was considered statistically significant. Volcano plots were constructed using a cut-off p value ≤ 0.05 for statistical significance and a $\geq$ 3-fold change in cytokine expression.

5.4. Results

5.4.1. Transplanted CD73+ Nrl-EGFP+ Photoreceptors Precursors Repopulate the ONL in Pde6brd10 Mice

Considering recent evidence of artefacts in photoreceptor transplant integration, we first sought to tackle any confounder of cytoplasm transfer. We chose the Pde6brd10 mouse model of retinal degeneration which bears a profound photoreceptor loss at P30, resulting in a single line of outer nuclear photoreceptor layer (ONL) at endpoint.(30–32) Given a four-week period necessary to evaluate graft survival, any additional EGFP-positive outer nuclear photoreceptor line in the recipient's ONL cannot be due cytoplasmic transfer. EGFP+ photoreceptor precursor donors were isolated from P4-7 Nrl-EGFP+ mice and positively selected using an anti-CD73 antibody and magnetic-activated cell sorting (MACS). The CD73 antigen, also known as ecto-5'-nucleotidase, is a cell surface molecule expressed in photoreceptor precursors.(33) MACS CD73-positive sorting has been proven to be effective to isolate a population of transplantable photoreceptor precursors.(34–37) Following MACS, we transplanted the CD73-positive photoreceptor precursors into the subretinal space of P30 Pde6brd10 mice. We observed that after four weeks, transplanted EGFP-positive CD73-positive MAC-sorted rod photoreceptor precursors repopulate the native ONL of Pde6brd10 mice (Supplementary Figure 1). Interestingly, we observed that with adequate imaging technique and laser exposure

levels, there was an appropriate visualization of grafted cells with no signs of background EGFP+ expression in the host retina. Interestingly cytoplasmic transfer in the inner nuclear layer (INL) could be seen with overexposed settings, consistent with previous publications.(38, 39)



Supplementary Figure 1. Photoreceptor Transplantation Repopulates the Outer Nuclear Layer. (A) Transplanted Nrl-EGFP-positive CD73-positive MAC-sorted rod photoreceptor precursors repopulate the native outer nuclear layer of Pde6brd10 mice after subretinal delivery (dotted box). (B-C) Selected area showing transplanted cells visualized under normal EGFP exposure settings, without saturated pixel values. (D) Overexposed settings display a sub retinal EGFP mass without distinguishable cell features and a spurious signal which could represent cytoplasmic transfer. Representative images of ≥ 3 independent experiments. Scale bar 100 µm (A), and 30 µm (B-D).

Disruption of outer limiting membrane (OLM) may enhance photoreceptor integration, even in wild-type animals.(40, 41) For this purpose, we transplanted Nrl-EGFP-positive CD73-positive rod photoreceptor precursors to the subretinal space of animals bearing the rd8 mutation in the Crb1 gene, hence with a disrupted OLM. We observed that transplanted photoreceptors to Crb1rd8 recipients failed to contact the INL or remaining ONL, most probably given the presence of the latter outer and inner segments in the host retina (Figure 1). Given the broad distance between graft and ONL/INL, synapsis were not only absent, but unlikely to occur. In contrast, grafted photoreceptors in the Pde6brd10 recipients showed close contact with the recipient INL/ONL and displayed synaptic ribbons with the host INL. Collectively, these results suggest that grafted photoreceptors require a pre-existing ONL degeneration to facilitate synapsis with the

recipient's INL or ONL. This phenomenon appears to occur regardless of the initial integrity of the OLM, which may already be disrupted in later stages of retinal degeneration.(42, 43)

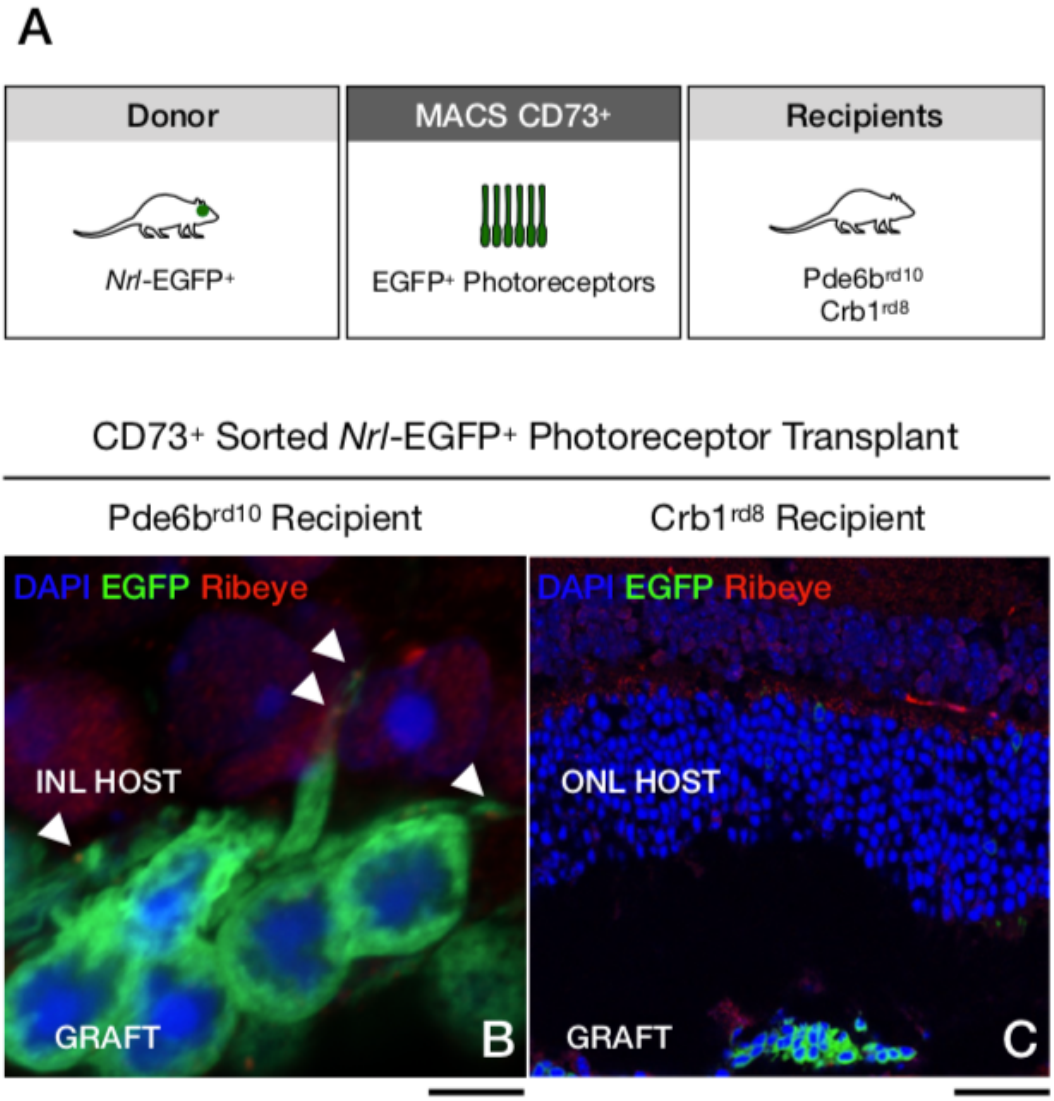
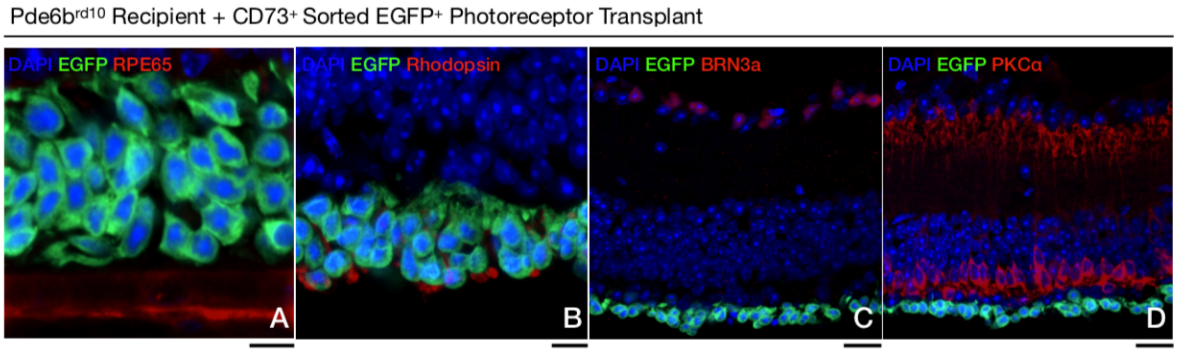


Figure 1. Graft-Host Synapsis is Facilitated by Pre-existent Retinal Degeneration.
(A) Outline of photoreceptor donors, sorting, and recipients used for the experiments. (B) Grafted photoreceptors in the Pde6brd10 recipients showing close contact with the recipient INL/ONL with synaptic ribbons with the host INL (arrows). (C) Grafted photoreceptors to Crb1rd8 recipients failed to contact the INL or remaining ONL (n = 7-8). Scale bar 5 µm (B) and 40 µm (C).

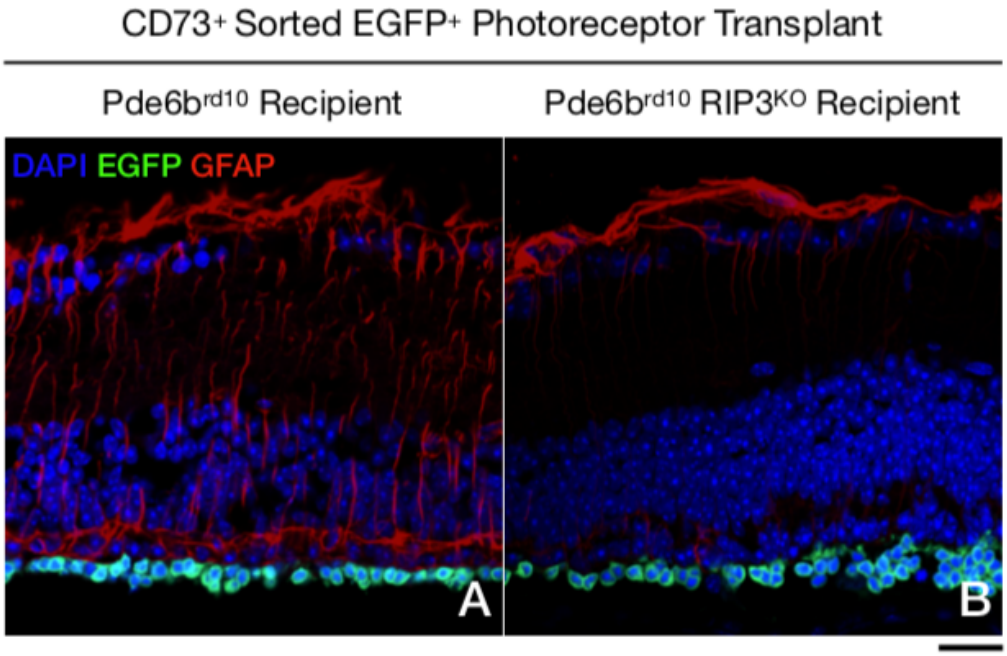
5.4.2. Photoreceptor RIPK3 Deletion Increases the Survival of Grafted Photoreceptor Precursors in Pde6brd10 Mice

RIPK3 is a key regulator of the necroptotic cell death pathway. Inasmuch, modulation of this pathway has been shown to increase cone survival in models of inherited retinal degeneration(12) and photoreceptors in experimental murine retinal detachment.(13) Moreover, given the potential role of cell death pathways in photoreceptor transplantation(44), we sought to examine whether RIPK3 modulation could enhance photoreceptor graft survival after subretinal implantation. Unfortunately, the crossing of Nrl-EGFP and RIP3KO mice was unsuccessful. To address this shortcoming, we bred EGFP-positive mice with RIP3KO and generated EGFP-positive RIPK3-deficient photoreceptors precursors. To positively select photoreceptors, we used the same CD73-positive MACS strategy as described above. Finally, the phenotype of sorted and transplanted EGFP-positive CD73-positive photoreceptor precursors was confirmed by lack of reactivity to RPE65, BRN3a, PKC-1α, and GFAP (Supplementary Figures 2 and 3). In addition, we did not observe any EGFP-negative cells expressing reactivity to these markers in the grafted area, which suggests that CD73-positive MAC-sorting can effectively enrich isolated rod photoreceptor precursors. Taken together, these results suggest that grafted EGFP-positive photoreceptor precursors repopulate the ONL and do not represent or de-differentiate to retinal pigment epithelial cells (RPE), ganglion cells, bipolar cells, astrocytes or Muller cells.



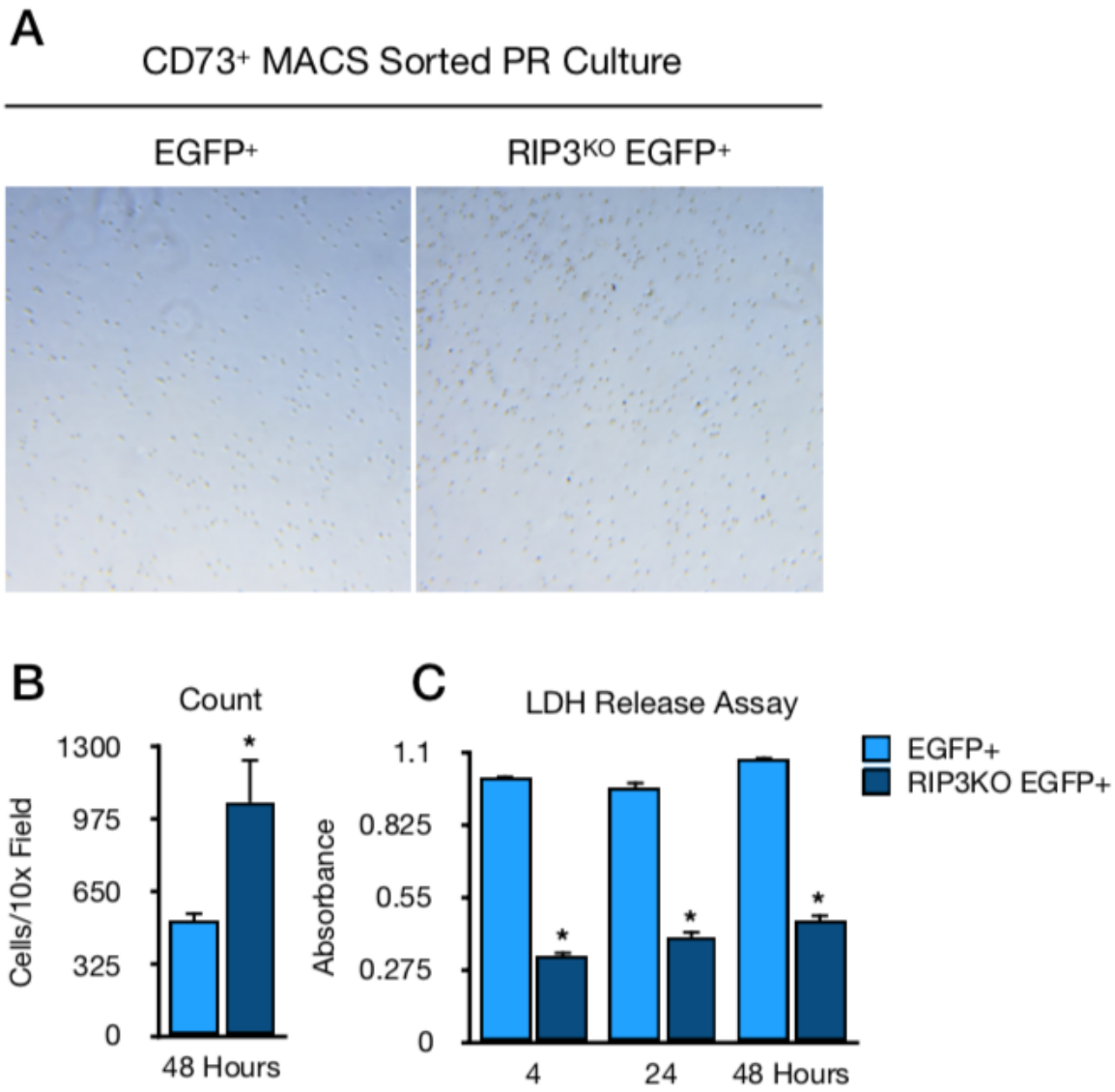
Supplementary Figure 2. Grafted EGFP-positive Photoreceptor Precursors Repopulate the ONL and Do Not De-Differentiate to Other Cell Types.

(A-D) The phenotype of sorted and transplanted EGFP-positive CD73-positive photoreceptor precursors was confirmed by lack of expression of RPE65 (A); BRN3a (C); and PKCα (D), and positive expression of Rhodopsin (B). Representative images of ≥ 3 independent experiments. Scale bar 10 µm (A-B) and 20 µm (C-D).



Supplementary Figure 3. RIPK3 Deletion Displays Reduced Retinal Gliosis After Transplantation. (A-B) Pde6brd10 RIP3KO recipients showed significantly lower gliosis in either RIP3WT or RIP3KO EGFP-positive transplants, when compared to Pde6brd10 RIP3WT recipients. Representative images of ≥ 3 independent experiments. Scale bar 20 µm (A-B).

Subsequently, we transplanted EGFP-positive (RIP3WT) or RIP3KO EGFP-positive rod precursors to Pde6brd10 recipients. We noticed that RIPK3 deletion in photoreceptor donors significantly increased the survival of these precursor cells in vitro (Supplementary Figure 4) and the overall graft survival at four weeks (4,152.44 ± 538.44 cells) when compared to RIP3WT donors (1,043.00 ± 238.37 cells, p = .009) (Figure 2B). Collectively, these results indicate that RIPK3 deletion in donor photoreceptor precursors is sufficient to enhance graft survival in Pde6brd10 recipients.



Supplementary Figure 4. RIPK3 Deletion in Donor Photoreceptor Enhances Survival.
(A) Isolated photoreceptors from EGFP RIP3WT and EGFP RIP3KO mice cultured for 48 hours. (B) Mean photoreceptor survival measured by quantitation of cells after timepoint. (C) LDH Assay performed using a non-radioactive cytotoxicity assay showing significant dampening of LDH release in RIP3KO EGFP cultured photoreceptors versus controls. Representative images of ≥ 3 independent experiments.

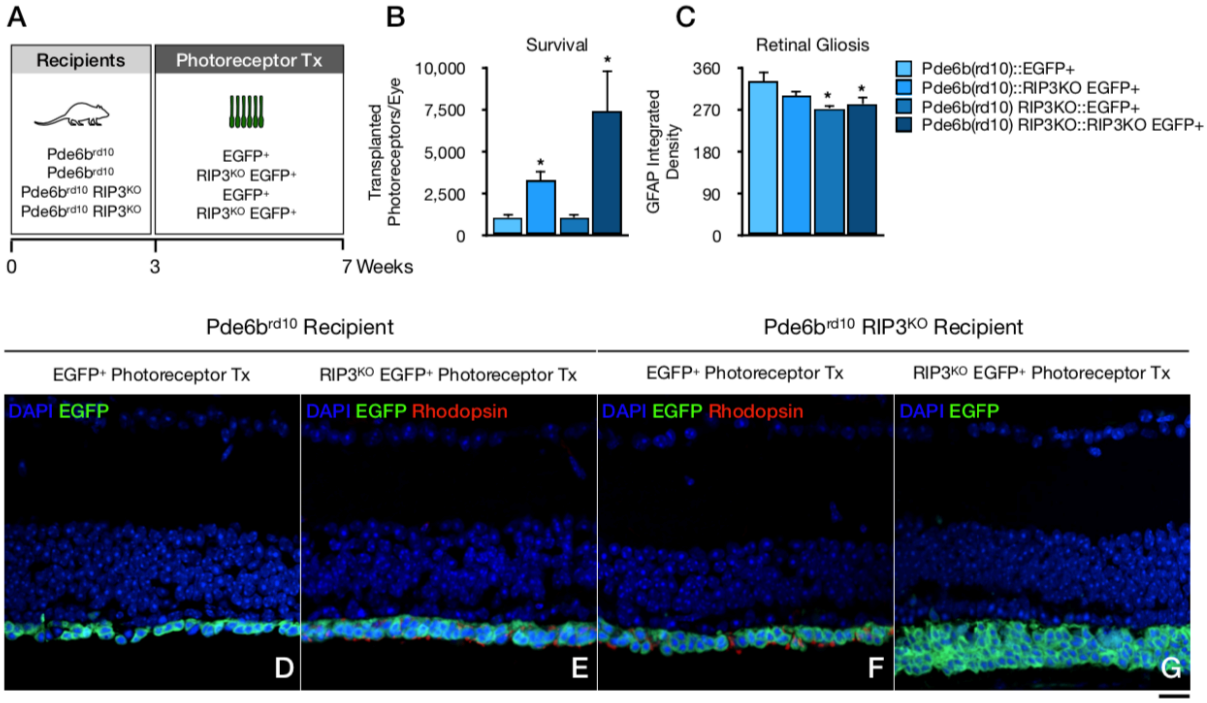


Figure 2. Recipient and Photoreceptor RIPK3 Deletion Maximizes the Survival of Grafted Photoreceptor Precursors.

(A) Outline of photoreceptor donors and recipients used for the experiments. (B) Mean photoreceptor survival measured by quantitation of transplanted photoreceptors/eye. (C) Mean retinal gliosis measured by quantitation of GFAP integrated density. (D-G) Photoreceptor transplantation of EGFP RIP3WT and EGFP RIP3KO photoreceptors to Pde6brd10 RIP3WT and Pde6brd10 RIP3KO recipients (n = 4-9). Scale bar 20 µm (D-G).

5.4.3. Recipient and Photoreceptor RIPK3 Deletion Maximizes the Survival of Grafted Photoreceptor Precursors

Our group has previously demonstrated that RIPK3 deletion in Pde6brd10 mice increase cone survival and display less microglial infiltration compared with wild-type controls at P42 and P56.(12) Taken together with the abovementioned results, we hypothesized that simultaneous donor and recipient RIPK3 kinase modulation could improve photoreceptor transplant graft survival. For this purpose, we used Pde6brd10 RIP3KO mice to determine whether recipient deletion of RIPK3 could influence graft survival. In this Pde6brd10 RIP3KO recipient group, RIPK3 deletion in photoreceptor donors maximized graft

survival (7,425.29 ± 2,415.23) when compared to RIP3WT donors (1,004.00 ± 199.67, p = .005) (Figure 2B). These results suggest that RIPK3 deletion in the recipient is not a sufficient condition to rescue wild-type donor photoreceptors yet appears to be necessary to maximize the overall RIP3KO photoreceptor survival.

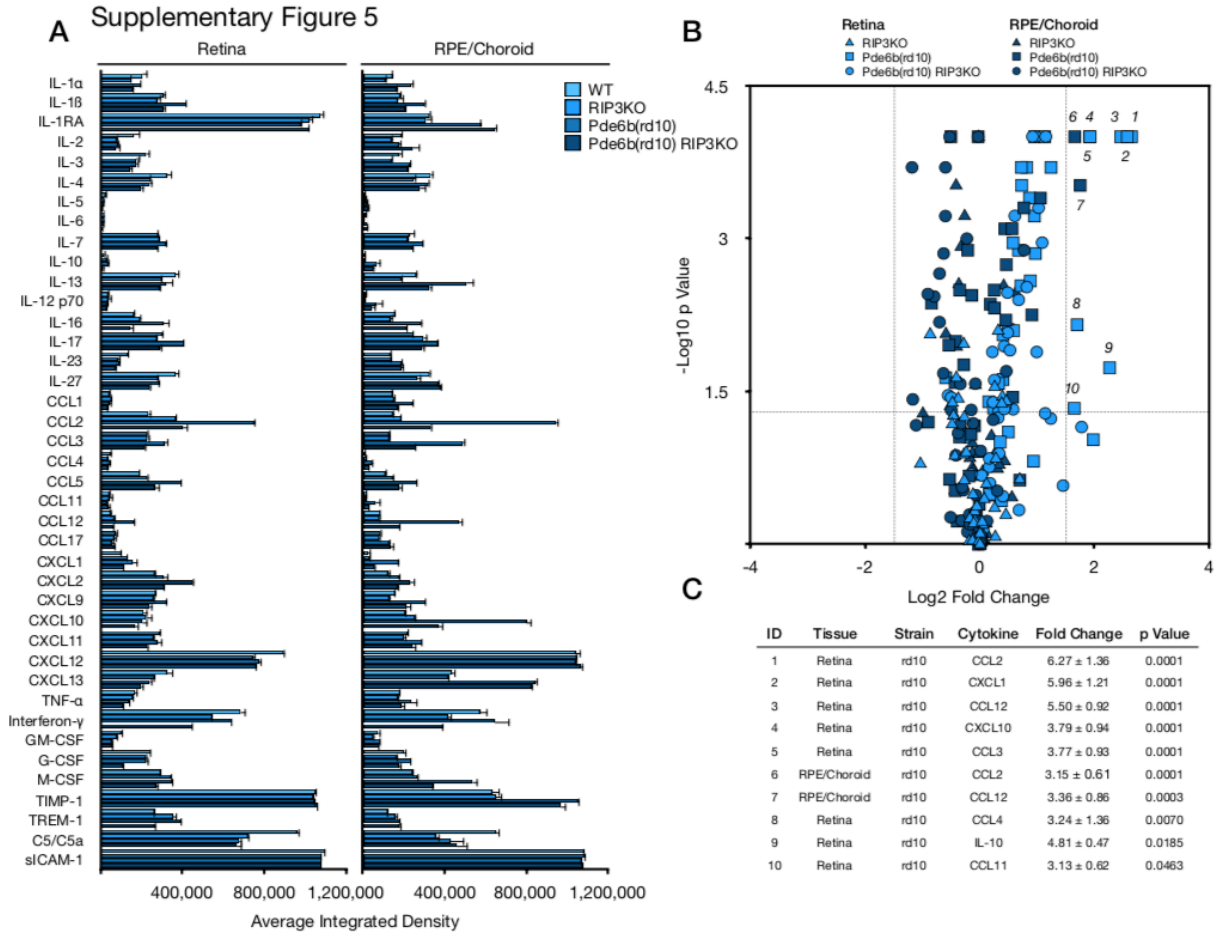
5.4.4. Recipient RIPK3 Deletion Reduces the Retinal Gliosis After Photoreceptor Precursor Transplantation

Photoreceptor death can cause secondary glial scarring in the retina. For this purpose, we assessed retinal gliosis by GFAP staining and quantitated the fluorescence integrated density. We observed that Pde6brd10 recipients receiving RIP3KO photoreceptors showed similar gliosis (330.28 ± 21.47) as compared to Pde6brd10 recipients receiving RIP3WT photoreceptors (299.92 ± 6.80, p = .160), (Figure 2C and Supplementary Figure 3). In contrast, Pde6brd10 RIP3KO recipients showed significantly lower gliosis in either RIP3WT (272.97 ± 10.66, p = .037) or RIP3KO (281.40 ± 15.02, p = .038) donors, when compared to Pde6brd10 recipients. This reduced gliosis in RIP3KO recipients could be either due to less cell death in the grafted area in these recipients or an inherent feature of the Pde6brd10 RIP3KO strain given the overall reduction of cone cell death.(12)

5.4.5. Recipient RIPK3 Deletion Reduces the Ocular Inflammatory Cytokines after Subretinal Delivery

To further investigate the effects of RIPK3 deletion in both graft and host in enhancing the survival of the transplanted cells, we assessed the effect on the inflammatory response. For this purpose, we first investigated the effects of recipient RIPK3 deletion on the inflammatory cytokine cascade. We analyzed the cytokine profile 24-hours after a PBS subretinal injection in wild-type, RIP3KO, Pde6brd10, and Pde6brd10 RIP3KO mice (Supplementary Figure 5). As expected, we observed that the degenerating Pde6brd10 mice showed a significant fold increase in the expression of the inflammatory cytokines CCL2, CCL3, CCL4, CCL11, CCL12, CXCL1, CXCL10, and IL-10 in the retina and CCL2 and CCL12 in the RPE/Choroid when compared to the wild-type, RIP3KO, and Pde6brd10 RIP3KO mice (Supplementary Figure 5C). Interestingly, this fold increase

was abrogated in Pde6brd10 RIP3KO mice. These results suggest that subretinal delivery by a retinal detachment in Pde6brd10 mice can expose transplanted photoreceptors to pro-inflammatory cytokines. Most importantly, RIPK3 deletion in this strain appears to dampen the pro-inflammatory milieu in this scenario.



