Statistical and Computational Analysis of Retinal Images to Understand Ocular Disorders

Thesis

Submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

MD. HASNAT ALI

ID No. **2013PHXF0505H**

Under the supervision of

Prof. M.B. Srinivas

&

Under the Co-supervision of

Prof. Saumyadipta Pyne &
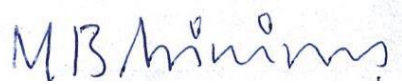
Dr. Sirisha Senthil

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI

2024

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI

CERTIFICATE

This is to certify that the thesis entitled, "**Statistical and Computational Analysis of Retinal Images to Understand Ocular Disorders**" submitted by **MD. HASNAT ALI** ID No. **2013PHXF0505H** in partial fulfillment of the requirements of Ph.D. of the Institute embodies original work done by him under our supervision.

Signature of the Supervisor
Name of the Supervisor – *Prof. MB Srinivas*
Designation - Professor, EEE Department, BITS Pilani, Hyderabad Campus.

Signature of the Co-Supervisor
Name of the Supervisor – *Prof. Saumyadipta Pyne*
Designation - Adjunct (Full) Professor, Department of Statistics and Applied Probability, University of California Santa Barbara, CA 93106.

Signature of the Co-Supervisor
Name of the Supervisor – *Dr. Sirisha Senthil*
Designation - Head, VST Centre for Glaucoma Care, LV Prasad Eye Institute, KAR Campus, Hyderabad.

Acknowledgments

Abstract

Progressive optic neuropathies such as glaucoma can cause irreversible blindness, especially when left untreated or diagnosed late. Indeed, early detection and management hold the key to slowing the progressive loss of vision and preventing blindness due to many chronic and age-related degenerative eye diseases. Glaucoma, for instance, is the second-leading cause of blindness worldwide.

In glaucoma, the structural damage is often characterized by neuroretinal rim (NRR) thinning of the optic nerve head, and other clinical parameters. Baseline structural heterogeneity in the eyes can play a key role in the progression of optic neuropathies, and present challenges to clinical decision-making. We generated a dataset of Optical Coherence Tomography (OCT) based high-resolution circular measurements on NRR phenotypes, along with other clinical covariates, of 3,973 healthy eyes as part of an established clinical cohort of Asian Indian participants. We introduced CIFU, a new computational pipeline for CIrcular FUnctional data modeling and analysis. We demonstrated CIFU by unsupervised circular functional clustering of the OCT NRR data, followed by meta-clustering to characterize the clusters using clinical covariates, and presented a circular visualization of the results. Upon stratification by age, we identified a healthy NRR phenotype cluster in the age group 40-49 years with predictive potential for glaucoma. Our dataset also addresses the disparity of representation of this particular population in normative OCT databases.

Using the OCT platform, we can generate high-resolution data on RNFLthickness in a given eye at a large number of angular-points defined on a circular coordinate system. However, insightful characteristics of OCT RNFL data such as its circularity and granularity generally remain unutilized in its common downstream statistical analyses. We present a new circular statistical framework that defines an Angular Decay function and thereby provides a directional representation of an eye in terms of its focused RNFL loss. By applying to a clinical cohort of Asian Indian eyes, the generated circular data was modeled with a finite mixture of von Mises distributions, which led to unsupervised identification in different age-groups of 2 recurrent clusters of glaucomatous eyes with distinct directional signatures of RNFL decay. New indices of global and local RNFL loss were computed for comparing the structural differences between these glaucoma clusters across the age-groups.