ID	Tissue	Strain	Cytokine	Fold Change	p Value
1	Retina	rd10	CCL2	6.27 ± 1.36	0.0001
2	Retina	rd10	CXCL1	5.96 ± 1.21	0.0001
3	Retina	rd10	CCL12	5.50 ± 0.92	0.0001
4	Retina	rd10	CXCL10	3.79 ± 0.94	0.0001
5	Retina	rd10	CCL3	3.77 ± 0.93	0.0001
6	RPE/Choroid	rd10	CCL2	3.15 ± 0.61	0.0001
7	RPE/Choroid	rd10	CCL12	3.36 ± 0.86	0.0003
8	Retina	rd10	CCL4	3.24 ± 1.36	0.0070
9	Retina	rd10	IL-10	4.81 ± 0.47	0.0185
10	Retina	rd10	CCL11	3.13 ± 0.62	0.0463

Supplementary Figure 5. Recipient RIPK3 Deletion Reduces Ocular Inflammatory Cytokines after Subretinal Delivery.

(A) Cytokine profile after subretinal injection. (B) Volcano plot constructed using a cut-off p value ≤ 0.05 for statistical significance and a ≥ 3-fold change in cytokine expression compared to wild-type expression. (C) List of cytokines showing a ≥ 3-fold change and statistically significant expression compared to wild-type.

5.4.6. Peripheral Immune RIPK3 Deletion in the Recipient Increases the Survival of Grafted Photoreceptor Precursors

Reduced cytokine cascade observed in RIPK3 deletion could be due to either RIPK3 deletion in the retina/RPE/choroid or due to the RIPK3 deletion in infiltrating peripheral immune cells. Thus, chimeric animals were created by means of bone marrow transplant (BMT). First, Pde6brd10 RIP3KO mice were subject to myelosuppressive conditioning with busulfan and subsequent BMT with EGFP-positive RIP3WT or EGFP-positive RIP3KO hematopoietic precursors through lateral tail vein injection (Figure 3). Following, we validated successful donor engraftment and immune reconstitution after 4-5 weeks following BMT by peripheral blood FACS analysis (Figure 3B). Next, chimeric recipients were subject to a subretinal transplant of RIP3KO photoreceptor precursors, given the maximum effect observed using these grafted cells. We noted that Pde6brd10 RIP3KO recipients which received RIP3WT BMT had a significantly lower photoreceptor graft survival (4,216 ± 1,550 cells) when compared to Pde6brd10 RIP3KO recipients which received a RIP3KO BMT (10,424 ± 1,580 cells, $p = .016$) (Figure 3C) suggesting a fundamental role of RIPK3 in peripheral immune cells for photoreceptor graft survival. In addition, Pde6brd10 RIP3KO recipients with RIP3WT BMT receiving RIP3KO photoreceptors showed similar graft survivals to Pde6brd10 RIP3WT recipients receiving RIP3KO photoreceptors, further supporting the notion that RIPK3 deletion in infiltrating peripheral immune cells is a large contributor to the maximum graft survival seen in Pde6brd10 RIP3KO recipients receiving RIP3KO photoreceptors. Moreover, we challenged this finding with a different approach. Using RIP3KO mice as recipients, we performed a BMT with either RIP3WT or RIP3KO donor hematopoietic precursors. Following immune reconstitution, we induced a retinal detachment and evaluated photoreceptor survival. We chose a retinal detachment model considering that photoreceptor transplantation is performed via an induced retinal detachment to deliver the cellular contents to the subretinal space. After retinal detachment EGFP-positive peripheral immune cells infiltrate the detached retina (Figure 3E). Most interestingly, RIP3KO recipients receiving a RIP3KO BMT presented a significantly higher photoreceptor cell count, thus less host cell loss in the detached retina at 7 days after

detachment (1015.48 ± 18.4 photoreceptors) compared to recipients receiving RIP3WT BMT (804.45 ± 20 photoreceptors, $p < .001$) (Figure 3D). Collectively, these results conclusively demonstrate that peripheral immune cell RIPK3 deletion is protective for photoreceptor survival for both the host and graft. Moreover, this effect is independent of photoreceptor transplantation or sorting method, as the peripheral protective effect is also observed in a retinal detachment model.

5.4.7. Systemic and Local Pharmacologic RIPK Inhibition Enhances Graft Survival

In view of the neuroprotective effect of RIPK3 deletion observed at photoreceptor, ocular, and peripheral immune level, we next hypothesized that pharmacologic RIP Kinase inhibition may increase graft survival. For this purpose, we used Necrostatin-1, a RIP1 kinase inhibitor which compromises RIPK3 phosphorylation and ultimately abrogates the necroptosis cascade, and a variant of which has been tested in clinical phase 1 trials for neurodegenerative disorders.(18) Necrostatin-1 was administered at the time of subretinal delivery of the photoreceptor precursors locally via subretinal injection (300 μM) and systemically by implantation of an osmotic drug delivery pump (15 mg/kg/day) in Pde6brd10 recipient mice subject to photoreceptors precursors transplant. Remarkably, systemic and local Necrostatin-1 delivery increased the overall graft survival (9,623 ± 2,055 cells) compared to vehicle-treated animals (4,151 ± 1,186 cells, $p = .045$) (Figure 4). Taken together, these results suggest that pharmacologic RIPK inhibition can replicate to a certain degree the protection of RIPK3 deletion at photoceptor, ocular, and systemic level. Most importantly, Necrostatin-1 treatment appears to be a plausible alternative in promoting photoreceptor graft survival.

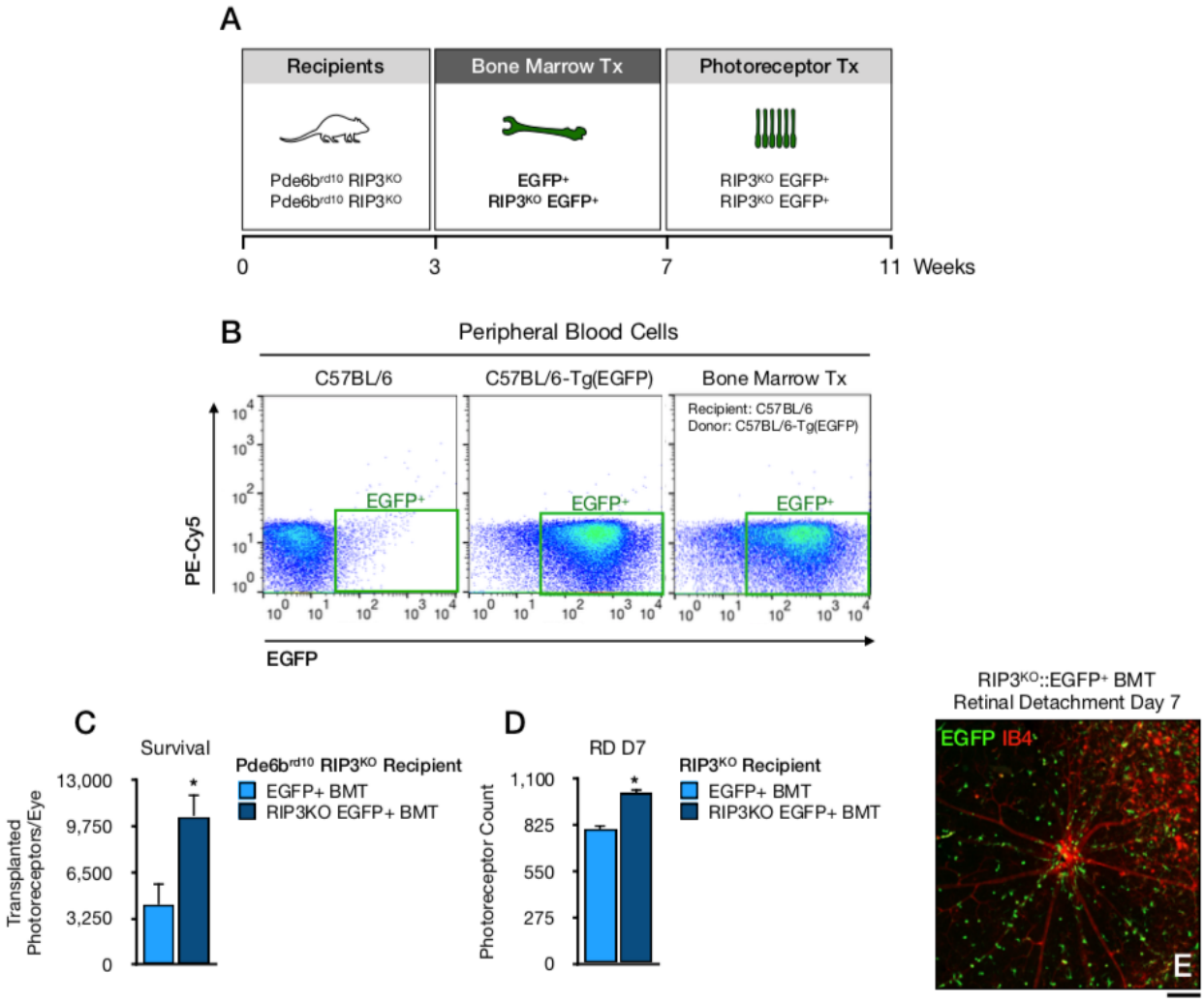


Figure 3. Recipient Peripheral Immune RIPK3 Deletion Increases Survival of Grafted Photoreceptor Precursors.

(A) Outline of bone marrow donors, bone marrow recipients, and photoreceptor donors used for the experiments. (B) Successful donor engraftment and immune reconstitution after 4 weeks following bone marrow transplantation. Analysis of EGFP expression in peripheral blood by FACS analysis. (C) Quantitation of photoreceptor survival in chimeric recipients' subject to subretinal transplant of RIP3KO EGFP-positive photoreceptor precursors (n = 6-8). (D) Quantitation of photoreceptor survival in chimeric recipients' subject to retinal detachment after 7 days (n = 6-8). (E) Representative retinal flat mount image of EGFP-positive peripheral immune cell infiltration of the retina in chimeric mice after retinal detachment. Scale bar 100 µm (E).

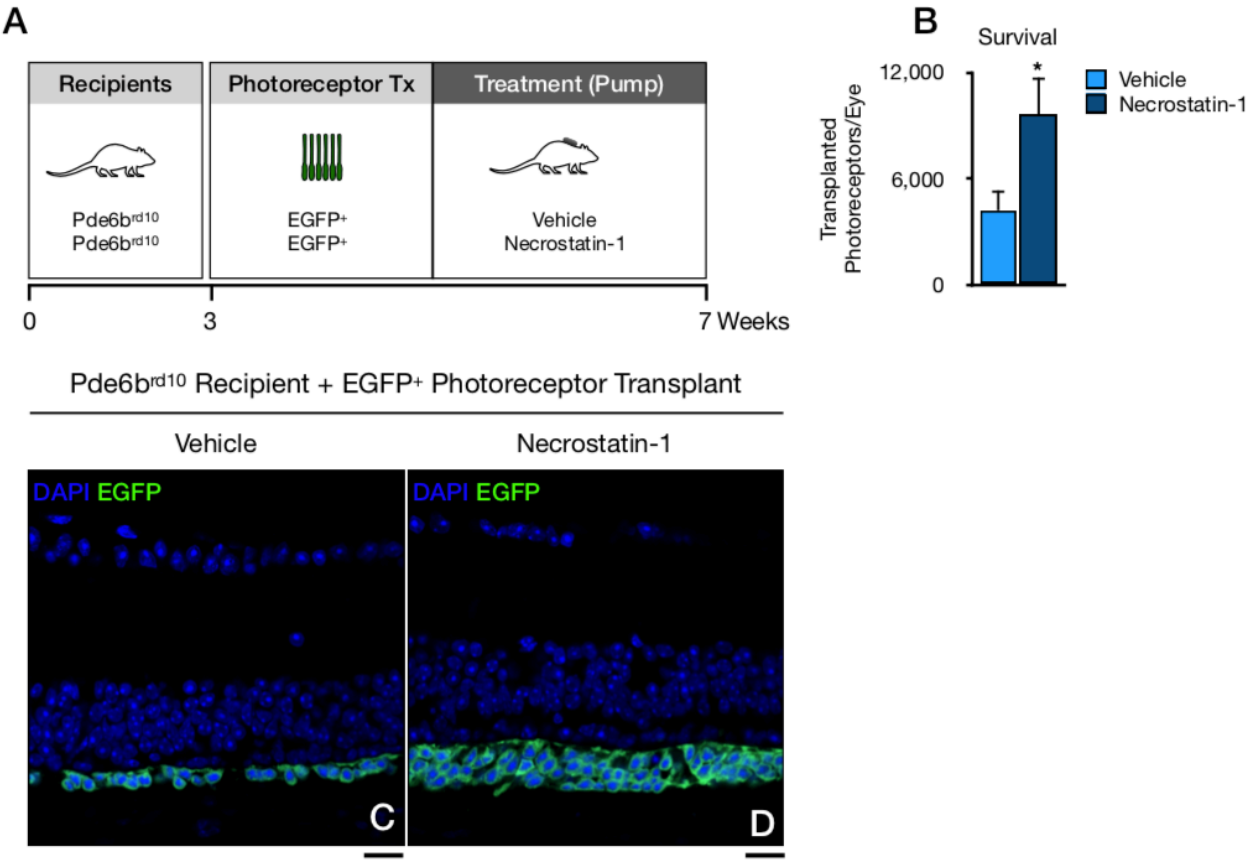


Figure 4. Systemic and Local Pharmacologic RIPK Inhibition Enhances Graft Survival.
(A) Outline of recipients, photoreceptor donors, and drugs used for the experiments. (B) Mean photoreceptor survival measured by quantitation of transplanted photoreceptors/eye. (C-D) Systemic and local Necrostatin-1 delivery increased the overall graft survival compared to vehicle-treated animals (n = 4-5). Scale bar 20 µm (A-B).

5.5. Discussion

In this work, we demonstrated that photoreceptor graft survival is enhanced by RIP3 kinase deficiency in donor photoreceptors and recipients. Most interestingly, recipient peripheral immune RIPK3 deletion was significantly protective for photoreceptor graft survival. Furthermore, these effects were replicated by pharmacologic inhibition of

RIPK1-RIPK3 axis. Altogether, these results strongly indicate the neuroprotective role of RIP Kinases in the retina through local and systemic interactions.

Our group has previously demonstrated in a model of retinal detachment that RIPK3 knock-out mice had less photoreceptor cell death and higher ONL thickness after injury.(13) In this work, RIPK3 deletion in donor photoreceptor precursors was sufficient to enhance graft survival in Pde6brd10 recipients. In addition, LDH release by cultured RIP3KO photoreceptors was significantly decreased compared to RIP3WT photoreceptors. These results in neurodegeneration and photoreceptor precursor transplantation appear to stand in agreement of previous published literature in non-neuronal solid organ transplant. As previously mentioned, Lau et al. have demonstrated that RIPK3-mediated necroptosis can promote inflammatory injury in a kidney transplant model.(22) Moreover, using a murine heart transplant model, Pavlosky et al. found that necroptotic cell death and release of the danger molecule high mobility group box 1 (HMGB1) were inhibited by genetic deletion of RIP3 kinase and by the RIP1 kinase inhibiting molecule Necrostatin-1.(23)

Previous work from our group and others has shown that after retinal detachment infiltrating cells migrating into the subretinal space contribute to neurodegeneration and cell loss.(17, 45–48) Interestingly, the number of immune CD11b-positive cells infiltrating the subretinal space was not significantly different in RIP3KO than in wild-type mice.(13, 17) However, western blot analysis showed that cleaved IL-1β production was significantly lower in RIP3KO in pooled samples containing subretinal fluid, retina, RPE, and choroid. We observed in this work that peripheral immune RIPK3 deletion was significantly protective for photoreceptor graft survival, and subsequently this finding was replicated in the retinal detachment model. We speculate that the reduced cleaved IL-1β production seen in RIP3KO pooled samples in our previous work may be due most likely to RIPK3 deletion in infiltrating peripheral immune cells.

Regarding myelosuppression strategies in bone marrow transplantation, Bruttger et al. have demonstrated that γ-irradiation can disrupt the blood-brain barrier leading to monocyte infiltration.(49) Recent evidence has shown that myelosuppressive

conditioning using busulfan preserved the retinal-blood barrier function without monocyte infiltration.(28, 50) Thus, we used busulfan as a conditioning agent prior to bone marrow transplantation to avoid blood-retinal barrier permeability changes.

Subretinal delivery is the preferred route to transplant photoreceptors precursors to the retina. Unfortunately, this approach induces a retinal detachment which triggers an inflammatory cascade.(45, 48, 51–53) We observed that Pde6brd10 mice which received a PBS subretinal injection had a significantly higher expression of CCL2, CCL3, CCL4, CCL11, CCL12, CXCL1, CXCL10, and IL-10 in the retina, and of CCL2 and CCL12 in the RPE/choroid, compared to the wild-type, RIP3KO, and Pde6brd10 RIP3KO mice. Thus, Pde6brd10 mice undergoing such profound retinal degeneration exposed transplanted photoreceptors to more inflammatory cytokines in the subretinal space. Interestingly, RIP3KO photoreceptors transplanted to Pde6brd10 recipients presented higher grafted survival that RIP3WT photoreceptors transplanted to these same recipients, which suggests that RIPK3 deletion in donors can aid in the tolerance to this pro-inflammatory milieu. Moreover, RIPK3 deletion appears to dampen the synthesis of these pro-inflammatory cytokines in the retina/subretinal space and RPE/choroid after subretinal injection. Unfortunately, in Pde6brd10 RIP3KO recipients which received RIP3WT photoreceptors, it appears that the less inflammatory milieu observed in this strain was not sufficient to rescue photoreceptors and promote a higher graft survival. In contrast, donor and recipient RIPK3 deletion maximized graft survival most likely due to combination of photoreceptor, retina, RPE/choroid, and peripheral immune effects.

In spite of the encouraging results seen with combined donor and recipient RIPK3 deletion, it is important to illustrate that most of the grafted photoreceptors fail to survive at four weeks post-transplant. Only 0.26% of the initial number of subretinally delivered photoreceptor precursors survived in Pde6brd10 recipients receiving RIP3WT photoreceptors. We did observe a ~ 4-fold increase (1.04%) in Pde6brd10 recipients receiving RIP3KO photoreceptors, and a ~7-fold increase (1.86%) in Pde6brd10 RIP3KO recipients receiving RIP3KO photoreceptors.

It is worth to mention that we did not observe migration of any EGFP+ cell to any other

retinal layer. Moreover, OLM disruption in the Crb1rd8 mice did not enhance photoreceptor contact with the recipient INL or ONL. Regarding previous evidence, it appears that the increased photoreceptor integration observed with OLM disruption(40, 41) may have been facilitated by expedited cytoplasmic GFP transfer through this disrupted membrane. Considering our results, photoreceptors delivered to the subretinal space do not migrate through retinal layers.

Regarding the procedure to isolate, sort, and transplant photoreceptors, there are key points to mention. First, our experimental design was subject to modification given the difficulty creating the Nrl-EGFP RIP3KO strain. Nonetheless, we followed the same MACS CD73-positive enrichment protocol as described above for EGFP and RIP3KO EGFP-positive animals. Second, the procedure was performed in equal conditions for all donors and recipients. Most importantly, we did not observe any EGFP-negative cells expressing reactivity to retinal or RPE markers in the grafted area, which indicates that transplanted cells are indeed rod photoreceptors. Another key point is that even though MACS confers a less traumatic sorting process for photoreceptors which may dampen the occurrence of cell death(25, 44), a complete elution of the reagents used is unlikely. More specific, the enriched photoreceptors may still carry trace CD73 antibody molecules and magnetic beads, which could be subsequently injected to the subretinal space. Despite this theoretical limitation, MACS has already been used in humans to enrich both embryonic stem cells in an implantable cardiac patch, and peripheral stem cells after high-dose chemotherapy.(54–56) To date, no long-term side effects have been reported after MACS procedures.(56)

In summary, targeting the RIPK3 pathway not only in transplanted cells but also in the recipient can help to promote graft survival. Furthermore, combination of immunomodulatory and neuroprotective strategies targeting necroptotic pathways can aid to restore photoreceptor loss in retinal degenerations. Despite this advancement, more work is required in this field to tackle the low survival rate of grafted cells.

5.6. Acknowledgements

Authors contributions: D.E.M. and D.G.V. designed the experiments. D.E.M. and L.G.B. performed the experiments. D.E.M and D.G.V. analyzed the data and wrote the paper. J.W.M. provided guidance and edited versions of the paper.

5.7. Funding Support

Funding Support: This work was supported by Bayer® Healthcare GOAP Ophthalmology Award (DEM), Alfonso Martin Escudero Foundation (LGB), NEI R21EY023079-01 A1 and R01EY02536201 (DGV), Yeatts Family foundation (DGV), Martin S Loeffler Renewed Hope Foundation (DGV), ARI Young Investigator Award (DGV), RPB Physician-Scientist Award (DGV), and NEI Grant EY014104 (MEEI Core Grant). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

5.8. Financial Disclosures

The authors have no relevant financial relationships or interests to disclose.

5.9. Chapter Summary in Spanish

5.9.1. Introducción

Las degeneraciones de la retina son un grupo heterogéneo de enfermedades, que incluyen la degeneración macular relacionada con la edad (AMD) y la retinitis pigmentosa (RP),

que se caracterizan por la muerte celular de los fotorreceptores y la pérdida severa de la visión que afecta a millones de personas sin una terapia efectiva.(1-4) Existen varias estrategias terapéuticas, entre las cuales se incluyen la administración de fármacos, la terapia génica y el trasplante de fotorreceptores. Entre estas estrategias, la terapia génica ha demostrado ser la más prometedora, pero este enfoque requiere que las células sigan siendo viables en la retina para transducir el gen de interés.(5) Esto da como resultado una ventana terapéutica estrecha para estas enfermedades neurodegenerativas. Para superar esta limitación, el trasplante de fotorreceptores se presenta como una estrategia de reparación para abordar las retinas degeneradas o ya degeneradas. Desafortunadamente, los esfuerzos para lograr estrategias de trasplante de fotorreceptores se han visto limitados por una eficiencia extremadamente limitada.(6–8) Por lo tanto, es crucial investigar las condiciones necesarias para optimizar el trasplante de fotorreceptores.

La muerte celular y el rechazo inmunológico son la principal limitación el trasplante de órganos y células.(9, 10) Por lo tanto, es razonable suponer que abordar estas dos variables relacionadas podría mejorar los resultados del trasplante al promover la supervivencia del injerto y reducir el rechazo inmunológico. Trabajos recientes han resaltado un papel importante de la proteína quinasa 3 que interactúa con el receptor (*receptor-interacting protein kinase 3*, RIP3/RIPK3) en la muerte celular necrótica programada (necroptosis) y la inflamación.(11-17) Esta isoforma es un miembro de la familia de quinasas RIP, la cual impulsa la vía de necroptosis desencadenada por el factor de necrosis tumoral (*tumoral necrosis factor*, TNF) principalmente a través de receptores *toll-like* TLR3 y TLR4, entre otros.(18) RIPK3 no solo ha demostrado mediar la muerte celular necrótica programada, sino que también ha demostrado funciones independientes de la necroptosis, como la activación del inflamasoma NLRP3.(19–21) Además, los estudios de trasplante de órganos sólidos han demostrado que el rechazo del aloinjerto puede ser mediado por la necroptosis y la inflamación vía RIPK3. (22, 23) A pesar de esta evidencia alentadora, aún se desconoce si la modulación selectiva de la vía RIPK3 podría mejorar el resultado del trasplante de fotorreceptores en las retinas degeneradas.

En este trabajo, investigamos el papel de la vía RIPK3 en la supervivencia de los trasplantes de fotorreceptores en un modelo animal de degeneración retiniana hereditaria. Usando quimeras de trasplante de médula ósea, investigamos el papel de las quinasas RIP en las células inmunes. Observamos que la combinación de la deficiencia genética simultanea de RIPK3 en donantes de fotorreceptores y en huéspedes maximiza la supervivencia del injerto. Además, confirmamos con experimentos de trasplante de médula ósea que la deficiencia de RIPK3 en células inmunitarias periféricas tiene un efecto protector para la supervivencia de los fotorreceptores. Además, los efectos neuroprotectores de RIPK3 pueden replicarse en parte por la inhibición farmacológica de la RIPK1. En conjunto, estos resultados indican que las estrategias neuroprotectoras en el trasplante de fotorreceptores pueden ayudar en las terapias de restauración de la retina.

5.9.2. Resultados

Teniendo en cuenta la evidencia reciente de artefactos en la integración del trasplante de fotorreceptores, primero buscamos abordar cualquier factor de confusión sobre la transferencia de citoplasma entre injerto y huésped. Elegimos el modelo de ratón *Pde6brd10* de degeneración retiniana que tiene una profunda pérdida de fotorreceptores en el día postnatal 30 (P30), lo que resulta en una sola línea celular en la capa nuclear externa de la retina (*outer nuclear layer*, ONL) en el tiempo final del experimento. (30-32) Dado un período de cuatro semanas, necesario para evaluar la supervivencia del injerto, cualquier línea adicional de fotorreceptores nucleares externos positivos para proteína verde fluorescente mejorada (*enhanced green fluorescent protein*, EGFP) en la ONL del receptor no puede deberse a la transferencia citoplasmática. Los donantes precursores de fotorreceptores EGFP+ se aislaron de ratones P4-7 *Nrl (neural retina leucine zipper)*-EGFP + y se seleccionaron positivamente usando un anticuerpo anti-CD73 y posteriormente enriquecidos mediante selección de células activadas magnéticamente (*magnetic-activated cell sorting*, MACS). El antígeno CD73, también conocido como ecto-5'-nucleotidasa, es una molécula de la superficie celular expresada en precursores de fotorreceptores.(33) Se ha demostrado que el enriquecimiento celular positivo mediante MACS CD73 es eficaz para aislar una población de precursores de

fotorreceptores (bastones) trasplantables.(34– 37) Después de MACS, trasplantamos los precursores de fotorreceptores CD73+ en el espacio subretiniano de ratones P30 *Pde6brd10*. Observamos que después de cuatro semanas, los precursores de bastones trasplantados repueblan la ONL del huésped en ratones *Pde6brd10* (Figura 1 complementaria). Curiosamente, observamos que, con una técnica de imagen adecuada y niveles de exposición al láser, se consigue una visualización adecuada del injerto sin signos de expresión *background* EGFP + en la retina del huésped. Curiosamente, la transferencia citoplasmática en la capa nuclear interna (INL) podría verse con valores de laser sobreexpuestos, lo cual se asemeja a publicaciones previas.(38, 39)

Se especula que la interrupción de la membrana limitante externa (*outer limiting membrane*, OLM) puede mejorar la integración de los fotorreceptores, incluso en animales de tipo salvaje, según publicaciones previas.(40, 41) Para este propósito, trasplantamos precursores de fotorreceptores CD73-positivos positivos para *Nrl*-EGFP en el espacio subretiniano de los animales que portan mutación *rd8* en el gen *Crb1*, por lo tanto, con una OLM interrumpida. Observamos que los fotorreceptores trasplantados a ratones *Crb1rd8* no pueden contactar con la INL o ONL del huésped, lo más probable debido a la presencia en estos últimos de los segmentos externos e internos en la retina del huésped (Figura 1). Dada la amplia distancia entre el injerto y ONL/INL del huésped, la sinapsis no solo está ausente, sino que es poco probable que ocurriera. En contraste, los fotorreceptores injertados en los ratones *Pde6brd10* mostraron contacto cercano con la INL/ONL del huésped, y mostraron esférulas sinápticas CTBP2+ (*C-terminal-binding protein 2*) con la INL. En conjunto, estos resultados sugieren que los fotorreceptores trasplantados requieren una degeneración de ONL preexistente para facilitar la sinapsis con la INL o ONL del huésped. Este fenómeno parece ocurrir independientemente de la integridad inicial de la OLM, que incluso ya puede estar interrumpida en las etapas posteriores de la degeneración de la retina.(42, 43)

La proteína RIPK3 es un regulador clave de la vía de muerte celular necroptótica. Trabajos anteriores han demostrado que la modulación de esta vía aumenta la supervivencia de los conos en modelos de degeneración retiniana hereditaria (12) y

fotorreceptores en el desprendimiento experimental de retina.(13) Además, dada la potencial función de las vías de muerte celular en el trasplante de fotorreceptores (44), Intentamos examinar si la modulación de la vía RIPK3 podría mejorar la supervivencia del injerto después de la implantación subretiniana. Desafortunadamente, el cruce de ratones *Nrl*-EGFP y aquellos con una deleción del gen RIP3K (*knock out*, KO) no tuvo éxito. Para abordar esta limitación, criamos ratones EGFP-positivos con RIP3KO y generamos precursores de fotorreceptores deficientes en RIPK3 y EGFP-positivos. Para seleccionar positivamente los fotorreceptores, utilizamos la misma estrategia MACS-CD73 como se describió anteriormente. Finalmente, el fenotipo de los precursores de fotorreceptores CD73 positivos para EGFP positivos y trasplantados se confirmó por falta de reactividad a marcadores clásicos como *retinal pigment epithelium-specific 65 kDa protein* (RPE65), *brain-specific homeobox/POU domain protein 3ª (BRN3A*), *protein kinase C alpha* (PKC-1α), y *glial fibrillary acidic protein* (GFAP) (Figuras complementarias 2 y 3).

Además, no observamos células EGFP negativas que expresen reactividad a estos marcadores en el área del injerto, lo que sugiere que la clasificación MACS-CD73 puede enriquecer eficazmente los precursores de bastones aislados. Tomados en conjunto, estos resultados sugieren que los precursores de fotorreceptores EGFP positivos injertados repueblan la ONL, y no representan o se diferencian en células epiteliales de pigmento retiniano (*retitnal pigment epithelium*, RPE), células ganglionares, células bipolares, astrocitos, o células Muller.

Posteriormente, aislamos y trasplantamos fotoreceptores EGFP+ de ratones *wild-type* (RIP3WT) o RIP3KO EGFP+ a huéspedes *Pde6brd10*. Notamos que la deleción de RIPK3 en donantes de fotorreceptores aumentó significativamente la supervivencia de los fotorreceptores in vitro (Figura 4 complementaria). Similarmente, la deleción de RIPK3 aumento significativamente la supervivencia general del injerto a las cuatro semanas (4,152.44 ± 538.44 células) en comparación con los donantes de RIP3WT (1,043.00 ± 238.37 células, p = .009) en huéspedes *Pde6brd10* (Figura 2B). En conjunto, estos resultados indican que la deleción de RIPK3 en los precursores de fotorreceptores

del donante es suficiente para mejorar la supervivencia del injerto en los receptores de Pde6brd10.

Trabajos previos ha demostrado previamente que la deleción de RIPK3 en ratones *Pde6brd10* aumenta la supervivencia de los conos y muestra menos infiltración microglial en comparación con los controles *wild-type* en P42 y P56. (12) En conjunto con los resultados mencionados anteriormente, planteamos la hipótesis de que la modulación simultánea de RIPK3 en donantes y huéspedes podría mejorar la supervivencia del injerto de trasplante de fotorreceptores. Para este propósito, utilizamos ratones *Pde6brd10* RIP3KO para determinar si la deleción de RIPK3 en el huésped podría influir en la supervivencia del injerto. En este grupo, el huésped *Pde6brd10* RIP3KO, la deleción de RIPK3 en donantes de fotorreceptores maximizó la supervivencia del injerto (7,425.29 ± 2,415.23) en comparación con los donantes RIP3WT (1,004.00 ± 199.67, $p = 0.005$) (Figura 2B). Estos resultados sugieren que la deleción de RIPK3 en el huésped no es una condición suficiente para rescatar fotorreceptores de donantes *wild-type*, pero parece ser necesaria para maximizar la supervivencia global de los fotorreceptores de RIP3KO.

La muerte por fotorreceptores puede causar cicatrices gliales secundarias en la retina. Para este propósito, evaluamos la gliosis retiniana mediante tinción con GFAP y cuantificamos la densidad integrada de la fluorescencia. Observamos que los huéspedes *Pde6brd10* que recibieron fotorreceptores RIP3KO mostraron una gliosis similar (330.28 ± 21.47) con los huéspedes *Pde6brd10* que recibieron fotorreceptores RIP3WT (299.92 ± 6.80, $p = 0.160$), (Figura 2C y Figura Complementaria 3). En contraste, los huéspedes *Pde6brd10* RIP3KO mostraron una gliosis significativamente menor en ambos donantes RIP3WT (272.97 ± 10.66, $p = 0.037$) o RIP3KO (281.40 ± 15.02, $p = 0.038$), en comparación con los huéspedes *Pde6brd10*. Esta reducción de la gliosis en los huéspedes RIP3KO podría deberse a una menor muerte celular en el injerto en estos huéspedes o debido a una característica inherente de la cepa *Pde6brd10* RIP3KO dada la reducción general de la muerte celular de los conos. (12)

Para investigar a fondo los efectos de la deleción de RIPK3 tanto en el injerto como en el

huésped, para mejorar la supervivencia de las células trasplantadas, evaluamos el efecto de la respuesta inflamatoria. Para este propósito, primero investigamos los efectos de la deleción de RIPK3 en la cascada de citoquinas inflamatorias. Analizamos el perfil de citoquinas 24 horas después de una inyección subretiniana de PBS en ratones *wild-type*, RIP3KO, *Pde6brd10* y *Pde6brd10* RIP3KO (Figura 5 complementaria). Como anticipamos, observamos que los ratones *Pde6brd10* mostraron un aumento significativo de la expresión de las citoquinas inflamatorias CCL2, CCL3, CCL4, CCL11, CCL12, CXCL1, CXCL10 e IL-10 en la retina y CCL2 y CCL12 en el RPE/Coroides en comparación con los ratones wild-type, RIP3KO, y Pde6brd10 RIP3KO (Figura 5C suplementaria). Curiosamente, este aumento se anuló en ratones *Pde6brd10* RIP3KO. Estos resultados sugieren que la administración del injerto en el espacio subretiniano subretiniano mediante un desprendimiento de retina en ratones *Pde6brd10,* puede exponer los fotorreceptores trasplantados a citoquinas proinflamatorias. Lo más importante, la deleción de RIPK3 en estos huéspedes parece minimizar el entorno proinflamatorio en este escenario.

La reducción de la cascada de citoquinas observada en la deleción de RIPK3 podría deberse a la eliminación de RIPK3 en la retina/RPE/coroides, o debido a la deleción de RIPK3 en las células inmunes periféricas infiltrantes. Pare dilucidar este interrogante, decidimos crear animales quiméricos mediante trasplante de médula ósea (*bone marrow transplant*, BMT). Primero, los ratones *Pde6brd10* RIP3KO fueron sometidos a acondicionamiento mielosupresor con Busulfán, y luego se realizo el BMT con precursores hematopoyéticos RIP3WT EGFP+ o RIP3KO EGFP+ a través de una inyección en la vena lateral de la cola (Figura 3). A continuación, validamos el éxito de la reconstitución inmune después de 4-5 semanas después del BMT mediante análisis de citometría de flujo (*fluorescence-activated cell sorting*, FACS) en sangre periférica (Figura 3B). A continuación, los huéspedes quiméricos se sometieron a un trasplante subretiniano de precursores de fotorreceptores RIP3KO, dado el efecto máximo observado utilizando estas células. Observamos que los huéspedes *Pde6brd10* RIP3KO que recibieron RIP3WT BMT tuvieron una supervivencia del injerto de fotorreceptores significativamente más baja (4.216 ± 1.550 células) en comparación con los receptores

Pde6brd10 RIP3KO que recibieron un RIP3KO BMT (10.424 ± 1.580 células, $p = 0.016$) (Figura 3C), lo que sugiere un papel fundamental de RIPK3 en las células inmunes periféricas para la supervivencia del injerto fotorreceptor. Además, los huéspedes *Pde6brd10* RIP3KO con RIP3WT BMT que recibieron fotorreceptores RIP3KO mostraron una supervivencia del injerto similar a los huéspedes *Pde6brd10* RIP3WT que reciben fotorreceptores RIP3KO, lo que respalda aún más la noción de que la deleción de RIPK3 en las células inmunes periféricas es un gran contribuyente de la máxima respuesta observada en huéspedes *Pde6brd10* RIP3O que recibieron fotorreceptores RIP3KO. A continuación, desafiamos este hallazgo con un enfoque diferente. Usando ratones RIP3KO como huéspedes, realizamos un BMT con precursores hematopoyéticos de donantes RIP3WT o RIP3KO. Después de la reconstitución inmune, realizamos un desprendimiento de retina y evaluamos la supervivencia de los fotorreceptores. Elegimos este modelo experimental teniendo en cuenta que el trasplante de fotorreceptores se realiza a través de un desprendimiento de retina inducido para entregar el contenido celular al espacio subretiniano. Después del desprendimiento de retina, las células inmunes periféricas EGFP+ infiltran en la retina desprendida (Figura 3E). Lo más interesante es que los huéspedes RIP3KO que recibieron un RIP3KO BMT presentaron un recuento de fotorreceptores significativamente más alto, por lo tanto, menos pérdida de células en la retina desprendida a los 7 días después del desprendimiento (1015.48 ± 18.4 fotorreceptores) en comparación con los huéspedes que recibieron RIP3WT BMT (804.45 ± 20 fotorreceptores, $p < 0.001$) (Figura 3D). Colectivamente, estos resultados demuestran de manera concluyente que la deleción de RIPK3 en células inmunes periféricas es protectora para la supervivencia de los fotorreceptores tanto para el huésped como en el injerto. Además, este efecto es independiente del trasplante de fotorreceptores o método de aislamiento, ya que el efecto protector periférico también se observa en un modelo de desprendimiento de retina.

En vista del efecto neuroprotector de la deleción de RIPK3 observada en fotorreceptores, a nivel ocular, y a nivel inmune periférico, a continuación planteamos la hipótesis de que la inhibición farmacológica de las quinasas RIP puede aumentar la supervivencia del injerto. Para este propósito, utilizamos Necrostatin-1, un inhibidor de la quinasa RIP1 que

compromete la fosforilación de RIPK3 y finalmente anula la cascada de necroptosis. Estas es una variante de la cual ha sido probada en ensayos clínicos de fase 1 para trastornos neurodegenerativos.(18) En nuestro caso, se administró Necrostatin-1 en el momento del suministro de los precursores de fotorreceptores localmente mediante inyección subretiniana (300 μM) y sistémicamente mediante la implantación de una bomba de suministro osmótico (15 mg/kg/día) en ratones huéspedes *Pde6brd10* sujetos a trasplante de precursores de fotorreceptores. Sorprendentemente, el suministro sistémico y local de Necrostatin-1 aumentó la supervivencia general del injerto (9.623 ± 2.055 células) en comparación con los animales tratados con vehículo (4.151 ± 1.186 células, $p = 0.045$) (Figura 4). Tomados en conjunto, estos resultados sugieren que la inhibición farmacológica de RIPK puede replicar en cierto grado la protección observada con la deleción de RIPK3 a nivel de fotorreceptor, ocular y sistémico. Lo más importante es que el tratamiento con Necrostatin-1 parece ser una alternativa plausible para promover la supervivencia del injerto de fotorreceptores.

5.9.3. Discusión

En este trabajo, demostramos que la supervivencia del injerto de fotorreceptores se ve reforzada por la deficiencia de la quinasa RIP3 en tanto fotorreceptores donantes como en huéspedes. Lo más interesante es que la deleción de RIPK3en células inmunes periféricas del huésped fue significativamente protectora para la supervivencia del injerto fotorreceptor. Además, estos efectos se replicaron por inhibición farmacológica del eje RIPK1-RIPK3. En conjunto, estos resultados indican fuertemente el papel neuroprotector de las quinasas RIP en la retina a través de interacciones locales y sistémicas.

Nuestro grupo ha demostrado previamente en un modelo de desprendimiento de retina que los ratones *knock-out* RIPK3 tienen menos muerte de fotorreceptores y un mayor grosor de ONL después de la lesión.(13) En nuestro caso, la liberación de LDH por los fotorreceptores RIP3KO cultivados fue significativamente menor en comparación con los fotorreceptores RIP3WT. Estos resultados en la neurodegeneración y el trasplante de precursores de fotorreceptores parecen estar de acuerdo con la literatura publicada anteriormente en el trasplante de órganos sólidos no neuronales. Como se mencionó

anteriormente, Lau et al. han demostrado que la necroptosis mediada por RIPK3 puede promover una respuesta inflamatoria en un modelo de trasplante de riñón.(22) Además, utilizando un modelo de trasplante de corazón murino, Pavlosky et al. descubrieron que la muerte celular necroptótica y la liberación de la proteína *high mobility group box 1* (HMGB1) fueron inhibidas por la eliminación genética de la quinasa RIP3 y por la molécula inhibidora de la quinasa RIP1 Necrostatin-1.(23)

El trabajo previo de nuestro grupo y otros ha demostrado que después del desprendimiento de retina, las células infiltrantes que migran hacia el espacio subretiniano contribuyen a la neurodegeneración y la pérdida celular.(17, 45-48) Curiosamente, el número de células inmunes CD11b positivas que infiltraron el espacio subretiniano no fue significativamente diferente en RIP3KO que en ratones wild-type.(13, 17) Sin embargo, el análisis de Western blot mostró que la producción de IL-1β escindida fue significativamente menor en RIP3KO en muestras agrupadas que contenían fluido subretiniano, retina, RPE y coroides. Observamos en este trabajo que la deleción de RIPK3 en células inmunes periféricas fue significativamente protectora para la supervivencia del injerto, y posteriormente este hallazgo se repitió en el modelo de desprendimiento de retina. Especulamos que la producción reducida de IL-1β escindida observada en las muestras agrupadas de RIP3KO en trabajo anteriores probablemente se deba a la deleción de RIPK3 en las células inmunes periféricas infiltrantes.

Con respecto a las estrategias de mielosupresión en el trasplante de médula ósea, Bruttger et al. han demostrado que la irradiación gamma (γ) puede alterar la barrera hematoencefálica, lo cual conduce a la infiltración de monocitos.(49) La evidencia reciente ha demostrado que el acondicionamiento mielosupresor con Busulfán conserva la función de barrera hemato-retiniana sin infiltración de monocitos.(28, 50) Por lo tanto, utilizamos Busulfan como agente acondicionador antes del trasplante de médula ósea para evitar cambios en la permeabilidad de la barrera hemato-retiniana.

La entrega subretiniana es la ruta preferida para trasplantar fotorreceptores a la retina. Este enfoque induce un desprendimiento de retina que termina por desencadenas una cascada inflamatoria.(45, 48, 51-53) Observamos que los ratones *Pde6brd10* que

recibieron una inyección subretiniana de *phosphate-buffered saline* (PBS) tenían una expresión significativamente mayor de CCL2, CCL3, CCL4, CCL11, CCL12, CXCL1, CXCL10 e IL-10 en la retina, y de CCL2 y CCL12 en RPE/Coroides. Por lo tanto, los ratones *Pde6brd10* expusieron los fotorreceptores trasplantados a incluso más citoquinas inflamatorias en el espacio subretiniano. Curiosamente, los fotorreceptores RIP3KO trasplantados a huéspedes *Pde6brd10* presentaron una mayor supervivencia injertada que los fotorreceptores RIP3WT trasplantados a estos mismos huéspedes, lo que sugiere que la deleción de RIPK3 en los donantes puede ayudar a la tolerancia a este entorno proinflamatorio. Además, la deleción de RIPK3 parece amortiguar la síntesis de estas citocinas proinflamatorias en la retina/espacio subretiniano y RPE/Coroides después de la inyección subretiniana. Desafortunadamente, en los huéspedes *Pde6brd10* RIP3KO que recibieron fotorreceptores RIP3WT, el ambiente menos inflamatorio observado en esta cepa no fue suficiente para rescatar fotorreceptores y promover una mayor supervivencia del injerto. Por el contrario, la eliminación de RIPK3 del donante y del huésped maximizó la supervivencia del injerto probablemente debido a la combinación del efecto a nivel de fotorreceptores, retina, RPE/Coroides, e inmune periférico.

A pesar de los resultados alentadores observados con la deleción combinada de RIPK3 en donantes y huéspedes, es importante aclarar que la mayoría de los fotorreceptores injertados no logran sobrevivir a las cuatro semanas después del trasplante. Solo el 0.26% del número inicial de precursores de fotorreceptores administrados por vía subretiniana sobrevivió en huéspedes *Pde6brd10* que recibieron fotorreceptores RIP3WT. Observamos un aumento de ~ 4 veces (1.04%) en los huéspedes de *Pde6brd10* que recibieron fotorreceptores RIP3KO, y un aumento de ~ 7 veces (1.86%) en los huéspedes de *Pde6brd10* RIP3KO que recibieron fotorreceptores RIP3KO.

Vale la pena mencionar que no observamos migración de ninguna célula EGFP + a ninguna otra capa retiniana. Además, la interrupción de OLM en los ratones *Crb1rd8* no mejoró el contacto del fotorreceptor con la INL u ONL del huésped. Con respecto a la evidencia previa, parece que la mayor integración de fotorreceptores observada con la disrupción de la OLM (40, 41) puede haberse facilitado por la transferencia acelerada de

GFP citoplasmática a través de esta membrana ya alterada. Teniendo en cuenta nuestros resultados, los fotorreceptores suministrados al espacio subretiniano no migran a través de las capas retinianas.

Con respecto al procedimiento para aislar, clasificar y trasplantar fotorreceptores, hay puntos clave que mencionar. Primero, nuestro diseño experimental debió ser sujeto a modificaciones dada la dificultad de crear la cepa *Nrl*-EGFP RIP3KO. Sin embargo, seguimos el mismo protocolo de enriquecimiento positivo MACS-CD73 como se describió anteriormente para los animales EGFP+ y RIP3KO EGFP+. En segundo lugar, el procedimiento se realizó en igualdad de condiciones para todos los donantes y huéspedes. Lo que es más importante, no observamos células EGFP negativas que expresen reactividad a los marcadores retinianos o RPE en el área injertada, lo que indica que las células trasplantadas son fotorreceptores. Otro punto clave es que, a pesar de que MACS confiere un proceso de clasificación menos traumático para los fotorreceptores que puede reducir la muerte celular (25, 44), es poco probable una elución completa de los reactivos utilizados. Más específicamente, los fotorreceptores enriquecidos aún pueden transportar moléculas de anticuerpo CD73 y microesferas magnéticas, que posteriormente podrían inyectarse en el espacio subretiniano. A pesar de estas limitaciones, MACS ya se ha utilizado en humanos para enriquecer tanto las células madre embrionarias en un parche cardíaco implantable, como las células madre periféricas después de la quimioterapia en altas dosis.(54-56) Hasta la fecha, no se han reportado efectos secundarios a largo plazo después de procedimientos que usen MACS. (56)

En resumen, abordar la vía RIPK3 no solo en las células trasplantadas sino también en el huésped puede ayudar a promover la supervivencia del injerto. Además, la combinación de estrategias inmunomoduladoras y neuroprotectoras dirigidas a las vías necroptóticas puede ayudar a restaurar la pérdida de fotorreceptores en las enfermedades degenerativas de la retina. A pesar de este avance, se requiere más trabajo en este campo para abordar la baja tasa de supervivencia del injerto.

Chapter VI

Photoreceptor Cell Death: Discussion

Chapter VI. Discussion

In this work, we describe a two-tined goal mission to improve our diagnostic and therapeutic approaches in retinal degeneration. First, we sought to reconceptualize cell death quantitation from a manual to an automated method. Second, we aimed to further improve replacement therapies in retinal degenerations. We managed to address a technical bottleneck in estimating retinal cell death, and subsequently optimized the survival of transplanted photoreceptors as a treatment modality for retinal degenerations.

Manual quantitation had been performed to date to quantitate cell death and retinal thickness. However, this process is prone to measurement errors, not entirely reproducible, and bears a significant time burden and resource allocation. Therefore, an automated approach is of interest to address these difficulties.

For our first aim, we hypothesized that the development of automated methods to quantitate cell death could improve precision, accuracy, and reduce time burden. For this purpose, we developed an automated ImageJ plugin, TUNEL Cell Counter, which quantitated images acquired from fluorescent based TUNEL cell death assay.

All in our dataset were processed by the TUNEL Cell Counter, with reliable ONL and INL segmentation and TUNEL-positive cell labelling. When comparing ImageJ macro and observers' performance for area and TUNEL+ cells, we observed no statistically significant differences in TUNEL+ cell counts and ratio of TUNEL+ cells/area between the experienced observer and macro. In contrast, significant differences were observed in these variables between the inexperienced observer and macro ($p < .001$). This suggests that the TUNEL Cell Counter does not perform differently to an experienced observer.

Following, we determined the agreement between TUNEL Cell with observers. As visualized in Bland-Altman plots, counts for the TUNEL Cell Counter presented a smaller bias value and a more grouped and narrow confidence interval observers. In addition, TUNEL Cell Counter was able to detect true and statistically significant differences between experimental groups in an independent validation dataset.

With regard to the time burden, TUNEL Cell Counter needed less than 3 seconds per image, which represents, a 60- and 160-fold faster than the experienced and inexperienced observers, respectively.

In summary, agreement analysis, experimental validation, and time burden analysis proves that TUNEL Cell Counter is superior to manual quantitation. Taken together, this confirms our hypothesis that automated cell death quantitation performance is superior to manual quantitation. We expect that advent of this tool will provide accurate, precise, and fast cell death quantitation. Moreover, we currently provide TUNEL Cell Counter freely to the scientific community as open-source.

For our second aim, and given our experience with TUNEL Cell Counter, we hypothesized that the development of similar automated tools to quantitate retinal thickness could optimize our current evaluation of retinal thinning and degeneration. For this purpose, we developed a second automated ImageJ plugin, ThicknessTool, with a similar validation framework.

We designed mock image datasets with masks of known area thickness and rotation to calibrate the algorithm. ThicknessTool accurately measured thickness with zero bias, and no significant difference between known and measured values. Moreover, when processing retinal images, ONL and INL thickness measurements are not significantly different from any observers' average. A deeper analysis showed that manual quantitation lacked repeatability, with poor accuracy and reproducibility.

Agreement analyses and Bland-Altman plots showed that ThicknessTool can measure ONL and INL with lower bias than observers' average against each other. Moreover, in an experimental validation dataset, ThicknessTool detected significant true thinning beyond manual measurement capabilities. The processing time for the ThicknessTool depends on the caliper number, and orbits around couple seconds per area in standard conditions. This time burden is negligible compared to manual observers.

In summary, calibration and agreement analysis, combined with experimental validation demonstrate that ThicknessTool is superior to manual quantitation. This confirms our hypothesis that automated retinal thickness quantitation performance is superior to manual quantitation. Similarly, we currently provide ThicknessTool freely to the scientific community as open-source.

Finally, for our third aim, our objective was to further contribute to our mission of restoring sight in retinal degenerations. In this case, not only by developing reliable technical methods to estimate cell death, thinning, and ultimately retinal degeneration; but we were also strongly committed to provide a therapeutic approach to improve vision in these diseases. Inasmuch, we optimized a translational approach to repopulate retinal degeneration.

Given recent evidence on the role of receptor-interacting protein kinase 3 (RIP3/RIPK3) in necrotic cell death and inflammation, we hypothesized that selective targeting of the RIPK3 pathway could enhance the outcome of single-cell photoreceptor transplant to degenerated retinas.

We found that grafted photoreceptors required a pre-existing retinal degeneration to facilitate synapsis with the host inner or outer nuclear layer. Moreover, this appeared to occur regardless of the initial integrity of the outer limiting membrane. Selective RIPK3 deletion in donor photoreceptor precursors was sufficient to enhance graft survival in rd10 recipients. However, RIPK3 deletion in the recipient was not sufficient to rescue wild-type donor photoreceptors. Most importantly, RIPK3 deletion in donor photoreceptors and recipients maximized graft survival. Bone marrow transplantation experiments demonstrated that peripheral immune cell RIPK3 deletion was protective for photoreceptor survival for both the graft and recipients. Finally, a pharmacologic RIPK inhibition with Necrostatin-1 promoted photoceptor survival.

In summary, these results indicate that RIP3K modulation in donor photoreceptors, hosts, and peripheral immune cells, maximize the number of grafted photoreceptors in the subretinal space at endpoint. This confirms our hypothesis that RIP3K modulation in

hosts and/or donors improves the survival of grafted photoreceptors. These results are indeed promising and may broaden the therapeutic alternatives for patients with retinal degenerations.

In summary, we believe this work can further improve the diagnostic objectivity in retinal research and improve therapeutic strategies in photoreceptor transplantation to aid in restoration therapies. We hope this will lead to better alternatives for patients suffering from retinal degenerations.

Chapter VII

Photoreceptor Cell Death: References

Chapter VII. References

7.1. Chapter I

1. Masland RH. The fundamental plan of the retina. Nat Neurosci 2001;4:877–886.

2. Masland RH. The Neuronal Organization of the Retina. Neuron 2012;76:266–280. Available at: http://dx.doi.org/10.1016/j.neuron.2012.10.002.

3. Hoon M, Okawa H, Della Santina L, Wong ROL. Functional architecture of the retina: Development and disease. Prog Retin Eye Res 2014;42:44–84. Available at: http://dx.doi.org/10.1016/j.preteyeres.2014.06.003.

4. Peynshaert K, Devoldere J, Minnaert AK, et al. Morphology and Composition of the Inner Limiting Membrane: Species-Specific Variations and Relevance toward Drug Delivery Research. Curr Eye Res 2019;44:465–475. Available at: https://doi.org/10.1080/02713683.2019.1565890.

5. Jonas JB, Dichtl A. Evaluation of the retinal nerve fiber layer. Surv Ophthalmol 1996;40:369–378.

6. Curcio CA, Allen KA. Topography of ganglion cells in human retina. J Comp Neurol 1990;300:5–25.

7. Jones MK, Lu B, Girman S, Wang S. Cell-based therapeutic strategies for replacement and preservation in retinal degenerative diseases. Prog Retin Eye Res 2017;58:1–27. Available at: http://dx.doi.org/10.1016/j.preteyeres.2017.01.004.

8. McCudden CR, Hains MD, Kimple RJ, et al. G-protein signaling: back to the future. Cell Mol Life Sci 2005;62:551–77. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15747061.

9. Lamb TD, Pugh EN. Phototransduction, dark adaption, and rhodopsin regeneration: The proctor lecture. Invest Ophthalmol Vis Sci 2006;47:5138–5152.

10. Raviola E, Gilula NB. Gap junctions between photoreceptor cells in the vertebrate retina. Proc Natl Acad Sci U S A 1973.

11. Strettoi E, Dacheux RF, Raviola E. Synaptic connections of rod bipolar cells in the inner

plexiform layer of the rabbit retina. J Comp Neurol 1990.

12. Boycott BB, Wässle H. Morphological Classification of Bipolar Cells of the Primate Retina. Eur J Neurosci 1991.

13. Euler T, Wässle H. Immunocytochemical identification of cone bipolar cells in the rat retina. J Comp Neurol 1995.

14. Schneeweis DM, Schnapf JL. Photovoltage of rods and cones in the macaque retina. Science (80-) 1995.

15. Soucy E, Wang Y, Nirenberg S, et al. A novel signaling pathway from rod photoreceptors to ganglion cells in mammalian retina. Neuron 1998.

16. Hack I, Peichl L, Brandstätter JH. An alternative pathway for rod signals in the rodent retina: Rod photoreceptors, cone bipolar cells, and the localization of glutamate receptors. Proc Natl Acad Sci U S A 1999.

17. Tsukamoto Y, Morigiwa K, Ueda M, Sterling P. Microcircuits for night vision in mouse retina. J Neurosci 2001.

18. Fyk-Kolodziej B, Qin P, Pourcho RG. Identification of a cone bipolar cell in cat retina which has input from both rod and cone photoreceptors. J Comp Neurol 2003.

19. Li W, Keung JW, Massey SC. Direct synaptic connections between rods and OFF cone bipolar cells in the rabbit retina. J Comp Neurol 2004.

20. Famiglietti E V, Kolb H. A bistratified amacrine cell and synaptic cirucitry in the inner plexiform layer of the retina. Brain Res 1975;84:293–300. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1111833.

21. Pan F, Toychiev A, Zhang Y, et al. Inhibitory masking controls the threshold sensitivity of retinal ganglion cells. J Physiol 2016;594:6679–6699.

22. Völgyi B, Deans MR, Paul DL, Bloomfield SA. Convergence and segregation of the multiple rod pathways in mammalian retina. J Neurosci 2004;24:11182–11192.

23. Murakami Y, Matsumoto H, Roh M, et al. Receptor interacting protein kinase mediates

necrotic cone but not rod cell death in a mouse model of inherited degeneration. Proc Natl Acad Sci 2012;109:14598–14603. Available at: www.pnas.org/cgi/doi/10.1073/pnas.1206937109 [Accessed January 21, 2019].

24. Trichonas G, Murakami Y, Thanos A, et al. Receptor interacting protein kinases mediate retinal detachment-induced photoreceptor necrosis and compensate for inhibition of apoptosis. Proc Natl Acad Sci 2010;107:21695–21700. Available at: www.pnas.org/cgi/doi/10.1073/pnas.1009179107 [Accessed January 21, 2019].

25. Barber AC, Hippert C, Duran Y, et al. Repair of the degenerate retina by photoreceptor transplantation. Proc Natl Acad Sci 2013;110:354–359. Available at: www.pnas.org/cgi/doi/10.1073/pnas.1212677110 [Accessed January 25, 2019].

7.2. Chapter III

1. Cuenca N, Fernández-Sánchez L, Campello L, et al. Cellular responses following retinal injuries and therapeutic approaches for neurodegenerative diseases. Prog Retin Eye Res. 2014;43:17-75..

2. Murakami Y, Matsumoto H, Roh M, et al. Programmed necrosis, not apoptosis, is a key mediator of cell loss and DAMP-mediated inflammation in dsRNA-induced retinal degeneration. Cell Death Differ. 2014;21(2):270-277.

3. Murakami Y, Matsumoto H, Roh M. Receptor interacting protein kinase mediates necrotic cone but not rod cell death in a mouse model of inherited degeneration. Proc Natl Acad Sci U S A. 2012;109(36):14598–14603.

4. Johnson TV, DeKorver NW, Levasseur VA, et al. Identification of retinal ganglion cell neuroprotection conferred by platelet-derived growth factor through analysis of the mesenchymal stem cell secretome. Brain. 2014;137(Pt 2):503-519.

5. Galluzzi L, Kroemer G. Necroptosis: a specialized pathway of programmed necrosis.

Cell. 2008;135(7):1161-1163.

6. Gavrieli Y, Sherman Y, Ben-Sasson SA. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. J Cell Biol. 1992;119(3):493-501.

7. Liming P, Bradley CJ, Liu JJ. The correlativity analysis of six methods of detecting apoptosis. Chin Med Sci J. 1999;14(3):145-151.

8. Grasl-Kraupp B, Ruttkay-Nedecky B, Koudelka H, Bukowska K, Bursch W, Schulte-Hermann R. In situ detection of fragmented DNA (TUNEL assay) fails to discriminate among apoptosis, necrosis, and autolytic cell death: a cautionary note. Hepatology. 1995;21(5):1465-1468.

9. Pasparakis M, Vandenabeele P. Necroptosis and its role in inflammation. Nature. 2015;517(7534):311-320.

10. Trichonas G, Manola A, Morizane Y, et al. A novel nonradioactive method to evaluate vascular barrier breakdown and leakage. Invest Ophthalmol Vis Sci. 2010;13(3):317-323.

11. Negoescu A, Guillermet C, Lorimier P, Brambilla E, Labat-Moleur F. Importance of DNA fragmentation in apoptosis with regard to TUNEL specificity. Biomed Pharmacother. 1998;52(6):252-258.

12. Matsumoto H, Miller JW, Vavvas DG. Retinal Detachment Model in Rodents by Subretinal Injection of Sodium Hyaluronate. JoVE. 2013;(79).

13. Matsumoto H, Murakami Y, Kataoka K, et al. Mammalian STE20-like kinase 2, not kinase 1, mediates photoreceptor cell death during retinal detachment. Cell Death Dis. 2014;5(5):e1269.

14. Lunscher WH, Beddoes MP. Optimal edge detector design I: parameter selection and noise effects. IEEE Trans Pattern Anal Mach Intell. 1986;8(2):164-177.

15. Tsai W-H. Moment-preserving thresolding: A new approach. GMIP. 1985;29(3):377-

393.

16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-310.

17. Filippi M, Horsfield MA, Rovaris M, et al. Intraobserver and interobserver variability in schemes for estimating volume of brain lesions on MR images in multiple sclerosis. AJNR Am J Neuroradiol. 1998;19(2):239-244

18. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet. 1995;346(8982):1085-1087.

19. Chan TF, Chiu K, Lok CKM, Lau HW, So K-F, Chang RC-C. Morphometric analyses of retinal sections. JoVE. 2012;(60).

20. Bland JM, Altman DG. Agreed statistics: measurement method comparison. Anesthesiology. 2012;116(1):182-185.

7.3. Chapter IV

1. Cook, B., Lewis, G. P., Fisher, S. K. & Adler, R. Apoptotic photoreceptor degeneration in experimental retinal detachment. Investig. Ophthalmol. Vis. Sci. 36, 990–6 (1995).

2. Arroyo, J. G., Yang, L., Bula, D. & Chen, D. F. Photoreceptor apoptosis in human retinal detachment. Am. J. Ophthalmol. 139, 605–10 (2005).

3. Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. J. Cell Biol. 119, 493–501 (1992).

4. Maidana, D. E. et al. A novel imagej macro for automated cell death quantitation in the retina. Investig. Ophthalmol. Vis. Sci. 56, 6701–6708 (2015).

5. Byun, J. et al. Automated tool for the detection of cell nuclei in digital microscopic images: Application to retinal images. Mol. Vis. 12, 949–960 (2006).

6. Trichonas, G. et al. Receptor interacting protein kinases mediate retinal detachment-induced photoreceptor necrosis and compensate for inhibition of apoptosis. Proc Natl Acad Sci. 107, 21695–21700 (2010).

7. Matsumoto, H. et al. Strain difference in photoreceptor cell death after retinal detachment in mice. Invest Ophthalmol Vis Sci. 55, 4165–4174 (2014).

8. Murakami, Y. et al. Receptor interacting protein kinase mediates necrotic cone but not rod cell death in a mouse model of inherited degeneration. Proc Natl Acad Sci. 109, 14598–14603 (2012).

9. Byun, J. et al. Quantifying structural distortions in retinal tissue before and after injury. in Workshop on Multiscale Biological Imaging, Data Mining & Informatics, Santa Barbara, CA, USA 10–11 (2006).

10. Arganda-Carreras, I., Fernández-González, R., Muñoz-Barrutia, A. & Ortiz-De-Solorzano, C. 3D reconstruction of histological sections: Application to mammary gland tissue. Microsc Res Tech. 73, 1019–29 (2010).

11. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1, 307–10 (1986).

12. Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. Stat. Methods Med. Res. 8, 135–60 (1999).

13. Troscianko, J. & Stevens, M. Image calibration and analysis toolbox - a free software suite for objectively measuring reflectance, colour and pattern. Methods Ecol. Evol. 6, 1320–1331 (2015).

14. Kim, H.-Y. Statistical notes for clinical researchers: Type I and type II errors in statistical decision. Restor. Dent. Endod. 40, 249 (2015).

15. Rothman, K. J. Curbing type i and type II errors. Eur. J. Epidemiol. 25, 223–224 (2010).

16. Bland, J. & Altman, D. Measuring agreement in method comparison studies with heteroscedastic measurements. Stat. Methods Med. Res. 8, 135–160 (1999).

17. Giavarina, D. Understanding Bland Altman analysis. Biochem. Medica 25, 141–151 (2015).

18. Cox, D. & Lewis, P. The Statistical Analysis of Series of Events. (1966).

7.4. Chapter V

1. J. L. Duncan, et al., Inherited Retinal Degenerations: Current Landscape and Knowledge Gaps. Transl. Vis. Sci. Technol. 7, 6 (2018).

2. R. R. A. Bourne, et al., Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe in 2015: magnitude, temporal trends and projections. Br. J. Ophthalmol. 102, 575–585 (2018).

3. W. L. Wong, et al., Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis. Lancet Glob. Heal. 2, e106–e116 (2014).

4. C. G. Owen, et al., The estimated prevalence and incidence of late stage age related macular degeneration in the UK. Br. J. Ophthalmol. 96, 752–756 (2012).

5. R. Kumar-Singh, Barriers for retinal gene therapy: Separating fact from fiction. Vision Res. 48, 1671–1680 (2008).

6. T. Santos-Ferreira, et al., Retinal transplantation of photoreceptors results in donor-host cytoplasmic exchange. Nat. Commun. 7 (2016).

7. R. A. Pearson, et al., Donor and host photoreceptors engage in material transfer following transplantation of post-mitotic photoreceptor precursors. Nat. Commun. 7 (2016).

8. M. S. Singh, et al., Transplanted photoreceptor precursors transfer proteins to host photoreceptors by a mechanism of cytoplasmic fusion. Nat. Commun. 7 (2016).

9. N. Pallet, M. Dieudé, J. Cailhier, M. Hébert, The molecular legacy of apoptosis in transplantation. Am. J. Transplant. 12, 1378–1384 (2012).

10. N. Zavazava, D. Kabelitz, Alloreactivity and apoptosis in graft rejection and transplantation tolerance. J Leukoc Biol 68, 167–174 (2000).

11. K. Moriwaki, F. K. M. Chan, RIP3: A molecular switch for necrosis and inflammation. Genes Dev. 27, 1640–1649 (2013).

12. Y. Murakami, et al., Receptor interacting protein kinase mediates necrotic cone but not rod cell death in a mouse model of inherited degeneration. Proc. Natl. Acad. Sci. 109, 14598–14603 (2012).

13. G. Trichonas, et al., Receptor interacting protein kinases mediate retinal detachment-induced photoreceptor necrosis and compensate for inhibition of apoptosis. Proc. Natl. Acad. Sci. 107, 21695–21700 (2010).

14. Y. Murakami, et al., Programmed necrosis, not apoptosis, is a key mediator of cell loss and DAMP-mediated inflammation in dsRNA-induced retinal degeneration. Cell Death Differ. 21, 270–7 (2014).

15. Y. Murakami, J. W. Miller, D. G. Vavvas, RIP kinase-mediated necrosis as an alternative mechanisms of photoreceptor death. Oncotarget 2, 497–509 (2011).

16. A. Al-Moujahed, et al., Receptor interacting protein kinase 3 (RIP3) regulates iPSCs generation through modulating cell cycle progression genes. Stem Cell Res. 35, 101387 (2019).

17. K. Kataoka, et al., Macrophage- and RIP3-dependent inflammasome activation exacerbates retinal detachment-induced photoreceptor cell death. Cell Death Dis. 6, e1731 (2015).

18. Y. Cho, et al., Phosphorylation-Driven Assembly of the RIP1-RIP3 Complex Regulates Programmed Necrosis and Virus-Induced Inflammation. Cell 137, 1112–1123 (2009).

19. K. Newton, et al., RIPK3 deficiency or catalytically inactive RIPK1 provides greater benefit than MLKL deficiency in mouse models of inflammation and tissue injury. Cell Death Differ. 23, 1565–1576 (2016).

20. M. Gerlic, et al., RIPK3 promotes cell death and NLRP3 inflammasome activation in the absence of MLKL. Nat. Commun. (2015).

21. X. Wang, et al., RNA viruses promote activation of the NLRP3 inflammasome through a RIP1-RIP3-DRP1 signaling pathway. Nat. Immunol. (2014) https:/doi.org/10.1038/ni.3015.

22. A. Lau, et al., RIPK3-mediated necroptosis promotes donor kidney inflammatory injury and reduces allograft survival. Am. J. Transplant. 13, 2805–2818 (2013).

23. A. Pavlosky, et al., RIPK3-mediated necroptosis regulates cardiac allograft rejection. Am. J. Transplant. 14, 1778–1790 (2014).

24. M. S. Singh, et al., Reversal of end-stage retinal degeneration and restoration of visual function by photoreceptor transplantation. Proc. Natl. Acad. Sci. 110, 1101–1106 (2013).

25. D. Eberle, S. Schubert, K. Postel, D. Corbeil, M. Ader, Increased integration of transplanted CD73-positive photoreceptor precursors into adult mouse retina. Investig. Ophthalmol. Vis. Sci. 52, 6462–6471 (2011).

26. J. Gust, T. A. Reh, Adult donor rod photoreceptors integrate into the mature mouse retina. Investig. Ophthalmol. Vis. Sci. 52, 5266–5272 (2011).

27. S. Ashizuka, W. H. Peranteau, S. Hayashi, A. W. Flake, Busulfan-conditioned bone marrow transplantation results in high-level allogeneic chimerism in mice made tolerant by in utero hematopoietic cell transplantation. Exp. Hematol. 34, 359–368 (2006).

28. E. I. Paschalis, et al., Permanent neuroglial remodeling of the retina following infiltration of CSF1R inhibition-resistant peripheral monocytes. Proc. Natl. Acad. Sci. 115, 201807123 (2018).

29. K. Peake, et al., Busulfan as a Myelosuppressive Agent for Generating Stable High-level Bone Marrow Chimerism in Mice. J. Vis. Exp., 1–7 (2015).

30. S. Rösch, S. Johnen, F. Müller, C. Pfarrer, P. Walter, Correlations between ERG, OCT, and anatomical findings in the rd10 mouse. J. Ophthalmol. 2014, 874751 (2014).

31. R. Barhoum, et al., Functional and structural modifications during retinal degeneration in the rd10 mouse. Neuroscience 155, 698–713 (2008).

32. C. Gargini, E. Terzibasi, F. Mazzoni, E. Strettoi, Retinal organization in the retinal degeneration 10 (rd10) mutant mouse: A morphological and ERG study. J. Comp. Neurol. 500, 222–238 (2007).

33. H. Koso, et al., CD73, a novel cell surface antigen that characterizes retinal photoreceptor precursor cells. Investig. Ophthalmol. Vis. Sci. 50, 5411–5418 (2009).

34. G. Gagliardi, et al., Characterization and Transplantation of CD73-Positive Photoreceptors Isolated from Human iPSC-Derived Retinal Organoids. Stem Cell Reports 11, 665–680 (2018).

35. J. Lakowski, et al., Transplantation of Photoreceptor Precursors Isolated via a Cell Surface Biomarker Panel from Embryonic Stem Cell-Derived Self-Forming Retina. Stem Cells 33, 2469–2482 (2015).

36. T. Santos-Ferreira, et al., Stem Cell–Derived Photoreceptor Transplants Differentially Integrate Into Mouse Models of Cone-Rod Dystrophy. Investig. Opthalmology Vis. Sci.

57, 3509 (2016).

37. D. Eberle, T. Santos-Ferreira, S. Grahl, M. Ader, Subretinal Transplantation of MACS Purified Photoreceptor Precursor Cells into the Adult Mouse Retina. J. Vis. Exp., e50932 (2014).

38. R. E. MacLaren, et al., Retinal repair by transplantation of photoreceptor precursors. Nature 444, 203–207 (2006).

39. E. L. West, R. A. Pearson, R. E. Maclaren, J. C. Sowden, R. R. Ali, Cell transplantation strategies for retinal repair. Prog. Brain Res. 6123, 3–21 (2012).

40. R. A. Pearson, et al., Targeted Disruption of Outer Limiting Membrane Junctional Proteins (Crb1 and ZO-1) Increases Integration of Transplanted Photoreceptor Precursors Into the Adult Wild-Type and Degenerating Retina. Cell Transplant. 19, 487–503 (2010).

41. E. L. West, et al., Pharmacological disruption of the outer limiting membrane leads to increased retinal integration of transplanted photoreceptor precursors. Exp. Eye Res. (2008) https:/doi.org/10.1016/j.exer.2008.01.004.

42. L. Fernández-Sánchez, P. Lax, L. Campello, I. Pinilla, N. Cuenca, Astrocytes and Müller Cell Alterations During Retinal Degeneration in a Transgenic Rat Model of Retinitis Pigmentosa. Front. Cell. Neurosci. 9, 1–16 (2015).

43. C. Hippert, et al., Müller glia activation in response to inherited retinal degeneration is highly varied and disease-specific. PLoS One 10, 1–27 (2015).

44. S. J. Gasparini, S. Llonch, O. Borsch, M. Ader, Transplantation of photoreceptors into the degenerative retina: Current state and future perspectives. Prog. Retin. Eye Res., 1–37 (2018).

45. T. Nakazawa, et al., Monocyte chemoattractant protein 1 mediates retinal detachment-induced photoreceptor apoptosis. Proc. Natl. Acad. Sci. 104, 2425–2430 (2007).

46. H. Matsumoto, et al., Mammalian STE20-like kinase 2, not kinase 1, mediates

photoreceptor cell death during retinal detachment. Cell Death Dis. 5, e1269–e1269 (2014).

47. X. Wang, et al., Rapid monocyte infiltration following retinal detachment is dependent on non-canonical IL6 signaling through gp130. J. Neuroinflammation (2017) https:/doi.org/10.1186/s12974-017-0886-6.

48. H. Matsumoto, et al., Strain difference in photoreceptor cell death after retinal detachment in mice. Investig. Ophthalmol. Vis. Sci. 55, 4165–4174 (2014).

49. J. Bruttger, et al., Genetic Cell Ablation Reveals Clusters of Local Self-Renewing Microglia in the Mammalian Central Nervous System. Immunity 43, 92–107 (2015).

50. E. I. Paschalis, et al., The Role of Microglia and Peripheral Monocytes in Retinal Damage after Corneal Chemical Injury. Am. J. Pathol. 188, 1580–1596 (2018).

51. T. Nakazawa, et al., Tumor necrosis factor-α mediates photoreceptor death in a rodent model of retinal detachment. Investig. Ophthalmol. Vis. Sci. (2011) https:/doi.org/10.1167/iovs.10-6509.

52. M. I. Roh, Y. Murakami, A. Thanos, D. G. Vavvas, J. W. Miller, Edaravone, an ROS scavenger, ameliorates photoreceptor cell death after experimental retinal detachment. Investig. Ophthalmol. Vis. Sci. (2011) https:/doi.org/10.1167/iovs.10-6797.

53. Y. Murakami, et al., Photoreceptor cell death and rescue in retinal detachment and degenerations. Prog. Retin. Eye Res. 37, 114–40 (2013).

54. P. Menasché, et al., Human embryonic stem cell-derived cardiac progenitors for severe heart failure treatment: First clinical case report. Eur. Heart J. 36, 2011–2017 (2015).

55. R. Handgretinger, et al., Isolation and transplantation of autologous peripheral CD34+ progenitor cells highly purified by magnetic-activated cell sorting. Bone Marrow Transplant. 21, 987–993 (1998).

56. G. Gagliardi, et al., Characterization and Transplantation of CD73-Positive Photoreceptors Isolated from Human iPSC-Derived Retinal Organoids. Stem Cell Reports 11, 665–680 (2018).

Appendix A

Photoreceptor Cell Death: Publications

Appendix A. Publications

1. Maidana DE, Tsoka P, Tian B, Dib B, Matsumoto H, Kataoka K, Lin H, Miller JW, Vavvas DG. A Novel ImageJ Macro for Automated Cell Death Quantitation in the Retina. *Invest Ophthalmol Vis Sci*. 2015 Oct;56(11):6701-8.

2. Maidana DE, Notomi S, Ueta T, Zhou T, Joseph D, Kosmidou C, Caminal-Mitjana JM, Miller JW, Vavvas DG. ThicknessTool: Automated ImageJ Retinal Layer Thickness and Profile in Digital Images. Accepted for publication in *Scientific Reports* (*In Press*, 2020).

A Novel ImageJ Macro for Automated Cell Death Quantitation in the Retina

Daniel E. Maidana, Pavlina Tsoka, Bo Tian, Bernard Dib, Hidetaka Matsumoto, Keiko Kataoka, Haijiang Lin, Joan W. Miller, and Demetrios G. Vavvas

Retina Service, Angiogenesis Lab, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts, United States

Correspondence: Demetrios G. Vavvas, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114, USA; Demetrios_Vavvas@MEEI.HARVARD.EDU.

Submitted: June 30, 2015
Accepted: August 31, 2015

Citation: Maidana DE, Tsoka P, Tian B, et al. A novel ImageJ macro for automated cell death quantitation in the retina. *Invest Ophthalmol Vis Sci*. 2015;56:6701–6708. DOI:10.1167/iovs.15-17599

Purpose. TUNEL assay is widely used to evaluate cell death. Quantification of TUNEL-positive (TUNEL$^+$) cells in tissue sections is usually performed manually, ideally by two masked observers. This process is time consuming, prone to measurement errors, and not entirely reproducible. In this paper, we describe an automated quantification approach to address these difficulties.

Methods. We developed an ImageJ macro to quantitate cell death by TUNEL assay in retinal cross-section images. The script was coded using IJ1 programming language. To validate this tool, we selected a dataset of TUNEL assay digital images, calculated layer area and cell count manually (done by two observers), and compared measurements between observers and macro results.

Results. The automated macro segmented outer nuclear layer (ONL) and inner nuclear layer (INL) successfully. Automated TUNEL$^+$ cell counts were in-between counts of inexperienced and experienced observers. The intraobserver coefficient of variation (COV) ranged from 13.09% to 25.20%. The COV between both observers was 51.11 $\pm$ 25.83% for the ONL and 56.07 $\pm$ 24.03% for the INL. Comparing observers' results with macro results, COV was 23.37 $\pm$ 15.97% for the ONL and 23.44 $\pm$ 18.56% for the INL.

Conclusions. We developed and validated an ImageJ macro that can be used as an accurate and precise quantitative tool for retina researchers to achieve repeatable, unbiased, fast, and accurate cell death quantitation. We believe that this standardized measurement tool could be advantageous to compare results across different research groups, as it is freely available as open source.

Keywords: cell death, inner nuclear layer, NMDA, outer nuclear layer, retina, retinal detachment, TUNEL assay

In recent years, research of degenerative retinal diseases has been focused on understanding cell death and ways of interfering with it. Several experimental animal models of diabetic retinopathy, retinitis pigmentosa, retinal detachment, and glaucoma, have revealed that cell death in the retina can occur through different modalities, including apoptosis and necroptosis among others, and encompasses multiple inducers and pathways and displays different morphologic characteristics.[1–5] From the many tools used to characterize and quantify cell death, the terminal deoxynucleotidyl transferase (dUTP) nick end labeling (TUNEL) assay is likely the most widespread method. This assay detects fragmented or nicked DNA by means of a terminal deoxynucleotidyl transferase enzyme, which incorporates fluorescent-labeled dUTPs in damaged nucleic acid regions. First described by Gavrieli et al.,[6] this assay has been a sensitive method for detecting apoptosis.[7] However, as TUNEL assay detects DNA fragments regardless of the induced cell death pathway, it cannot distinguish between apoptotic or other forms of programmed cell death.[3,8–10] Despite its limitations, TUNEL assay remains the most widespread method used to screen for any form of programmed cell death, to date.

In order to quantitate TUNEL assay results in tissue sections, manual counting of TUNEL-positive (TUNEL$^+$) cells, ideally by two masked observers is performed.[11] Observers count TUNEL$^+$ cells and measure the respective retinal area, and results are expressed as either TUNEL$^+$ cells/area or TUNEL$^+$ cells/total cells. However, this process is time consuming, prone to measurement errors, and not entirely reproducible. Therefore, an automated approach is of interest to address these difficulties. To the best of our knowledge, there is no automated cell death quantitation method available, capable of segmenting retinal layers and counting individual TUNEL$^+$ cells in the retina without performance variability.

The purpose of this study was to introduce a novel ImageJ macro for the automated detection of cell death in retinal cross-sections. In addition, we studied differences in grading between an experienced and an inexperienced observer. Because the introduction of a new measurement method requires that it be parallel with the gold standard, or commonly accepted method, we validated this macro by comparing its performance relative to that of two observers. We found that the ImageJ macro can achieve a fast and accurate retinal cell death quantitation.

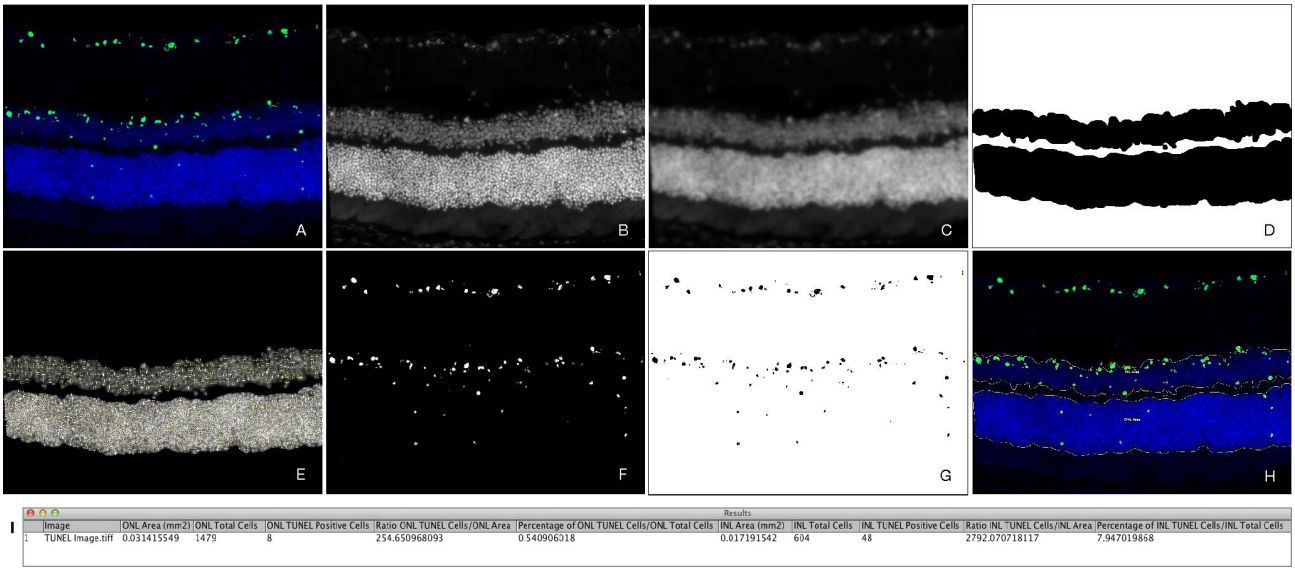


Figure 1. Segmentation of ONL, INL, and quantitation of TUNEL-positive cells by ImageJ macro. From the native RGB image (**A**), an 8-bit blue channel was extracted (**B**), and a Gaussian blur filter was applied (**C**). Using the Tsai moment-preserving thresholding method, we segmented ONL and INL layers (**D**) and, by determining the local maxima corresponding to cell nuclei, we identified individual cells at these layers (**E**). The 8-bit green channel was extracted, and background noise was subtracted (**F**). Using the Tsai moment-preserving thresholding method and binary watershed segmentation, we counted TUNEL$^+$ cells (**G**). A JPEG image overlay was automatically created and exported for visual assessment of the quantitation (**H**). Results from the macro were reported (**I**) as area (mm^2), total cells (count), TUNEL$^+$ cells (count), TUNEL$^+$ cells-to-area (count/mm^2) ratio, and percentage of TUNEL$^+$ cells over total cells (%), for both ONL and INL.

Materials and Methods

Fluorescence Microscopy Image Database

Digital images of TUNEL-stained retinal cross-sections were obtained from the Angiogenesis Laboratory fluorescence microscopy database. In order to quantitate TUNEL$^+$ cells on the outer nuclear layer (ONL), we randomly selected 90 images from an experimental murine retinal detachment model which induces photoreceptor cell death by subretinal injection of sodium hyaluronate.[12] For experimental comparison between groups, we selected 30 images from an experimental group (mammalian sterile 20-like kinase 2 [MST2] knock-out mice) and compared them to a control group (C57BL/6 mice), as previously published.[13] To assess inner nuclear layer (INL) TUNEL$^+$ cells, we randomly selected 90 images from an experimental *N*-methyl-D-aspartic acid (NMDA)-induced excitotoxicity murine model that uses two different drug doses (10 and 100 nM). Eyes from these experimental models were enucleated, and 7- to 10-µm thin sections were cut using a model CM1850 cryostat (Leica Microsystems, Bannockburn, IL, USA). TUNEL staining was performed with ApopTag fluorescein direct in situ apoptosis detection kit (Millipore, Bedford, MA, USA), according to manufacturer's instructions. Prior to being mounted, cell nuclei were stained with Quinolinium, 4-[3-(3-methyl-2(3H)-benzothiazolylidene)-1-propenyl]-1-[3-(trimethylammonio)propyl]-, di-iodide (TO-PRO 3; Life Technologies, Grand Island, NY, USA).

Because we sought to build an assorted and representative sample, we included images from the aforesaid models acquired by multiple independent researchers, which used different acquisition protocols. These images were acquired using fluorescence microscopy (Axio Imager model M2; Carl Zeiss, Inc., Thornwood, NY, USA) with a 20×/0.8 numerical aperture air objective, as we considered this magnification was the most suitable to assess a reasonably large area without missing details in cell morphology. Image exclusion criteria were poor TUNEL staining quality, uneven image focus, ONL or INL layers not distinguishable, retina cross-section not centered on the image frame, or significant shadowing. Finally, selected digital images were exported as 24-bit red-green-blue (RGB) images in an uncompressed tagged image file format (TIFF) for further analysis.

ImageJ Macro Script Programming

To develop an automated method to quantitate cell death by TUNEL assay in retinal cross-section images, we designed a macro for the ImageJ platform (version 1.49, http://imagej.nih.gov/ij/; provided in the public domain by the National Institutes of Health, Bethesda, MD, USA) and coded the script using IJ1 programming language on Fiji image processing software (http://fiji.sc/Fiji, in the public domain). After TIFF files importation, the most relevant image processing steps are detailed as follows.

Image Scale Setup. The image spatial scale (pixel-to-mm^2) was acquired from the microscope bundled software (Zen version 1.1; Carl Zeiss, Inc.) in order to express results in mm^2. This scale value can be easily obtained from image metadata.

ONL and INL Segmentation. From the native RGB digital image (Fig. 1A), an 8-bit blue channel was extracted, and a Gaussian blur filter was applied (Figs. 1B, 1C) in order to reduce noise and detail.[14] After contrast enhancement, we used Tsai moment-preserving thresholding method to obtain a binary image.[15] Following several erosion iterations to achieve a fitted segmentation, outlier pixels were removed to a specific radius. The largest area in the image was presumed to be the ONL area. Then, the segmented ONL region of interest was subtracted from the processed blue channel. The INL area was segmented following the same approach as that used for the ONL area (Fig. 1D).

ONL and INL Total Cells. Outer nuclear layer and INL regions of interest were used to determine specific areas where total cells were counted. By determining the local maxima corresponding to cell nuclei, individual cells at these layers were identified and recorded (Fig. 1E).

Investigative Ophthalmology & Visual Science

TABLE. Outcome Variables for Observers and ImageJ Macro

Variable	ImageJ Macro		Experienced Observer			Inexperienced Observer		
	Mean ± SD	Range, min–max	Mean ± SD	Range, min–max	*P* Value*	Mean ± SD	Range, min–max	*P* Value*
Outer nuclear layer, $n = 90$								
Area, mm^2	0.028 ± 0.004	0.020–0.037	0.030 ± 0.004	0.022–0.040	0.019	0.028 ± 0.004	0.021–0.037	0.958
TUNEL$^+$ cells, count	32.4 ± 19.6	3–97	40.5 ± 27.0	3.5–138	0.121	25.0 ± 19.0	2–89.5	0.194
TUNEL$^+$ cells:area ratio, count/mm^2	1197 ± 833	108–4742	1403 ± 1045	121–5938	0.552	915 ± 772	66–4019	0.230
Inner nuclear layer, $n = 90$								
Area, mm^2	0.016 ± 0.003	0.011–0.025	0.018 ± 0.003	0.012–0.027	0.009	0.018 ± 0.003	0.011–0.026	0.286
TUNEL$^+$ cells, count	30.6 ± 15.3	5–74	35.9 ± 19.6	6–98	0.204	19.9 ± 12.4	2–58.5	<0.001
TUNEL$^+$ cells:area ratio, count/mm^2	1805 ± 797	357–3551	1922 ± 986	319–4332	0.878	1106 ± 616	97–2854	<0.001
Total burden time per image, s†	2.82 ± 0.04	2.77–2.88	130.60 ± 17.96	95–196.50	<0.0001	340.40 ± 49.07	195.50–450.50	<0.0001

Values for experienced and inexperienced observers represent averages between first and second measurements. Statistical significance between macro and experienced or inexperienced observers was calculated by 1-way analysis of variance (ANOVA) with post hoc Tukey honest significant differences test for comparison between groups. SD, standard deviation; TUNEL$^+$, TUNEL positive.

* A *P* value of <0.05 was considered statistically significant.

† Total burden time per image represents total time required for ONL and INL segmentation and TUNEL$^+$ cells counting.

ONL and INL TUNEL$^+$ Cells. From the native RGB digital image, an 8-bit green channel was extracted, and background noise was subtracted, in order to achieve a cleaner image (Fig. 1F). Following that, we applied Tsai moment-preserving thresholding method to obtain a binary image and used binary watershed segmentation to separate any contiguous cells (Fig. 1G). This filter first calculates the Euclidian distance map and finds the ultimate eroded points. Following, it dilates each of the latter as far as possible, either until the edge of the particle is reached, or the edge touches a region of another. Individual cells were counted according to specific size and circularity.

After image processing, the macro generated a report with 10 quantitative outcome variables (Fig. 1I), which included (1) ONL area (mm^2); (2) ONL total cells (count); (3) ONL TUNEL$^+$ cells (count); (4) ONL TUNEL$^+$ cells to ONL area (count/mm^2) ratio; (5) percentage of ONL TUNEL$^+$ cells/ONL total cells (%); (6) INL area (mm^2); (7) INL total cells (count); (8) INL TUNEL$^+$ cells (count); (9) ratio of INL TUNEL$^+$ cells to INL area (count/mm^2); and (10) percentage of INL TUNEL$^+$ cells/INL total cells (%). Additionally, for a qualitative analysis of the script performance, a JPEG (Joint Photographic Experts Group) image overlay was automatically created and exported for visual assessment of segmented areas and quantitated TUNEL$^+$ cells (Fig. 1H).

Script Training, Validation, and Statistical Analysis

To calibrate and train the macro to experimental observer values, we first defined specific criteria to count TUNEL$^+$ cells, including fluorescent signal colocalization between green and blue channels and nuclei-like size and shape, and if TUNEL$^+$ cells appeared clumped, cell number was counted according to clump size. Most important, we encouraged observers to keep the same criteria for all images. Initially, images were assessed by an independent observer (BD) who manually segmented ONL and INL areas and counted TUNEL$^+$ cells by using ImageJ (version 1.48). Script parameters were fine tuned to these values. Following that, 2 masked observers, 1 experienced (PT) and 1 inexperienced observer (BT), assessed the image dataset using the same approach. Observers were given duplicate images within the set, without their knowledge, in order to assess measurement reproducibility in a blind manner. Outcome variables were recorded as was time required to quantitate each image.

Statistical analysis was performed with Prism version 6 software (GraphPad, San Diego, CA, USA) and JMP Pro software version 11.2.0 (SAS, Carey, NC, USA). Normality was assessed with Shapiro-Wilk test. Statistical significance for differences between groups was determined with Wilcoxon signed-rank test for matched pairs and 1-way ANOVA with Tukey post hoc correction for multiple comparisons. For correlation analysis, values were fitted to linear regression, and results were expressed as Spearman ρ correlation coefficient and R^2 goodness of fit for the model. Intra- and interobserver agreement analyses werre achieved by plotting the mean differences between measurements relative to their means, using Bland-Altman plot method.[16] Coefficients of variation (COV) were calculated as differences between measurements over relative to their means: $[(A - B)/([A + B]/2) \times 100]$.[17] A and B represent either first and second measurement in intraobserver agreement, or different observers in interobserver agreement. Results were expressed as means ± standard deviations (SD). A P value of <0.05 was considered statistically significant.

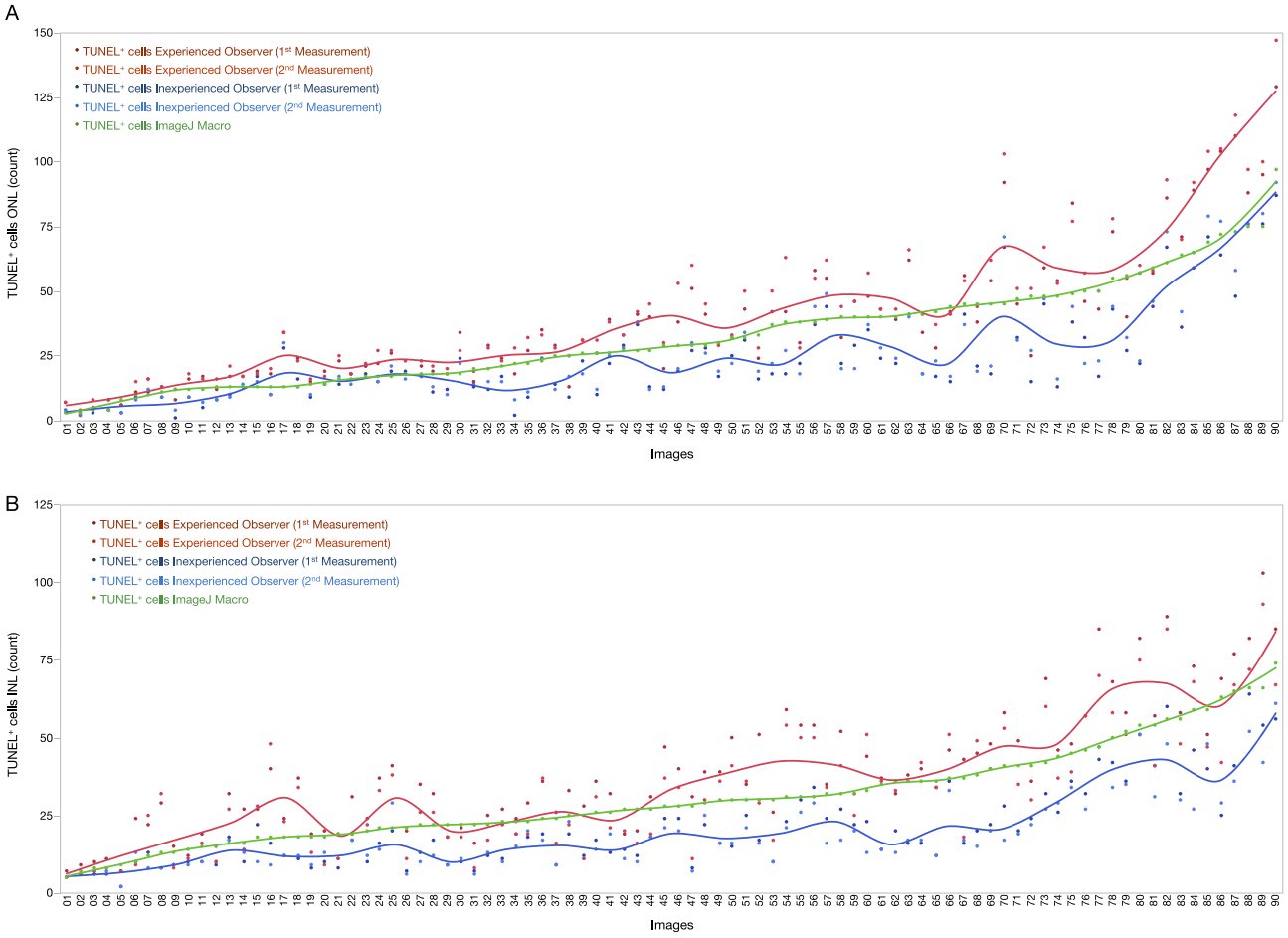


Figure 2. TUNEL-positive cell counts for the macro and mean values of inexperienced and experienced observers. Cell counts for the ImageJ macro and first and second measurements from observers plotted for every image, for ONL (**A**) and INL (**B**) datasets. A *colored trendline* represents mean variations for each observer. Macro TUNEL$^+$ counts are located between observer's measurements.

Results

All images that matched our selection criteria and were included in the dataset were processed by the ImageJ macro, with accurate segmentation of ONL and INL areas and TUNEL$^+$ cell count, as seen in the exported JPEG image overlay (Supplementary Fig. S1B). Additionally, all defined variables were quantitated and reported.

ImageJ Macro and Observers' Performance For Area and TUNEL-Positive Cell Density

Results from observers and macro measurements can be seen in the Table. ONL area measured by the automated macro (0.028 mm^2) was similar to that measured by the inexperienced observer (0.028 mm^2, $P = 0.958$) and 10% less than that measured by the experienced observer (0.030 mm^2, $P = 0.019$). No statistically significant differences were observed in TUNEL$^+$ cell counts and ratio of TUNEL$^+$ cells to area between the observers and macro. INL area measured by the macro (0.016 mm^2) was similar to that of the inexperienced observer (0.018 mm^2, $P = 0.286$) and 10% less than that of the experienced observer (0.018 mm^2, $P = 0.009$). No statistically significant differences were observed in TUNEL$^+$ cell counts and ratio of TUNEL$^+$ cells to area between the experienced observer and macro. In contrast, significant differences were seen between the variables of the inexperienced observer and those of the the macro ($P < 0.001$).

We calculated observers' total burden time per image by adding the time needed for ONL and INL segmentation and TUNEL$^+$ cell counting. On average, an experienced observer needed 2.16 minutes per image, whereas the inexperienced observer needed more than 5.66 minutes. Macro needed less than 3 seconds per image with batch processing (using 2.8-GHz Core i7 iMac model computer; Apple Co., Cupertino, CA, USA), which was 60- and 160-fold faster than the experienced and inexperienced observers, respectively (Supplementary Fig. S4).

Next, ONL and INL TUNEL$^+$ cell values were plotted for each image, observers, and automated macro. TUNEL$^+$ cell values for the macro were located between the lower values of the inexperienced observer and the higher counts of the experienced observer, with only few exceptions corresponding to images with significant background noise, due to lack of postprocessing in the native image (Figs. 2A, 2B).

Intraobserver Agreement and Correlation

Since repeatability is a crucial element for any measurement method, we first sought to determine intraobserver variability. To answer this question, observers quantitated 180 images within the dataset, twice, in a masked fashion. Values from the first and second measurements were compared for each image.

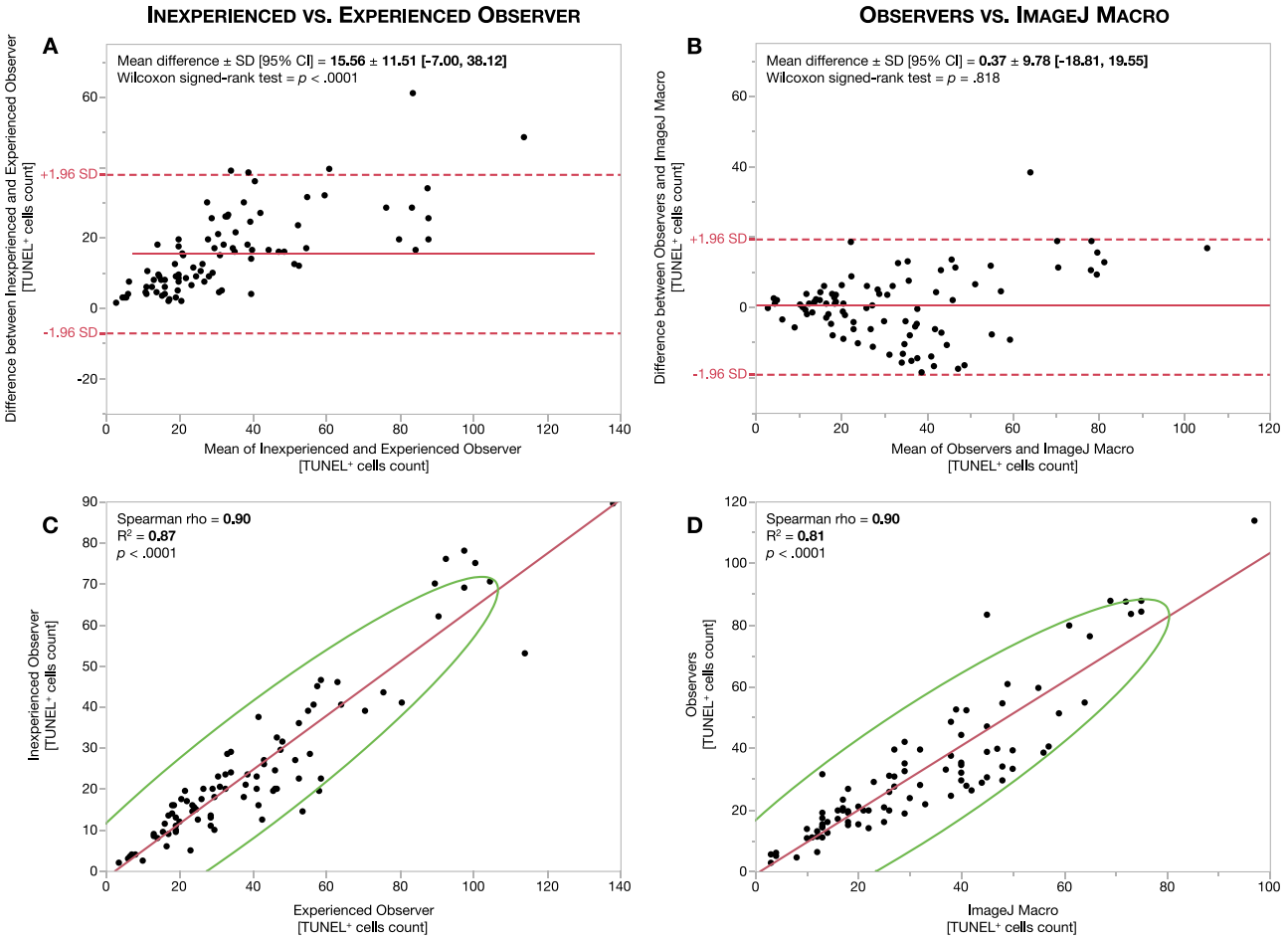


FIGURE 3. Outer nuclear layer inter-observer agreement and correlation. Bland-Altman plots for ONL measurement agreement between observers (**A**) and between observers and ImageJ macro (**B**) are shown. *Red lines* represent mean differences or bias. *Red dashed lines* represent 95% limit of agreement (±1.96 SD). Correlation analysis between observers (**C**) and between observers and ImageJ macro (**D**). *Red lines* represent the fitted linear regression *trendline*. *Green lines* represent the 95% confidence interval area.

We then plotted their differences relative to the means, using Bland-Altman plots. The COV of TUNEL$^+$ cell counts between these two measurements for the ONL dataset was 15.21 ± 19.69% for the inexperienced observer and 13.09 ± 13.84% for the experienced observer (Supplementary Table S1; Supplementary Figs. S2A, S2B). For the INL dataset, the COV was 16.41 ± 15.39% for the inexperienced observer and 25.20 ± 21.82% for the experienced observer (Supplementary Table S1; Supplementary Figs. S2B, S3A). Spearman ρ correlation coefficient between first and second measurements from the inexperienced observer was 0.95 in ONL images and 0.92 in INL images (Supplementary Table S1; Supplementary Figs. S2C–S3C). This coefficient was 0.96 in ONL images and 0.93 in INL images for the experienced observer (Supplementary Table S1, Supplementary Figs. S2D–S3D).

Interobserver Agreement and Correlation

To determine interobserver variability, we averaged the first and second measurements of each observer and obtained their respective mean TUNEL$^+$ cell count. COV of variation of TUNEL$^+$ cell counts between the two observers was 51.11 ± 25.83% for the ONL dataset (Supplementary Table S1; Fig. 3A) and 56.07 ± 24.03% for the INL dataset (Supplementary Table S1; Fig. 4A). Following, we sought to determine whether the ImageJ macro would agree with the observers. We used the limits of agreement between these two observers to define the range to which the macro TUNEL$^+$ cell counts should fit. Of note, these ranges were estimated for the whole dataset and should not be extrapolated or predicted as per image. Compared to the COV for the observers, the COV for the macro was 23.37 ± 15.97% for the ONL (Supplementary Table S1; Fig. 3B) and 23.44 ± 18.56% for the INL (Supplementary Table S1; Fig. 4B). These results can be visualized in Bland-Altman plots, as the values for the macro relative to those for the observers presented a smaller bias value and a more grouped and narrower confidence interval than that existing between observers. In addition, Spearman ρ coefficient between observers was 0.90 for the ONL and 0.89 for the INL and 0.90 for the ONL and 0.86 for the INL between observers and macro (Supplementary Table S1; Figs. 3C, 3D, 4C, 4D).

Experimental Testing

To confirm its validation and to further test performance, we sought to determine whether the macro was able to detect statistically significant differences, if present, between experimental groups. We randomly selected 30 images from a control group and 30 from an experimental group and compared results between observers and macro.

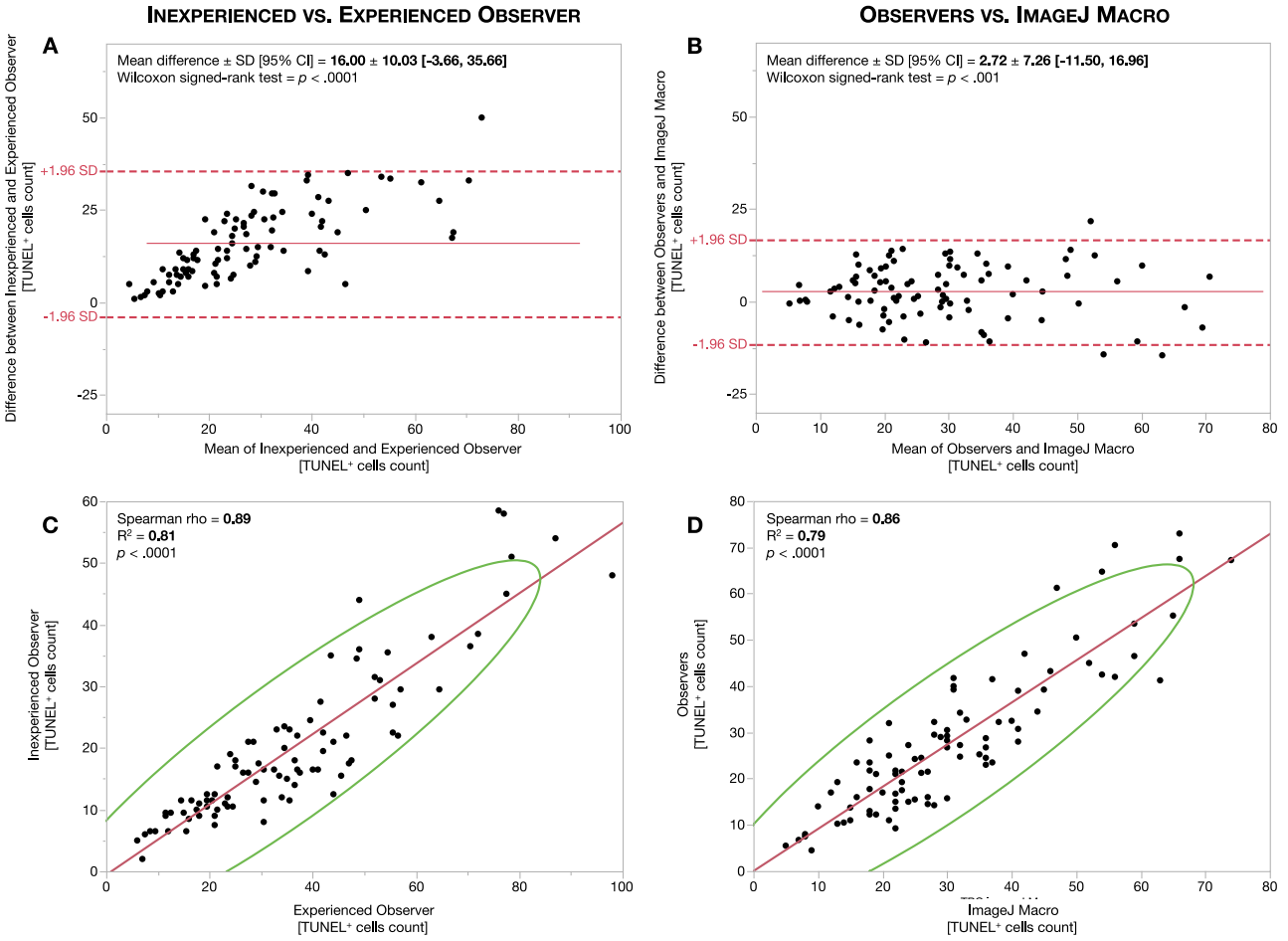


FIGURE 4. Inner nuclear layer interobserver agreement and correlation. Bland-Altman plots for INL measurement agreement between observers (**A**) and between observers and ImageJ macro (**B**). *Red lines* represent mean differences or bias. *Red dashed lines* represent the 95% limit of agreement (±1.96 SD). Correlation analysis between observers (**C**) and between observers and ImageJ macro (**D**). *Red lines* represent the fitted linear regression *trendline*. *Green lines* represent the 95% confidence interval area.

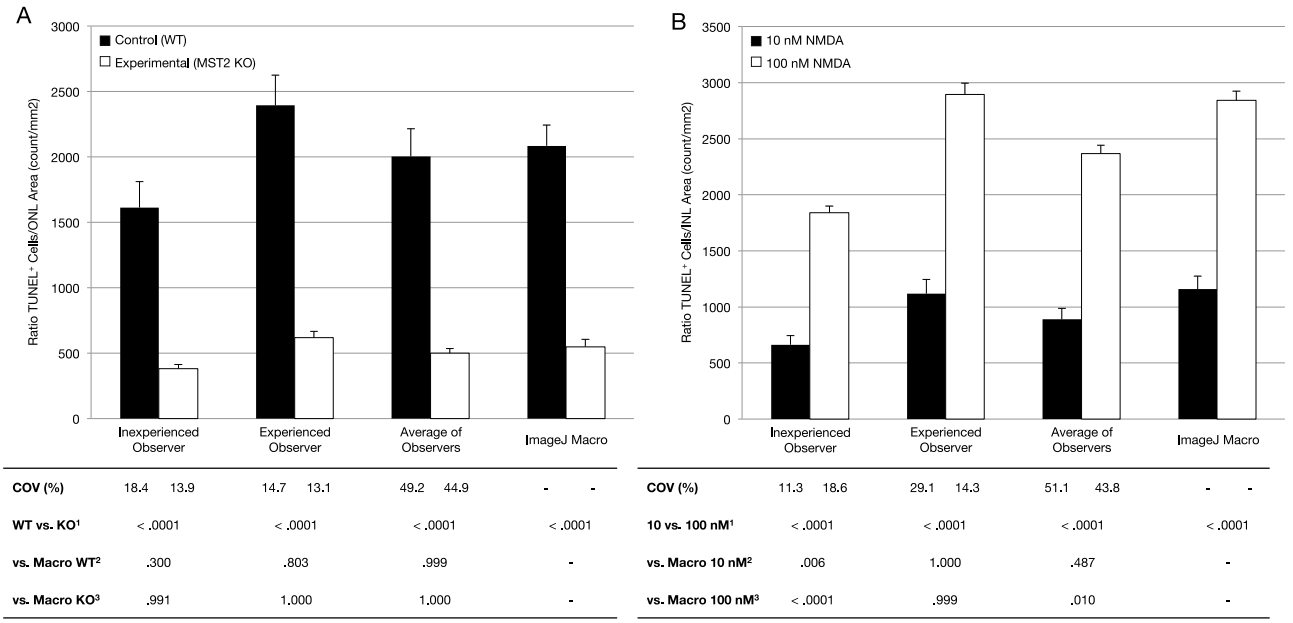
In the ONL dataset (Fig. 5A), all observers and macro were able to detect a statistically significant difference between the experimental group and their respective control ($P < 0.0001$). No statistical differences were observed in control and experimental groups between macro and average of all observers. However, there were small but statistically significant differences in the control values obtained by the different observers ($P = 0.004$) but no statistical significant differences obtained by the different observers in the experimental group ($P = 0.943$).

In the INL dataset (Fig. 5B), all observers and macro were able to detect a statistically significant difference between the two experimental groups ($P < 0.0001$). No statistical differences were observed for both 10 and 100 nM groups between the experienced observer and macro. There was though a statistically significant difference in the absolute values obtained for the 10 nM group ($P = 0.018$) and 100 nM group ($P < 0.0001$) between the two observers.

DISCUSSION

In this study, we developed an ImageJ macro that was able to segment ONL and INL retinal layers and count TUNEL$^+$ cells accurately. The reported parameters were able to fit within the established limits of agreement and were highly correlated with the observers' measurements. In addition, the macro was able to detect statistically significant differences between experimental groups. Collectively, these results validate the macro as a measurement tool for automated cell death quantitation in the retina. We speculate that the advent of this tool can aid in achieving accurate, repeatable, fast, and unbiased measurements.

A moderate degree of variability was observed between repeated measurements in both the inexperienced and experienced observers, as expected. Collectively, both observers presented good reproducibility as shown by bias value and range of limits of agreement. Bland-Altman plots for the inexperienced observer displayed a lower bias value and narrower limits of agreement compared to the experienced observer, most likely due to the time invested in the analysis and higher TUNEL$^+$ cell counts by the experienced observer. Still, both of the observers were able to detect statistically significant differences among experimental groups, and the magnitude of these difference was higher for the experienced observer. In contrast, Spearman correlation coefficient for both of the observers showed similar values between groups. As Bland and Altman previously described, statistical hypothesis tests and correlation coefficients are not suitable statistical



COV (%)	18.4	13.9	14.7	13.1	49.2	44.9	-	-
WT vs. KO[1]	< .0001		< .0001		< .0001		< .0001	
vs. Macro WT[2]	.300		.803		.999		-	
vs. Macro KO[3]	.991		1.000		1.000		-	

COV (%)	11.3	18.6	29.1	14.3	51.1	43.8	-	-
10 vs. 100 nM[1]	< .0001		< .0001		< .0001		< .0001	
vs. Macro 10 nM[2]	.006		1.000		.487		-	
vs. Macro 100 nM[3]	< .0001		.999		.010		-	

Figure 5. Experimental testing of ImageJ macro. From the ONL dataset, 30 images from a WT control group and 30 from an experimental (MST2 knock-out) group were selected (**A**). From the INL dataset, 30 images from a 10 nM NMDA group and 30 from a 100 nM NMDA group were selected (**B**). *Columns* represent mean and standard errors of the mean values for the TUNEL+ cells-to-area per observer ratio. [1]Wilcoxon signed rank test between groups within the same observer. [2]One-way ANOVA with Tukey post hoc correction for comparison between all control groups or 10 nM against the macro. [3]One-way ANOVA with Tukey post hoc correction for comparison between all experimental groups or 100 nM against the macro.

tests for comparing the agreement of two measurement methods. The former only determines whether the means of groups are equal or not. The latter measures the linear dependence between two variables, hence a paralleled change in the measurement magnitude will not affect the correlation coefficient but will certainly affect agreement.[16,18] We suggest that, for the purpose of validating new measurement techniques, Bland-Altman plots and COV should be calculated, and a limit of agreement should be determined beforehand.

The ImageJ macro was able to detect layers and count TUNEL+ cells within the limits of agreement determined by both observers. The coefficients of variation between macro and average of observers were approximately half the value of those between observers, for both ONL and INL dataset. These values are supported by the narrower limits of agreement between macro and average of observers than between both observers, as seen in Bland-Altman plots. These premises argue in favor of the high level of accuracy of this measurement tool.[16] On a side note, TUNEL+ cell count variability was increased in INL images with more TUNEL+ cells, probably owing to the presence of cells of different sizes[19] and staining patterns. Nonetheless, because there is no absolute reference value in TUNEL+ cell counting, we believe it is crucial to follow consistent criteria when counting cells, in order to detect consistent differences between groups.

The burden time to analyze these images is significant. On average, an experienced observer needed more than 2 minutes per image, whereas an inexperienced observer required more than 5 minutes. In contrast, the macro required less than 3 seconds for each image analysis, when executed in batch mode. This advantage opens up the possibility of assessing a larger area of retina. We calculated that the time required to manually count two images per retina would still be higher than analyzing the entire retina with eight to nine images using ImageJ macro, even considering increased acquisition time. The ability to examine larger retina areas in the same amount of time will result in higher experimental objectivity.

While estimating expected sample size for each experimental group, the variability both of the biological model and the measurement tool should be taken into account. If the measurement tool used to quantitate outcome variables has significant variability, it increases the overall variability as bias and limits of agreement are also subject to sampling variation.[20] In essence, higher variability reduces the ability to detect statistical significance, increasing required sample size. Therefore, using an automated measurement tool can reduce bias and sample size.

Image Acquisition, Postprocessing Recommendations, and Limitations

To challenge the macro versatility, we included images taken by multiple observers using different acquisition parameters. However, we recommend a standardized acquisition and postprocessing approach customized for each researcher and equipment used and testing the macro prior to experimental analysis. Images should ideally be centered, with ONL and INL layers not merged, without shadows or uneven focus, to ensure an appropriate macro performance. Image background noise should be reduced as much as possible.

Study Limitations. Still, several limitations should be considered. First, macro results were compared against two observers. However, these observers performed with good reproducibility. Second, the macro assumes the largest nuclei area in the image to be the ONL area. This should be considered when assessing images from experimental neurodegenerative models with significant ONL/INL thinning. Finally, digital images should be acquired with a 20×/0.8 numerical aperture objective, as this is the chosen working magnification of the script for the aforementioned reasons. We are currently working on new macro versions to broaden the type of images suitable for automated analysis.

In summary, the developed and validated ImageJ macro can be an accurate and precise quantitative tool with which retina researchers can achieve repeatable, unbiased, fast, and above all, accurate cell death quantitation. We believe this standardized measurement tool could be advantageous to compare results across different research groups. The macro will be freely available to all as open source.

Acknowledgments

Supported by Bayer Healthcare Global Ophthalmology Awards Program (GOAP; DEM); the Yeatts Family Foundation (DGV, JWM); Loefflers family foundation (JWM,DGV); a 2013 Macula Society Research Grant award (DGV); a Bausch & Lomb Vitreoretinal Fellowship (HM); a Physician Scientist Award (DGV) and an unrestricted grant (JWM) from the Research to Prevent Blindness Foundation; NEI Grant R21EY023079-01A1 (DGV); and NEI Grant EY014104 (MEEI Core Grant). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Please visit http://imagej.net/TUNEL_Cell_Counter to download the macro and original image dataset. We added a detailed tutorial, setup instructions, and suggestions to improve performance.

Disclosure: **D.E. Maidana**, Bayer Healthcare (F); **P. Tsoka**, None; **B. Tian**, None; **B. Dib**, None; **H. Matsumoto**, None; **K. Kataoka**, None; **H. Lin**, None; **J.W. Miller**, None; **D.G. Vavvas**, None

References

1. Cuenca N, Fernández-Sánchez L, Campello L, et al. Cellular responses following retinal injuries and therapeutic approaches for neurodegenerative diseases. *Prog Retin Eye Res*. 2014;43:17–75.
2. Murakami Y, Matsumoto H, Roh M, et al. Programmed necrosis, not apoptosis, is a key mediator of cell loss and DAMP-mediated inflammation in dsRNA-induced retinal degeneration. *Cell Death Differ*. 2014;21:270–277.
3. Murakami Y, Matsumoto H, Roh M. Receptor interacting protein kinase mediates necrotic cone but not rod cell death in a mouse model of inherited degeneration. *Proc Natl Acad Sci U S A*. 2012;109:14598–14603.
4. Johnson TV, DeKorver NW, Levasseur VA, et al. Identification of retinal ganglion cell neuroprotection conferred by platelet-derived growth factor through analysis of the mesenchymal stem cell secretome. *Brain*. 2014;137(pt 2):503–519.
5. Galluzzi L, Kroemer G. Necroptosis: a specialized pathway of programmed necrosis. *Cell*. 2008;135:1161–1163.
6. Gavrieli Y, Sherman Y, Ben-Sasson SA. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J Cell Biol*. 1992;119:493–501.
7. Liming P, Bradley CJ, Liu JJ. The correlativity analysis of six methods of detecting apoptosis. *Chin Med Sci J*. 1999;14:145–151.
8. Grasl-Kraupp B, Ruttkay-Nedecky B, Koudelka H, Bukowska K, Bursch W, Schulte-Hermann R. In situ detection of fragmented DNA (TUNEL assay) fails to discriminate among apoptosis, necrosis, and autolytic cell death: a cautionary note. *Hepatology*. 1995;21:1465–1468.
9. Pasparakis M, Vandenabeele P. Necroptosis and its role in inflammation. *Nature*. 2015;517:311–320.
10. Trichonas G, Manola A, Morizane Y, et al. A novel nonradioactive method to evaluate vascular barrier breakdown and leakage. *Invest Ophthalmol Vis Sci*. 2010;13:317–323.
11. Negoescu A, Guillermet C, Lorimier P, Brambilla E, Labat-Moleur F. Importance of DNA fragmentation in apoptosis with regard to TUNEL specificity. *Biomed Pharmacother*. 1998;52:252–258.
12. Matsumoto H, Miller JW, Vavvas DG. Retinal detachment model in rodents by subretinal injection of sodium hyaluronate. *J Vis Exp*. 2013;79:50660.
13. Matsumoto H, Murakami Y, Kataoka K, et al. Mammalian STE20-like kinase 2, not kinase 1, mediates photoreceptor cell death during retinal detachment. *Cell Death Dis*. 2014;5:e1269.
14. Lunscher WH, Beddoes MP. Optimal edge detector design I: parameter selection and noise effects. *IEEE Trans Pattern Anal Mach Intell*. 1986;8:164–177.
15. Tsai W-H. Moment-preserving thresholding: a new approach. *Comput Vis Graph Image Process*. 1985;29:377–393.
16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
17. Filippi M, Horsfield MA, Rovaris M, et al. Intraobserver and interobserver variability in schemes for estimating volume of brain lesions on MR images in multiple sclerosis. *AJNR Am J Neuroradiol*. 1998;19:239–244.
18. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet*. 1995;346:1085–1087.
19. Chan TF, Chiu K, Lok CKM, et al. Morphometric analyses of retinal sections. *J Vis Exp*. 2012;60:3377.
20. Bland JM, Altman DG. Agreed statistics: measurement method comparison. *Anesthesiology*. 2012;116:182–185.

ThicknessTool: Automated ImageJ Retinal Layer Thickness and Profile in Digital Images.

Daniel E. Maidana[1,2,3], Shoji Notomi[1], Takashi Ueta[1], Tianna Zhou[1], Danica Joseph[1], Cassandra Kosmidou[1], Josep Maria Caminal-Mitjana[3], Joan W. Miller[1], and Demetrios G. Vavvas[1,*]

[1] From the Retina Service, Angiogenesis Lab, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA, USA

[2] From the Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL, USA

[3] From the Retina Service, Bellvitge Hospital, University of Barcelona, Barcelona, Spain.

[*] **Correspondence:** Demetrios G. Vavvas, MD, PhD
Massachusetts Eye and Ear Infirmary
243 Charles Street, Boston MA 02114, United States
Tel: +1 (617) 573-6874
e-mail: Demetrios_Vavvas@MEEI.HARVARD.EDU

Financial Disclosure: The authors have no relevant financial relationships or interests to disclose.

Funding Support: This work was supported by Bayer Healthcare Global Ophthalmology Awards Program (DEM); the Yeatts Family Foundation (DGV, JWM); Loefflers family foundation (JWM, DGV); a Macula Society Research Grant award (DGV); a Physician Scientist Award (DGV) and an unrestricted grant (JWM) from the Research to Prevent Blindness Foundation; NEI Grant R21EY023079-01A1 (DGV); and NEI Grant EY014104 (MEEI Core Grant). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Running Head: Automated Thickness Measurement in ImageJ

Supplementary Tables: 4

Supplementary Figures: 3

Supplementary Code: 1

ThicknessTool: Automated ImageJ Retinal Layer Thickness and Profile in Digital Images.

Daniel E. Maidana[1,2,3], Shoji Notomi[1], Takashi Ueta[1], Tianna Zhou[1,3], Danica Joseph[1], Cassandra Kosmidou[1], Josep Maria Caminal-Mitjana[3], Joan W. Miller[1], and Demetrios G. Vavvas[1,*]

[1] From the Retina Service, Angiogenesis Lab, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, MA, USA

[2] From the Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL, USA

[3] From the Retina Service, Bellvitge Hospital, University of Barcelona, Barcelona, Spain.

[*] **Correspondence:** Demetrios G. Vavvas, MD, PhD

Massachusetts Eye and Ear Infirmary

243 Charles Street, Boston MA 02114, United States

Tel: +1 (617) 573-6874

e-mail: Demetrios_Vavvas@MEEI.HARVARD.EDU

Financial Disclosure: The authors have no relevant financial relationships or interests to disclose.

Funding Support: This work was supported by Bayer Healthcare Global Ophthalmology Awards Program (DEM); the Yeatts Family Foundation (DGV, JWM); Loefflers family foundation (JWM, DGV); a Macula Society Research Grant award (DGV); a Physician Scientist Award (DGV) and an unrestricted grant (JWM) from the Research to Prevent Blindness Foundation; NEI Grant R21EY023079-01A1 (DGV); and NEI Grant EY014104 (MEEI Core Grant). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Running Head: Automated Thickness Measurement in ImageJ

Word count: 5,157 words

Figures: 8

Tables: 8

Abstract

Purpose: To develop an automated retina layer thickness measurement tool for the ImageJ platform, to quantitate nuclear layers following the retina contour.

Materials and Methods: We developed the ThicknessTool (TT), an automated thickness measurement plugin for the ImageJ platform. To calibrate TT, we created a calibration dataset of mock binary skeletonized mask images with increasing thickness masks and different rotations. Following, we created a training dataset and performed an agreement analysis of thickness measurements between TT and two masked manual observers. Finally, we tested the performance of TT measurements in a validation dataset of retinal detachment images.

Results: In the calibration dataset, there were no differences in layer thickness between measured and known thickness masks, with an overall coefficient of variation of 0.00%. Training dataset measurements of immunofluorescence retina nuclear layers disclosed no significant differences between TT and any observer's average outer nuclear layer (ONL) ($p = 0.998$), inner nuclear layer (INL) ($p = 0.807$), and ONL/INL ratio ($p = 0.944$) measurements. Agreement analysis showed that bias between TT vs. observers' mean was lower than between any observers' mean against each other in the ONL (0.77 ± 0.34 µm vs 3.25 ± 0.33 µm) and INL (1.59 ± 0.28 µm vs 2.82 ± 0.36 µm). Validation dataset showed that TT can detect significant and true ONL thinning ($p = 0.006$), more sensitive than manual measurement capabilities ($p = 0.069$).

Conclusions: ThicknessTool can measure retina nuclear layers thickness in a fast, accurate, and precise manner with multi-platform capabilities. In addition, the TT can be customized to user preferences and is freely available to download.

Keywords

Retina

Outer nuclear layer

Inner nuclear layer

Thickness

Photoreceptor

Cell death

Introduction

Progressive photoreceptor cell death is a significant culprit in retinal degenerative diseases. Inasmuch, quantitation of this cell loss has been addressed by a myriad of approaches, predominantly in experimental models of retinal diseases, such as retinal detachment.[1,2] Among these methods, photoreceptor cell death assays[3,4], outer nuclear layer cell counting[5], and outer nuclear layer thickness[6,7], have been used in animal models to quantitate photoreceptor degeneration. Given the technical and time constraints related to cell death assays and manual counting; to date, outer nuclear layer (ONL) thickness has been widely used as a practical proxy method to estimate the depth of photoreceptor cell death.[6–8]

The utility of retina ONL thickness quantitation to infer photoreceptor degeneration relies on manual calliper measurements. However, the advantage of this approach can be compromised as callipers are manually drawn and measured by an observer. In addition, given the curvature of the retina, manually positioning callipers perpendicular to the layer contour along the long axis can be difficult. Furthermore, given the high magnification of microscope images, a small area of interest is often analysed to expedite the analysis, which can lead to bias. As new imaging modalities also become available, a versatile method that can adapt to multiple imaging modalities is ideal. Therefore, a tool that can adapt to layer architecture and contour to measure thickness in a broad segment, in multiple layers, in either single images or large tiles, and in multiple platforms, is very compelling. Despite pioneering work done in this area[9], to the best of our knowledge, there is currently no freely available tool for the ImageJ platform to automatedly quantitate multiple layer thicknesses in large images and in different imaging modalities.

The purpose of this work was to develop an automated retinal layer thickness measurement tool for the ImageJ platform, which can quantitate nuclear layers following the retina contour, with callipers as close to 1-pixel to each other. For this purpose, we developed the ThicknessTool (TT) and validated its accuracy by objective calibration and agreement analysis with two masked observers. We found that this measurement tool can provide accurate and precise thickness measurements. In addition, TT can process images from multiple imaging modalities.

Materials and Methods

All animals used in experiments and breeding adhered to the statement of the Association for Research in Vision and Ophthalmology (ARVO). Animal protocols were reviewed and approved by the Animal Care Committee of the Massachusetts Eye and Ear Infirmary. To validate the ThicknessTool in experimental model, a retinal detachment was induced in eight-week-old C57BL/6J mice, as previously described.[10] C57BL/6J mice were purchased from The Jackson Laboratories and maintained in a standard 12-hour light/dark cycle.

Digital Image Dataset

Calibration Dataset. We created an image dataset of fifty 8-bit mock images of 1344 x 1024 pixels containing masks of increasing thickness, from 10 to 500-pixel thickness with 10-pixel increments. Following, we created a second dataset of images with a known mask area of 1344 x 200 pixels with 10º rotation intervals, plus 45º and 135º rotations for a total of 21 images. Finally, to emulate layer thinning and thickening, we created a single mock image containing a mask with abrupt thickness changes, with segments of 200, 150, and 250 pixels. All images were stored in TIFF format, and the measured thickness was compared to the known standard.

Calibration Criteria. We defined the overall mean thickness limit of agreement to 1 pixel, and the following criteria for the ThicknessTool script calibration: i) overall mean thickness equal to 200 ± 1 pixel; ii) index of dispersion equal to 0, as the given area is a rectangle and all callipers should be parallel to each other with equal thickness; iii) single calliper measurement at 0º, 45º, 90º, 135º, and 180º rotations equal to 200 pixels; and iv) single calliper measurement equal to 200 pixels ± 2 pixels at the remaining rotations. This difference corresponds to the aliasing effect on area edges at rotations other than the specified on iii). The occurrence of uneven jagged edges requires a broader margin of error (2 pixels) than images with straight (0º, 90º, 180º) or even (45º and 135º) edges. For comparisons, the measured mean was tested against a hypothesized mean of 200 pixels. An equivalence test with a two one-sided tests (TOST) approach was used to assess minimum and maximum single calliper measurements, and the threshold difference considered equivalent to no difference was defined as 1 for images with even edges and 2 for images with uneven edges.

Training Dataset. Immunofluorescence digital images of 4′,6-diamidino-2-phenylindole (DAPI)-stained retinal

cross-sections were obtained from the Angiogenesis Laboratory fluorescence microscopy database. We selected images from a murine retinal detachment model in order to evaluate TT performance with different ONL thicknesses and various retinal distortions. A total of 64 randomly selected images were included in this dataset and stored in TIFF format for further processing. Training dataset images were measured by an inexperienced observer who never performed the given task before, and by an experienced observer who had performed these measurements before. Images of murine retina displaying the outer (ONL) and inner nuclear layer (INL) were masked and given to observers. Each image was measured twice in a masked manner. Observers measured each layer with six individual callipers, placed across the retinal layer at their discretion. Single callipers were used to calculate mean, minimum, and maximum layer thickness.

Validation Dataset. We randomly selected 16 images from 8 eyes of an experimental murine retinal detachment model. The layer thickness was manually assessed by an observer, who measured ONL and INL in the detached and attached retina, with six individual callipers for each layer.

Thickness Script

We developed a script for the ImageJ platform (version 1.52p, http://imagej.nih. gov/ij/; provided in the public domain by the National Institutes of Health, Bethesda, MD, USA. To facilitate the use of the script, we designed a graphical user interface dialog to tailor performance parameters to users' preferences, including image scale, number of callipers, skeleton width, among others (Supplementary Figure 1).

Briefly, retinal layers were segmented using a previously validated image processing algorithm (Figure 1A-D).[4] From segmented layers, a skeletonized mask was created with AnalyzeSkeleton function, as previously described by Arganda-Carreras and coauthors.[11] Following, thickness callipers were constructed as follows (Figure 1D): i) The segmented mask skeleton was fragmented in segments by user-defined *n* number of callipers; ii) Centroids from these segmented skeletons were isolated as single points and defined by the coordinates (X_0, Y_0); iii) The skeleton segment angle was calculated by an user-defined grid which evaluated the angle of the segmented mask skeleton in a given area; iv) From each centroid (X_0, Y_0), a positive *final* vector and a negative *initial* vector were created perpendicularly to the segmented skeleton. The *initial* and *final* vectors were composed by a

corresponding horizontal component vector (X_{step}), and a vertical component vector (Y_{step}); v) The X_{step} and Y_{step} components were created for the *initial* and *final* resolved vector at a 1-pixel distance from the centroid in horizontal and vertical directions, respectively. The preliminary *initial* (X_i, Y_i) and *final* (X_f, Y_f) coordinates were created from the centroid as described in Equation 1; vi) Pixel values were evaluated in the retinal layer mask at *initial* (X_i, Y_i) and *final* (X_f, Y_f) coordinates. If the pixel value returned *zero*, this was interpreted as the retinal layer mask, and the X_{step} and Y_{step} components were increased by 1-pixel, to analyse the contiguous pixel; vii) This process was iterated until the pixel value was > 0, and this point was assumed as the *initial* (X_i, Y_i) and *final* (X_f, Y_f) vector endpoint; viii) Thickness was calculated as the rounded Euclidean distance between *initial* (X_i, Y_i) and *final* (X_f, Y_f), as seen in Equation 2.

$$(1) \quad X_i = round\left(\sin\left(\alpha \times \frac{\pi}{180}\right) \times X_{step} + X_0\right)$$

$$Y_i = round\left(Y_0 - \cos\left(\alpha \times \frac{\pi}{180}\right) \times Y_{step}\right)$$

$$X_f = round\left(X_0 - \sin\left(\alpha \times \frac{\pi}{180}\right) \times (X_{step} - \sqrt{1 + {Y_{step}}^2})\right)$$

$$Y_f = round\left(\cos\left(\alpha \times \frac{\pi}{180}\right) \times (Y_{step} - \sqrt{1 + {Y_{step}}^2}) + Y_0\right)$$

$$(2) \quad Thickness\ Caliper = \sqrt{(X_f - X_i)^2 + (Y_f - Y_i)^2}$$

Where *i* = initial; *f* = final; α = skeleton supplementary angle; X_0 = x-axis centroid coordinate; Y_0 = y-axis centroid coordinate; X_i = *initial* horizontal vector component endpoint x-axis coordinate; Y_i = *initial* vertical vector component endpoint y-axis coordinate; X_f = *final* horizontal vector component endpoint x-axis coordinate; Y_f = *final* vertical vector component endpoint y-axis coordinate; X_{step} = 1-pixel increment in the x-axis; Y_{step} = 1-pixel increment in the y-axis.

Statistical Analyses

Statistical analysis was performed with SAS software (2019 SAS, NC). Normality was assessed with Shapiro-Wilk test. Statistical significance for differences between groups was determined with t-test for matched

pairs or Wilcoxon signed-rank test for matched pairs, t-test for independent samples, and one-way ANOVA with Tukey *post hoc* correction for multiple comparisons. For correlation analysis, values were expressed as Pearson correlation coefficient (*r*) and statistical significance. Parallelism plots were constructed fitting a cubic spline with a default lambda of 0.05. Agreement analysis was performed by Bland-Altman plot method[12,13], and as previously described.[4] Coefficient of variation (CoV) was calculated as the difference between measurements over their mean [(A – B)/(A + B)/2] and CoV% as CoV*100. The index of dispersion was calculated as variance-to-mean ratio. Results are expressed as mean ± standard deviation (SD). A p value of < 0.050 was considered statistically significant.

Results

ThicknessTool Calibration

To calibrate the ThicknessTool algorithm, we created three mock image datasets with masks of known area thickness and rotation. First, the TT was able to accurately measure areas of increasing thickness, from 10 to 500 pixels, with a bias of 0.00 ± 0.00 (Supplementary Figure 2) and thus, as expected, the correlation coefficient for the known vs. measured thickness was 1.00 ($p < 0.001$). Following, we evaluated the performance of the algorithm at different layer rotations with 200-pixel masks at different angles. We predefined the known or theoretical mean, minimum, and maximum thickness to 200 pixels, and a coefficient of variation (CoV) of 0.00, and a margin of error of 1-pixel. As seen in Supplementary Table 1, the mean thickness measured by ThicknessTool was 199.88 ± 0.25 pixels, with a CoV of 0.00 for the overall sample. There were no statistically significant differences between the measured overall mean thickness and hypothesized mean ($p = 0.06$). Mean thickness for images with even and jagged edges fulfilled the calibration criteria as previously defined, with CoV of 0.00 (Supplementary Figure 2A-D). Equivalence Test of mean, minimum and maximum calliper thickness rejected both null hypotheses (both $p < 0.001$), indicating that the observed difference does not exceed the defined threshold. Finally, we tested TT capability to detect true thinning and thickening in a mock image with decreasing and increasing thickness, respectively. TT accurately measured varying thickness ($r = 1.00$, $p < 0.001$), with a bias of 0.00 ± 0.00 and no significant difference between known vs. measured markers ($p = 1.000$) (Supplementary Figure 2F). These results indicate that the ThicknessTool can automatedly measure the different thicknesses and at different rotations in an accurate and precise manner.

Training Dataset Mean Layer Thickness

To evaluate the performance of observers and TT across different images and thickness, we first constructed a parallelism plot of mean ONL and INL thickness. As seen in Figure 2, the inexperienced observer's mean measurements were consistently higher than the experienced and TT, which in return were similar and parallel to each other in both ONL and INL layers, suggesting a proportionate performance. Correlation coefficient (r) analysis between TT and observers' measurements were > 0.88 ($p < 0.001$) for the ONL and > 0.84 ($p < 0.001$) for the INL (Supplementary Table 2 and 3).

The mean ONL thickness was 65.26 ± 7.78 µm, 62.00 ± 7.14 µm, and 62.85 ± 7.33 µm for the inexperienced, experienced, and TT, respectively (Table 1). There were no significant differences between TT and any observer or average ONL measurements. The mean INL thickness was 39.13 ± 6.22 µm, 36.31 ± 5.17 µm, 36.12 ± 4.72 µm for the inexperienced, experienced and TT, respectively. In this group, there was only a significant difference between the inexperienced first measurement and TT (39.23 ± 6.30 µm vs. 36.12 ± 4.72 µm, $p = 0.044$), and no significant differences between TT and any observer average INL measurements. The average for the inexperienced vs. experienced was not significantly different, in either the ONL ($p = 0.219$) and INL ($p = 0.097$), with overlapping confidence intervals. Moreover, there were no significant differences between TT and any observer or average ONL/INL ratio measurements, with overlapping 95% confidence intervals. Altogether, these results suggest that no significant difference was evidenced regarding the observers' prior experience. However, TT ONL and INL measurements stand more closely to experienced observer's results. Moreover, TT measurements are not significantly different from any observers' ONL and INL average.

Training Dataset Minimum and Maximum Layer Thickness

Qualitative assessment of observers' thickness callipers showed no overlapping in repeated measurements, suggesting that measurements are not likely reproducible, as seen in Supplementary Figure 3. Moreover, calliper over and undershoot was observed, which can lead to over and under measurement, respectively. In addition, the thickness vectors were not always perpendicular to the layer angle, hence prone to spurious results. Regardless of the lack of statistical differences in mean ONL and INL thickness between observers' average and TT measurements, we analysed the minimum and maximum calliper values for each image.

First, we constructed a parallelism plot of the minimum and maximum ONL and INL thickness calliper per image and observer. As seen in Figure 3, the experienced observer had lower maximum and higher minimum ONL and INL values, without curve overlapping. Similarly, the inexperienced observer had higher minimum ONL and INL values without curve overlapping. These results suggest that experienced observers' measurements are not reproducible. In addition, there were no significant differences between TT and any observers' overall ONL minimum and maximum measurements (Table 2). However, we found significant differences across individual

observer's minimum and maximum measurements and between the inexperienced overall minimum INL and TT measurements ($p < 0.001$). In addition, the coefficient of variability between measurements was significantly different between inexperienced and observers' average measurements for both ONL and INL ($p < 0.001$) and between the experienced observer and observers' average measurements for the ONL ($p < 0.009$) (Supplementary Table 4). In summary, given the qualitative assessment of observers' callipers' lack of repeatability together with the quantitative analyses between manual callipers and a calibrated automated algorithm, results indicate that manual callipers measurements are not accurate nor reproducible.

Agreement Analysis between Manual and Thickness Measurements

We performed a bias analysis using Bland and Altman method.[13] For the ONL group, as seen in Table 3, we found statistically significant differences when comparing paired measurements, with a greater bias in inexperienced mean vs. experienced mean measurements (3.25 ± 0.33 µm, $p < 0.001$). In contrast, we found the lowest bias when comparing TT vs. observers' means (0.77 ± 0.34 µm, $p = 0.028$). For the INL group, we found statistically significant differences when comparing inexperienced mean vs. experienced mean ($p < 0.001$), TT vs. inexperienced mean ($p < 0.001$), and TT vs. observers' mean ($p < 0.001$). Most importantly, the bias between TT vs. observers' INL mean (1.59 ± 0.28 µm) was lower than the those between any observers mean against each other (2.82 ± 0.36 µm). In addition, Bland-Altman plots displayed random variability within measurements (Figure 4). In conclusion, these results indicate that ThicknessTool can measure ONL and INL with lower bias than observers' average against each other for both ONL and INL.

Validation of Thickness in a Retinal Detachment Model

To definitively validate the TT, we tested the tool in an experimental setting using a retinal detachment model, which causes photoreceptor cell death and subsequent ONL thinning and distortion (Figure 5). We quantitated the ONL and INL thickness manually and using TT (Table 4). Interestingly, manual measurements failed to show significant ONL thinning in the detached retina ($p = 0.069$), in comparison to TT ($p = 0.006$). No differences were observed in the INL between the attached and detached retina by both methods as expected. Both methods showed significant differences in the ONL/INL ratio in the detached retina. These results suggest that TT

can be more sensitive and detect significant true thinning beyond manual measurement capabilities. Collectively, these results suggest that TT is a sensitive, precise, and reliable tool to measure retinal nuclear layers in experimental models of disease.

Application of Thickness in Digital Images

In order to examine the applicability of TT in different image modalities, we quantitated the thickness in various digital images. As seen in Figure 6, TT was capable of quantitating the ONL thickness and corresponding profile in histology sections. Moreover, we challenged this tool in digital images of macular spectral-domain optical coherence tomography, with satisfactory results. Finally, we tested the performance in the quantitative assessment of retinal vessels in fluorescein angiogram images. The TT was able to draw accurate callipers with the correct vector in all cases. Collectively, these results indicate that ThicknessTool is a versatile tool with multi-platform capabilities.

Discussion

In this study, we developed a sensitive, precise, and reliable tool to measure retinal layers. The performance of this tool was tested in calibration, training, and validation datasets with adequate and reliable performance. Collectively, results validate ThicknessTool for automated retinal thickness measurement. We speculate that this tool can aid in achieving precise, accurate, and fast measurements in multi-platform digital images.

In the development of new measurement tools, calibration images are mandatory to benchmark and finetune the performance of image processing algorithms to a standard.[14] We designed mock images with various known thickness ranging from 10 pixels (given a single nuclei size in our images) to 500 pixels (twice the outer nuclear layer measured in our dataset). In addition, we created a dataset with multiple mask rotations to mimic the random mask orientations. This step is necessary as TT callipers are constructed in a vectorial manner, in number and angles as the user or region of interest requires. Moreover, we tested TT capability to detect true thinning and thickening in a mock image with decreasing and increasing thickness, respectively. We found that the TT can automatedly and accurately measure thickness when challenged with different thicknesses and rotations.

Accurate assessment of photoreceptor cell death currently relies on manual counting by masked observers. Ideally, two masked two observers quantitate retina thickness, and their average is used to determine the overall outcome. We found that single observers' measurements were prone to over and underestimation, regardless of their previous experience. In addition, observers draw incorrectly angled callipers, which can indirectly contribute to the abovementioned bias. Moreover, callipers were not located in corresponding locations in repeated same-image analysis, which dampens the reproducibility of the measurements. Altogether, this leads to a lack of accurate or reproducible retina thickness quantitation. This matter is of critical importance, as this bias can lead to α or β errors. For example, finding no difference in retinal thickness in cases where a neuroprotective treatment is effective, or even worse, finding a significant difference in retinal thickness in an ineffective treatment.[15,16] Such errors are not only misleading to the scientific community but also high priced.

Regardless of the performance of manual measurements to date, it remains the gold standard. For purposes of method agreement, we refrained from relying on mean comparison or correlation, as Bland and Altman have previously reported in their seminal work.[12,17] As authors describe, such agreement must be quantified, and cannot

be assessed by mean comparison or hypothesis testing, as it does not exist in a present or absent binary form.[13] In addition, high correlation does not necessarily imply a high agreement between two methods, as it quantifies the degree to which two methods are related.[18] Agreement analysis showed lower bias of TT against observers than observers' average against each other for both ONL and INL. Taken into consideration TT capability to detect true mask thinning and thickening, collectively, these results confer to TT a higher measurement agreement.

Moreover, the automated thickness analysis by TT yields a series of convenient metrics that remain hidden otherwise in manual measurements. On this note, Byun et al. have developed pioneering methods to estimate the ONL thickness exclusively.[9] However, their approach was not validated by thickness agreement analysis with manual observers. In their work, authors also described a method to quantify structural distortions in retinal tissue before and after injury, also known as distortion index. In essence, the authors stated that structural changes in the ONL are directly proportional to the distortion index. In our work, we estimated overall distortion in a similar approach using the index of dispersion (ID), calculated as the variance-to-mean ratio, as previously defined by Cox and Lewis.[19] Similarly, the index of dispersion allows the assessment of the overall calliper variation, which can provide detailed layer architecture data. Briefly, an ID of 0 represents no dispersion; therefore, we used the ID for calibration purposes, as mock images of known and constant thickness were uniform rectangles (ID = 0). We speculate that the index of dispersion can be equally used for further morphometric assessment.

Several limitations to this work should be considered. First, despite the calibration analysis performed beforehand, the validation of TT can only be performed against manual measurements. This process is cumbersome, as observers' measurements carry an intrinsic bias. However, observers showed comparable performance, with overlapping confidence intervals. Second, TT analysis should be performed in areas distant from image edges, as skeleton are trimmed to avoid calliper-to-end image measurements. Despite the fact that TT allows skeleton and edge trimming, users should evaluate processed images for qualitative control. Third, TT measures thickness in binary segmented masks. Therefore, its performance is as strong as the segmentation algorithm used. In our case, we build this tool based on a previously validated algorithm for retina layer segmentation.[4] We recommend users to first tailor and validate their segmentation algorithm for optimal performance. Finally, regardless of the potential image modality application, this tool has been validated only for ONL and INL measurements in DAPI nuclei-stained

murine retina cryosections. We believe future work can help expand the applicability of this tool for three-dimensional images, including upcoming imaging platforms.

In summary, we developed a fast, accurate, and precise retinal thickness measurement tool with multi-platform capabilities. In addition, the TT can be customized to user-defined preferences. This includes calliper number with a 1-pixel minimum distance between callipers, skeleton and edge trimming, and automated or manual layer segmentation. In addition, given the potential use with multi-platform capabilities, the source code can be easily modified to fit different applications. Finally, this tool is freely available to download and use. We expect this measurement tool could improve the outcomes and reduce bias in retina research.

Acknowledgements

Authors contributions: D.E.M and D.G.V. designed research. D.E.M. coded the ThicknessTool, performed the statistical analysis, and drafted the manuscript. S.N., T.U., T.Z., D.J., and C.K. measured retinal thickness from the digital image dataset and contributed to the manuscript. J.W.M., and D.G.V. reviewed the data and made substantial intellectual contributions to the research design and manuscript. Please visit https://github.com/DanielMaidana/ThicknessTool to download the ThicknessTool. ThicknessTool is licensed under the GNU General Public Licence.

References

1. Cook, B., Lewis, G. P., Fisher, S. K. & Adler, R. Apoptotic photoreceptor degeneration in experimental retinal detachment. *Invest Ophthalmol Vis Sci* **36**, 990–6 (1995).

2. Arroyo, J. G., Yang, L., Bula, D. & Chen, D. F. Photoreceptor apoptosis in human retinal detachment. *Am. J. Ophthalmol.* **139**, 605–10 (2005).

3. Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J. Cell Biol.* **119**, 493–501 (1992).

4. Maidana, D. E. *et al.* A Novel ImageJ Macro for Automated Cell Death Quantitation in the Retina. *Invest Ophthalmol Vis Sci* **56**, 6701–6708 (2015).

5. Byun, J. *et al.* Automated tool for the detection of cell nuclei in digital microscopic images: Application to retinal images. *Mol. Vis.* **12**, 949–960 (2006).

6. Trichonas, G. *et al.* Receptor interacting protein kinases mediate retinal detachment-induced photoreceptor necrosis and compensate for inhibition of apoptosis. *Proc Natl Acad Sci U S A.* **107**, 21695–21700 (2010).

7. Matsumoto, H. *et al.* Strain difference in photoreceptor cell death after retinal detachment in mice. *Invest Ophthalmol Vis Sci* **55**, 4165–4174 (2014).

8. Murakami, Y. *et al.* Receptor interacting protein kinase mediates necrotic cone but not rod cell death in a mouse model of inherited degeneration. *Proc Natl Acad Sci U S A.* **109**, 14598–14603 (2012).

9. Byun, J. *et al.* Quantifying structural distortions in retinal tissue before and after injury. in *Workshop on Multiscale Biological Imaging, Data Mining & Informatics, Santa Barbara, CA, USA* 10–11 (2006).

10. Matsumoto, H., Miller, J. W. & Vavvas, D. G. Retinal Detachment Model in Rodents by Subretinal Injection of Sodium Hyaluronate. *J. Vis. Exp.* (2013) doi:10.3791/50660.

11. Arganda-Carreras, I., Fernández-González, R., Muñoz-Barrutia, A. & Ortiz-De-Solorzano, C. 3D reconstruction of histological sections: Application to mammary gland tissue. *Microsc. Res. Tech.* **73**, 1019–29 (2010).

12. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–10 (1986).

13. Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–60 (1999).

14. Troscianko, J. & Stevens, M. Image calibration and analysis toolbox - a free software suite for objectively measuring reflectance, colour and pattern. *Methods Ecol. Evol.* **6**, 1320–1331 (2015).

15. Kim, H.-Y. Statistical notes for clinical researchers: Type I and type II errors in statistical decision. *Restor. Dent. Endod.* **40**, 249 (2015).

16. Rothman, K. J. Curbing type i and type II errors. *Eur. J. Epidemiol.* **25**, 223–224 (2010).

17. Bland, J. & Altman, D. Measuring agreement in method comparison studies with heteroscedastic measurements. *Stat. Methods Med. Res.* **8**, 135–160 (1999).

18. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Medica* **25**, 141–151 (2015).

19. Cox, D. & Lewis, P. *The Statistical Analysis of Series of Events*. (Springer Netherlands, 1966).

Tables

Figure Legends

Figure 1. Outline of automated segmentation, thickness calliper constructs, and thickness measurements by ThicknessTool. (A) Representative native image acquired from mouse retina section and immunostaining. (B) Segmented outer nuclear layer (ONL) and mask. (C) Cropped area of the outer nuclear layer (ONL) mask. The angle of each skeleton segment (θ) was calculated respective to the zero-plane. The calliper angle (α), and resulting vector, were calculated as the supplementary angle of (θ). (D) Digital image representation displaying a pixel-grid and averaged skeleton segment calculated for the ONL mask, resulting vectors, and calliper endpoints.

Figure 2. Parallelism plot of mean retinal thickness measurements of observers and ThicknessTool. Results were fitted to a cubic spline with a default lambda of 0.05. (A) Outer nuclear layer. (B) Inner nuclear layer. Dotted lines represent the first and second measurements by an inexperienced observer. Dashed lines represent the first and second measurements by an experienced observer. Solid lines represent measurements by ThicknessTool. Images are sorted by ThicknessTool values in ascending order. (n = 64)

Figure 3. Parallelism plot of minimum and maximum retinal thickness measurements of observers and ThicknessTool. Results were fitted to a cubic spline with a default lambda of 0.05. (A) Outer nuclear layer. (B) Inner nuclear layer. Dotted lines represent the first and second measurements by an inexperienced observer. Dashed lines represent the first and second measurements by an experienced observer. Solid lines represent measurements by ThicknessTool. (n = 64)

Figure 4. Agreement analysis of observers and ThicknessTool by Bland-Altman plots. Agreement between inexperienced and experienced average mean (A) ONL and (C) INL thickness. Agreement between observers and ThicknessTool (B) ONL and (D) INL thickness. (n = 64) SD = standard deviation; LoA = limit of agreement.

Figure 5. Validation of ThicknessTool in a retinal detachment model. (A) Representative image of mouse retina section of a retinal detachment displaying the attached (dotted lines box) and detached retina (solid lines box). (B) Representative image of ThicknessTool measurement of detached retina inner nuclear layer (INL) and outer nuclear layer (ONL) at 1-pixel interval callipers. (C) Outer nuclear layer (ONL) and (D) Inner nuclear layer (INL) thickness measurements from observer manual callipers and ThicknessTool in the attached and detached retina. (E)

ONL/ONL+INL ratio measurements of observer manual callipers and ThicknessTool. (n = 8 per group, *p ≤ .05)

Figure 6. ThicknessTool application in digital images. (A-B) Haematoxylin-eosin staining of retinal cryosection. (C-D) Macular spectral-domain optical coherence tomography scan showing a pigment epithelial detachment in age-related macular degeneration. (E-F) Fluorescein angiogram of retinal vessels.

Figure 1

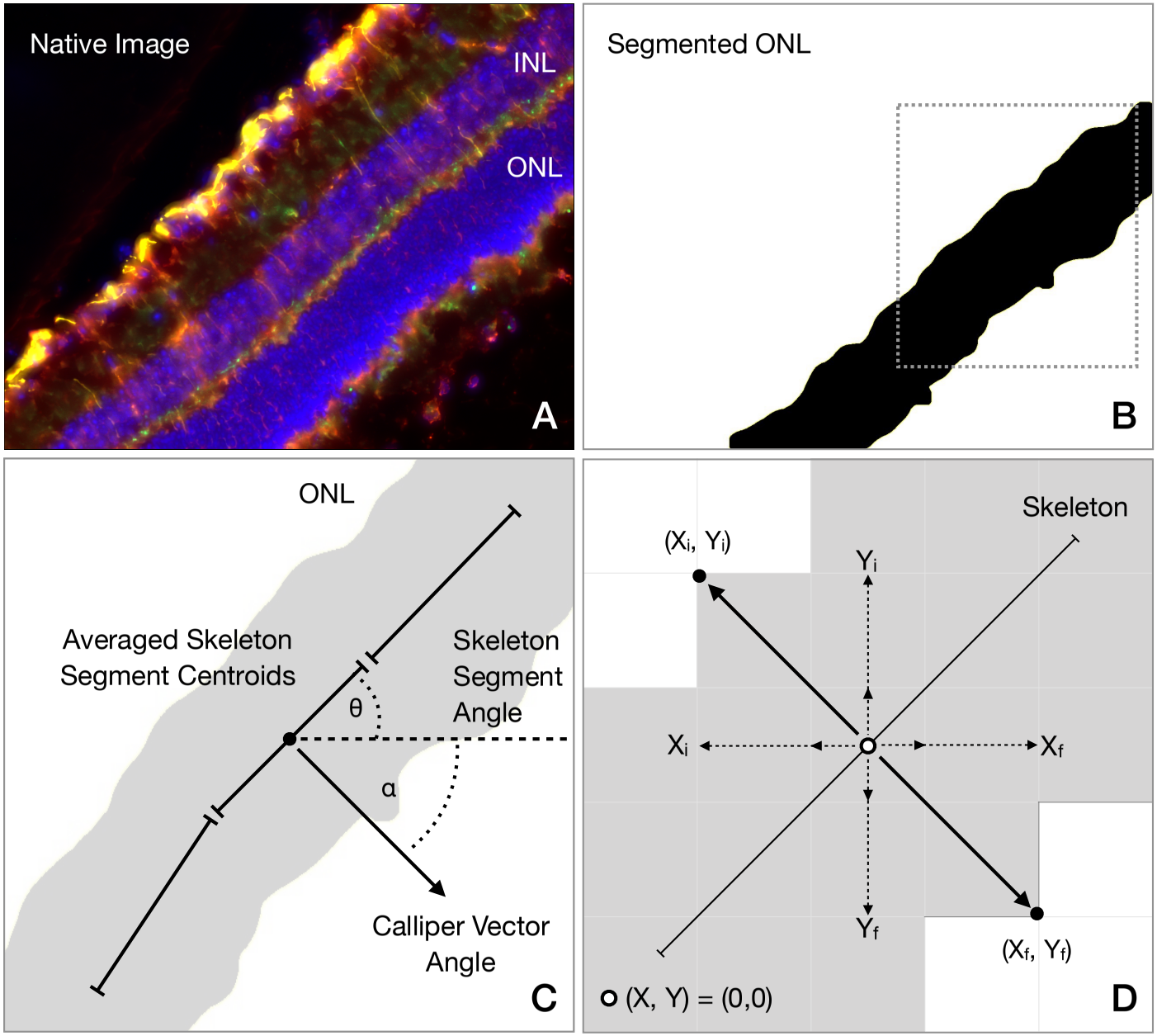

Native Image
INL
ONL
A
Segmented ONL
B
ONL
Averaged Skeleton
Segment Centroids
Skeleton
Segment
Angle
θ
α
Calliper Vector
Angle
C
Skeleton
(Xi, Yi)
Yi
Xi
Xf
Yf
(Xf, Yf)
(X, Y) = (0,0)
D

Figure 2

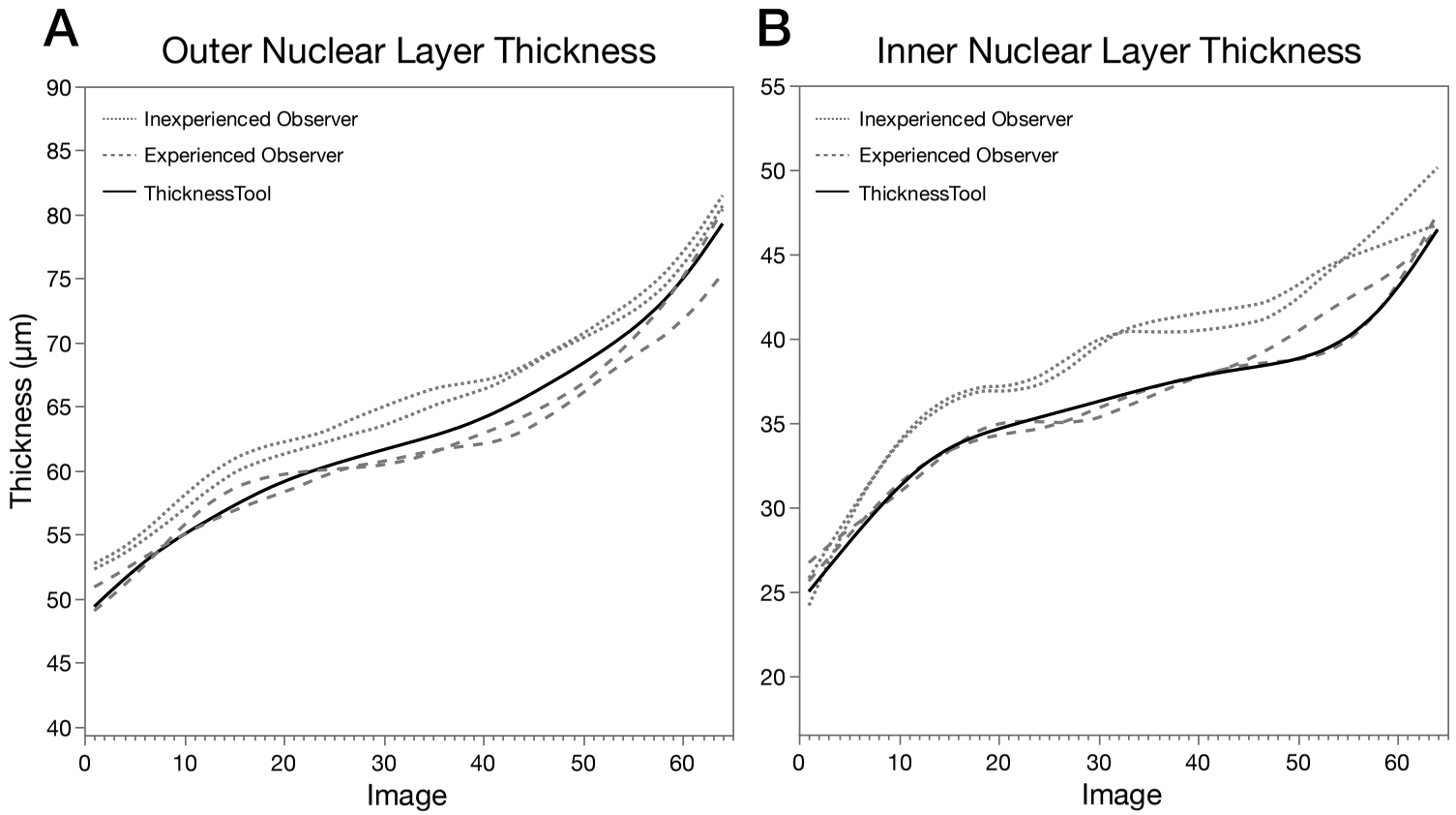

A
Outer Nuclear Layer Thickness
Inexperienced Observer
Experienced Observer
ThicknessTool
Thickness (µm)
Image
B
Inner Nuclear Layer Thickness
Inexperienced Observer
Experienced Observer
ThicknessTool
Image

Figure 3

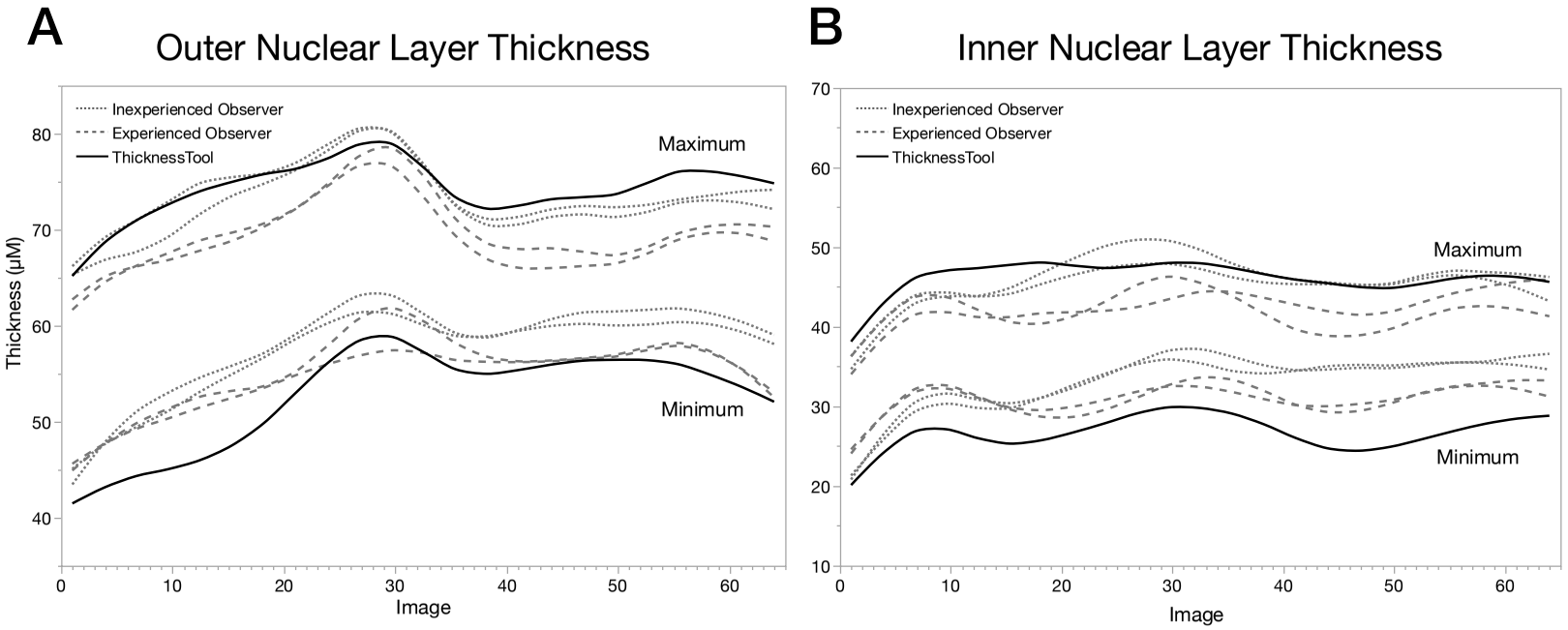

A
Outer Nuclear Layer Thickness
Inexperienced Observer
Experienced Observer
ThicknessTool
Maximum
Minimum
Thickness (µM)
Image
B
Inner Nuclear Layer Thickness
Inexperienced Observer
Experienced Observer
ThicknessTool
Maximum
Minimum
Image

Figure 4

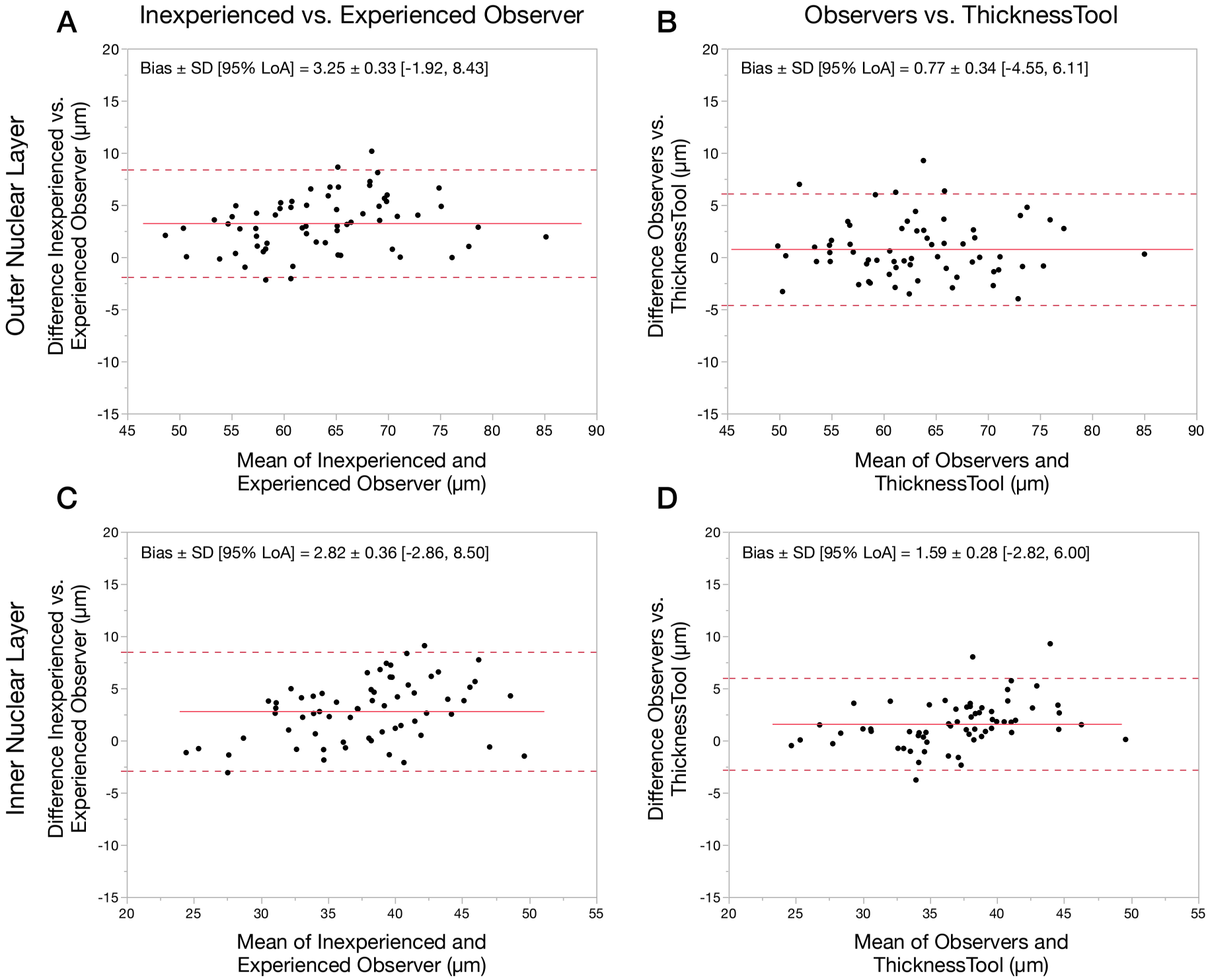

A
Inexperienced vs. Experienced Observer
Bias ± SD [95% LoA] = 3.25 ± 0.33 [-1.92, 8.43]
Outer Nuclear Layer
Difference Inexperienced vs. Experienced Observer (µm)
Mean of Inexperienced and Experienced Observer (µm)
B
Observers vs. ThicknessTool
Bias ± SD [95% LoA] = 0.77 ± 0.34 [-4.55, 6.11]
Difference Observers vs. ThicknessTool (µm)
Mean of Observers and ThicknessTool (µm)
C
Bias ± SD [95% LoA] = 2.82 ± 0.36 [-2.86, 8.50]
Inner Nuclear Layer
Difference Inexperienced vs. Experienced Observer (µm)
Mean of Inexperienced and Experienced Observer (µm)
D
Bias ± SD [95% LoA] = 1.59 ± 0.28 [-2.82, 6.00]
Difference Observers vs. ThicknessTool (µm)
Mean of Observers and ThicknessTool (µm)

Figure 5

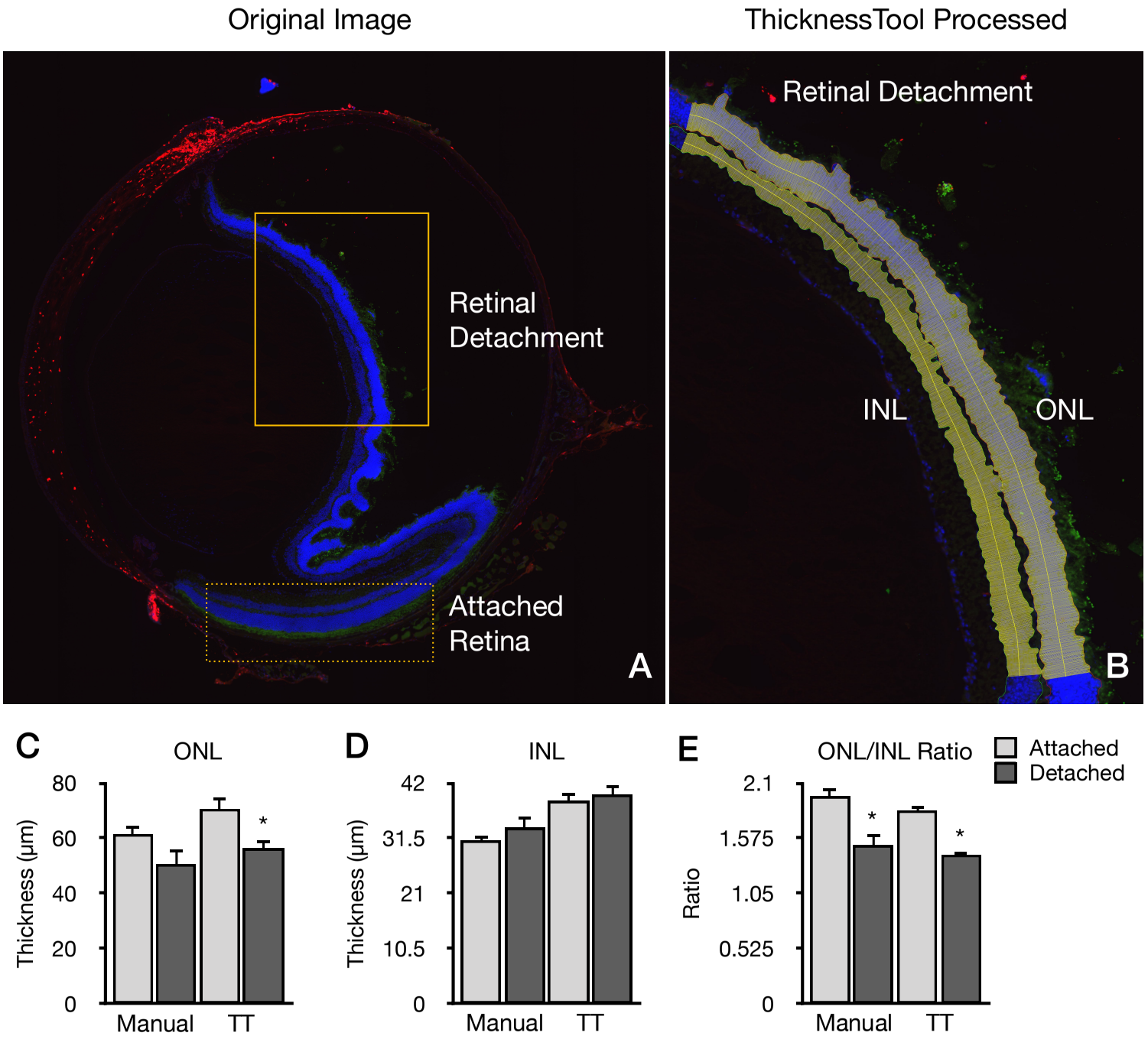
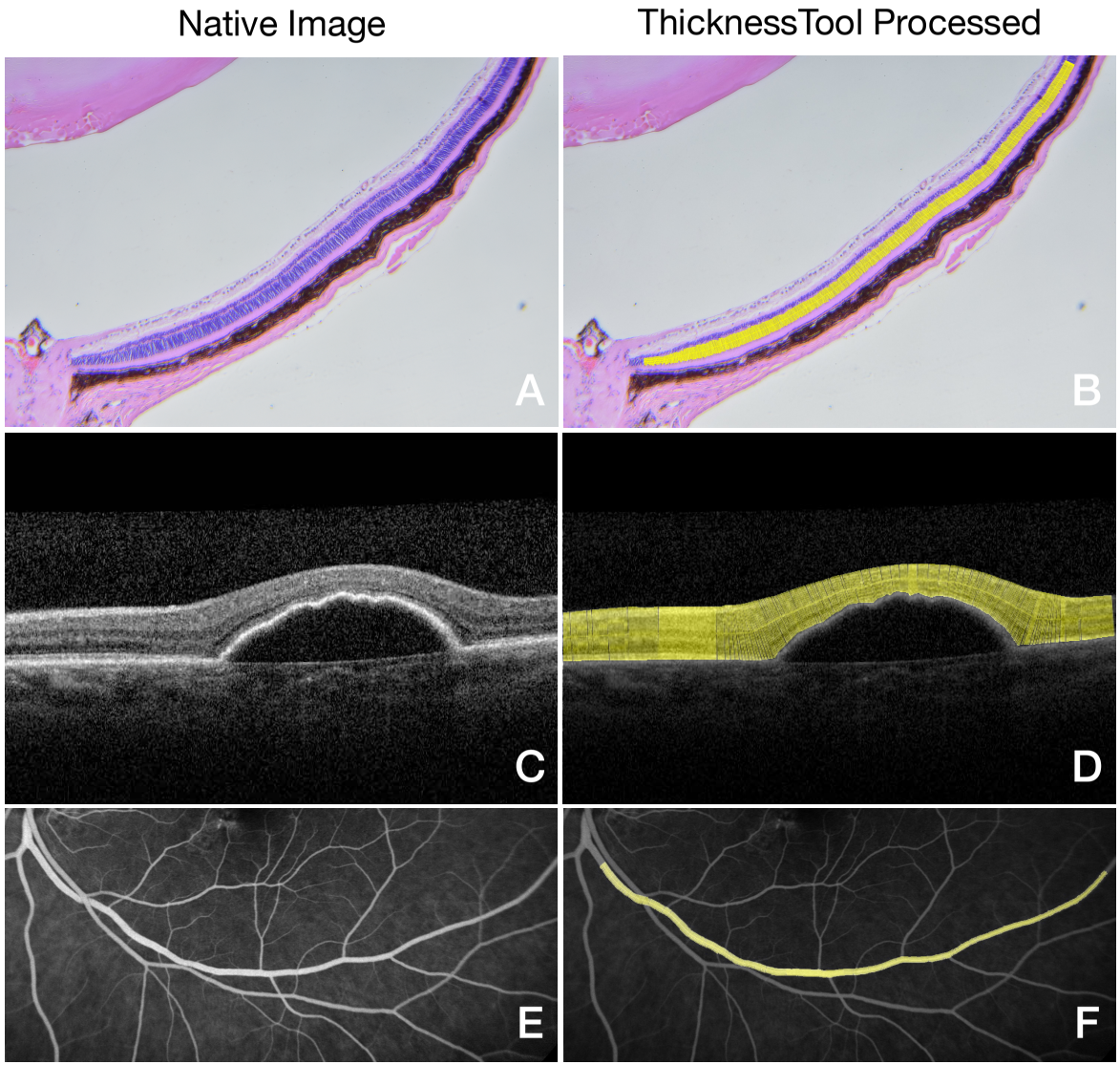

Original Image
ThicknessTool Processed
Retinal Detachment
Attached Retina
A
Retinal Detachment
INL
ONL
B
C
ONL
Thickness (µm)
80
60
40
20
0
Manual
TT
*
D
INL
Thickness (µm)
42
31.5
21
10.5
0
Manual
TT
E
ONL/INL Ratio
Attached
Detached
Ratio
2.1
1.575
1.05
0.525
0
Manual
TT
*
*

Figure 6



Supplementary Code: 1

Supplementary Figure Legends

Supplementary Table 1. Calibration dataset theoretical and ThicknessTool measurements.

Supplementary Table 2. Training dataset correlation analysis of outer nuclear layer thickness measurements.

Supplementary Table 3. Training dataset correlation analysis of inner nuclear layer thickness measurements.

Supplementary Table 4. Training dataset thickness measurement coefficient of variation between observers.

Supplementary Figure Legends

Supplementary Figure 1. Capture of the graphical user interface dialog for the ThicknessTool.

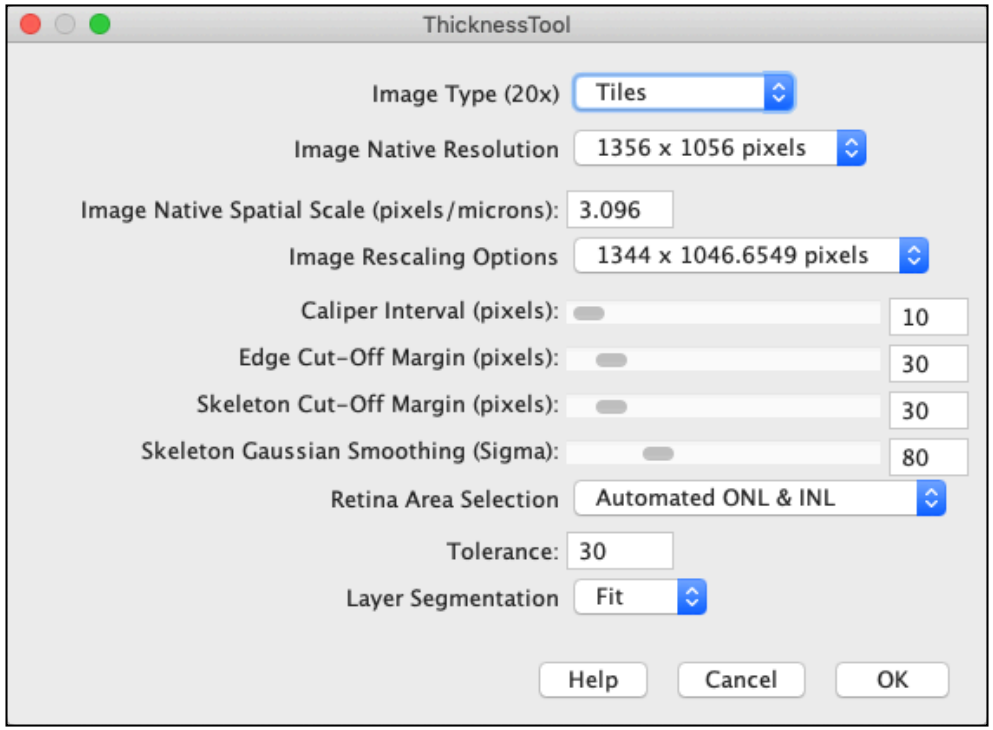
Supplementary Figure 2. Outline of ThicknessTool (TT) Calibration. (A) Representative calibration image of a known thickness (200 pixels). (B) Representative calibration image after processing with thickness callipers (200 pixels). (C) Cropped area in a 45º tilted image showing even edges. (D) Cropped area in a 30º tilted image showing jagged edges. (E) Known vs. measured thickness in images of increasing mask size. (F) Known vs. measured thickness in a mock image with abrupt thinning and thickening.

Supplementary Figure 3. Qualitative assessment of manual and automated thickness measurements of outer (ONL) and inner nuclear (INL) retina layers. (A) Representative native image of mouse retinal section. (B) Inexperienced observer's measurements, displaying a lack of reproducibility and oblique calliper vector (arrowheads). (C) Experienced observer's measurements, displaying a lack of reproducibility and overshooting calliper vector (arrowhead). (D) ThicknessTool measurement of the INL and ONL at 1-pixel interval.

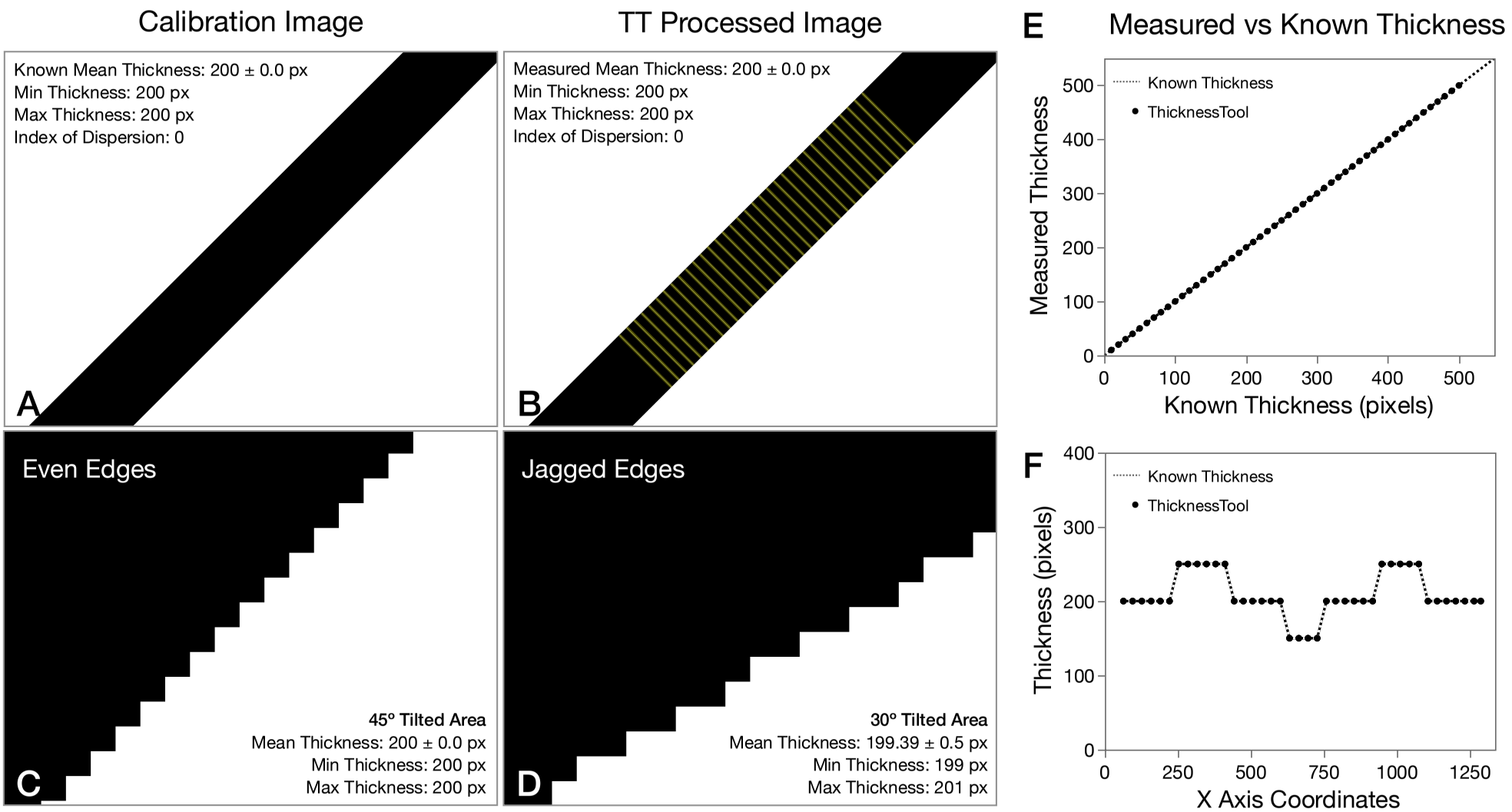
Supplementary Code

Supplementary Code 1. Source code of ThicknessTool.

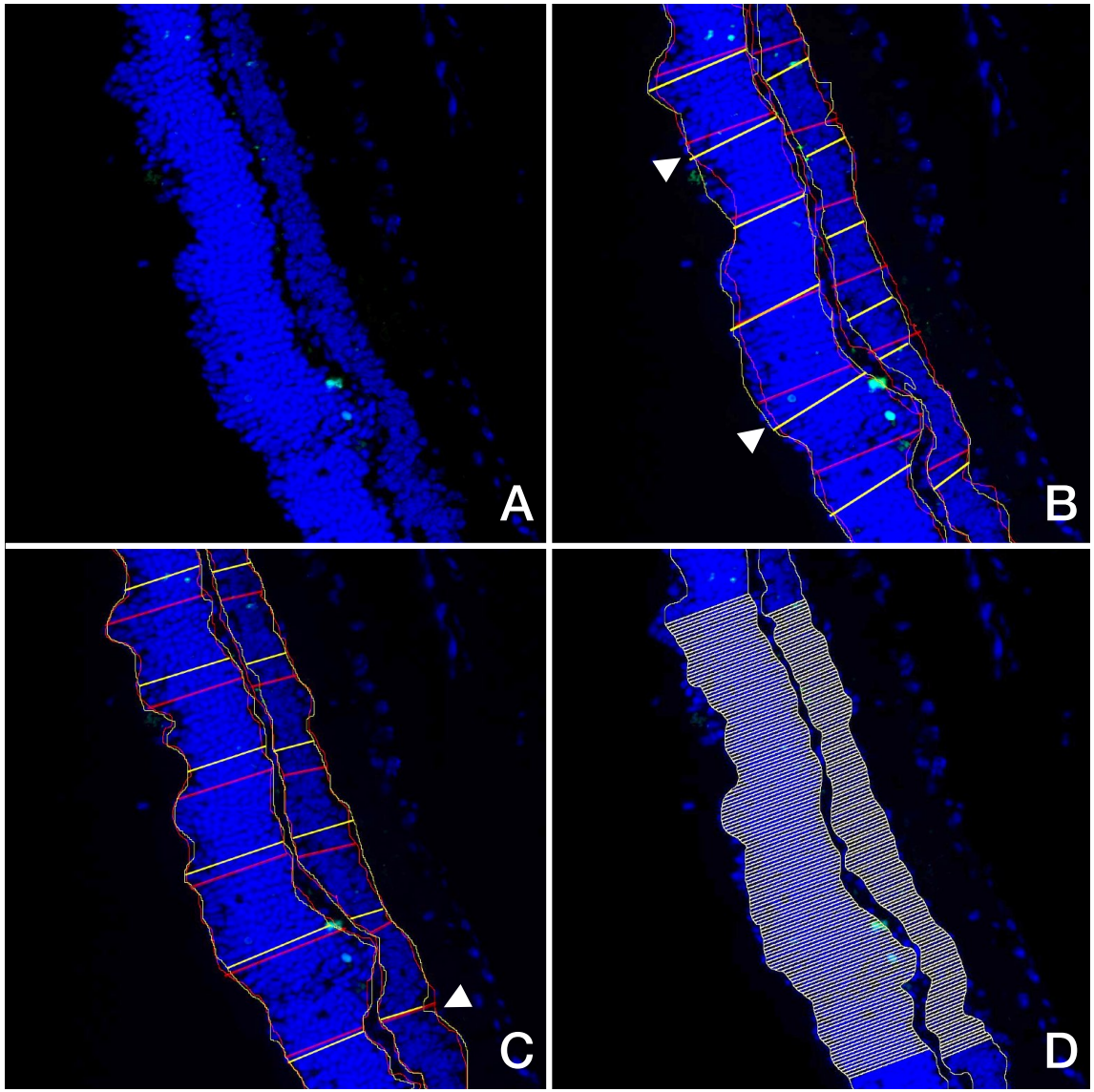
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Table 1. Calibration dataset theoretical and ThicknessTool measurements.

	Theoretical	**Overall (n = 21)**	**Straight/Even Edges (n = 5)**	**Jagged Edges (n = 16)**
Hypothesised Mean	-	200.00 ± 1	200.00 ± 1	200.00 ± 2
Mean Thickness	200.00 ± 0.00	199.88 ± 0.25	200.00 ± 0.00	199.85 ± 0.28
Minimum Thickness	200.00 ± 0.00	199.04 ± 0.74	200.00 ± 0.00	198.75 ± 0.74
Maximum Thickness	200.00 ± 0.00	200.57 ± 0.59	200.00 ± 0.00	200.75 ± 0.57
CoV Thickness	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
CoV Thickness (%)	0.00 ± 0.00	0.10 ± 0.09	0.00 ± 0.00	0.13 ± 0.07

CoV = coefficient of variation
Values shown are means ± standard deviations

Supplementary Table 2. Training dataset correlation analyses of outer nuclear layer thickness measurements.

	Inexperienced 1st	**Inexperienced 2nd**	**Inexperienced Mean**	**Experienced 1st**	**Experienced 2nd**	**Experienced Mean**	**Observers Mean**	**ThicknessTool**
Inexperienced 1st	-	0.96	0.99	0.90	0.91	0.92	0.97	0.88
Inexperienced 2nd	0.96	-	0.99	0.92	0.93	0.94	0.98	0.90
Inexperienced Mean	0.99	0.99	-	0.92	0.93	0.94	0.99	0.90
Experienced 1st	0.90	0.92	0.92	-	0.93	0.98	0.96	0.92
Experienced 2nd	0.91	0.93	0.93	0.93	-	0.98	0.97	0.90
Experienced Mean	0.92	0.94	0.94	0.98	0.98	-	0.98	0.93
Observers Mean	0.97	0.98	0.99	0.96	0.97	0.98	-	0.93
ThicknessTool	0.88	0.90	0.90	0.92	0.90	0.93	0.93	-

1st = first measurement; 2nd = second measurement
Values shown are Pearson correlation (r) coefficients (all coefficients $p < 0.001$)

Supplementary Table 3. Training dataset correlation analyses of inner nuclear layer thickness measurements.

	Inexperienced 1st	**Inexperienced 2nd**	**Inexperienced Mean**	**Experienced 1st**	**Experienced 2nd**	**Experienced Mean**	**Observers Mean**	**ThicknessTool**
Inexperienced 1st	-	0.94	0.98	0.86	0.80	0.88	0.97	0.86
Inexperienced 2nd	0.94	-	0.98	0.84	0.78	0.86	0.96	0.84
Inexperienced Mean	0.98	0.98	-	0.86	0.80	0.89	0.98	0.86
Experienced 1st	0.86	0.84	0.86	-	0.77	0.95	0.93	0.87
Experienced 2nd	0.80	0.78	0.80	0.77	-	0.94	0.89	0.87
Experienced Mean	0.88	0.86	0.89	0.95	0.94	-	0.97	0.92
Observers Mean	0.97	0.96	0.98	0.93	0.89	0.97	-	0.92
ThicknessTool	0.86	0.84	0.86	0.87	0.87	0.92	0.92	-

1st = first measurement; 2nd = second measurement
Values shown are Pearson correlation (r) coefficients (all coefficients $p < 0.001$)

Supplementary Table 4. Training dataset thickness measurement coefficient of variation between observers.

	Outer Nuclear Layer		**Inner Nuclear Layer**	
	Mean ± SD	*P* value*	Mean ± SD	*P* value*
Inexperienced	2.77 ± 2.52	**< 0.001**	4.82 ± 3.45	**< 0.001**
Experienced	3.81 ± 2.83	**0.009**	7.80 ± 6.42	0.602
Observers Average	5.39 ± 3.55	-	8.70 ± 5.55	-

SD = standard deviation; * = comparison for pair vs observers average using Tukey-Kramer HSD test
Values shown are means ± standard deviations