Table of Contents

Contents

List of Tables

List of Figures

xiii

List of Abbreviations and Symbols

FDA	:	Functional data analysis
OCT	:	Optical coherence tomography
OCTA	:	Optical coherence tomography- Angiography
ONH	:	Optic nerve head
IOP	:	Intraocular pressure
POAG	:	Primary open angle glaucoma
PACG	:	Primary angle closure glaucoma
CCT	:	Central corneal thickness
SE	:	Spherical equivalent
RNFL	:	Retinal nerve fibre layer
CDR	:	Cup to disc ratio
SDOCT	:	Spectral-domain optical coherence tomography
mm	:	Millimeter
3-D	:	Three dimensional
μm	:	Micrometers
TDOCT	:	Time-domain optical coherence tomography
LVPEI-GLEAM	:	L V Prasad Eye Institute-Glaucoma Epidemiology And Molecular-genetics
NRR	:	Neuroretinal rim
CIFU	:	CIrcular FUnctional data modeling and analysis
HVF	:	Humphrey visual fields
eCDF	:	empirical Cumulative Distribution Function
IPL	:	Inner plexiform layer
GCL	:	Ganglion cell layer
RGC	:	Retinal ganglion cells
VF	:	Visual field
BMO	:	Bruch's membrane opening
TSNIT	:	Temporal-Superior-Nasal-Inferior-Temporal
MRW	:	Minimum rim width
LOGES	:	Longitudinal Glaucoma Evaluation Study
OD	:	Right eyes
OS	:	Left eyes
SAP	:	Standard automated perimetry
PAM	:	Partitioning Around Medoids
AIC	:	Akaike Information Criterion
BIC	:	Bayesian Information Criterion

ICL	:	Integrated Complete Likelihood
ASW	:	Average Silhouette Width
DFM	:	Discriminative functional mixture
SD	:	Standard deviation

Chapter 1

Introduction to glaucoma and the analysis of the optic nerve head, along with the assessment of the retinal nerve fiber layer and employment of Optical Coherence Tomography (OCT) for evaluation.

1.1   Introduction

Glaucoma comprises a collection of disorders featuring advancing optic neuropathy, marked by the deterioration of retinal ganglion cells and retinal nerve fiber layers. This deterioration leads to modifications in the optic nerve head (ONH) alongside corresponding visual field (VF) impairments [1] . Elevated intraocular pressure (IOP) stands as a significant risk element for the development of glaucoma, leading to harm to the optic nerve and subsequent depletion of retinal ganglion cells [2]. Untreated or late-diagnosed Glaucoma can lead to permanent vision loss, making early detection crucial for slowing its advancement and averting blindness arising from numerous age-related degenerative eye disorders. With approximately 80 million individuals affected in 2020, Glaucoma stands as the world's second most prevalent contributor to blindness, a number projected to surpass 111 million by 2040 [3].

1.2   Categorization of Glaucoma

Glaucomas can be categorized into two primary divisions based on their underlying anatomical and pathophysiological characteristics: open-angle glaucoma and angle-closure glaucoma. These divisions further encompass two major subtypes known as primary and secondary glaucoma. The idiopathic or primary forms of open-angle and closed-angle glaucoma manifest without discernible causative factors. Conversely, secondary glaucoma arises due to identifiable causes or a convergence of ocular and systemic factors, resulting in elevated IOP

that ultimately culminates in irreversible damage to the optic nerve.

Open Angel Glaucoma

Primary open-angle glaucoma (POAG)
Increased IOP with progression of optic nerve

Normal-tension glaucoma (NTG)
Normal IOP with progression, optic neuropathy

Secondary open-angle glaucoma
The cause of elevated IOP and/or optic neuropathy is
defined (i.e, pseudoexfoliation, pigment dispersion
syndrome, steroid use)

Angele Closure glaucoma

Primary angle closure glaucoma (PAC)

Acute Closure of anterior chamber
angel with sudden increase in IOP

Chonic Closure of anterior
chamber angle with gradual
increase of IOP or development of
peripheral anterior synechiae.

Secondary closed angle glaucoma
Closure of anterior chamber angel with
increased IOP due to identifiable causes

1.3   Figure 1-1 Categorization of Glaucoma

1.3.1 Open-angle glaucoma

Primary open-angle glaucoma (POAG) is characterized by the presence of glaucomatous damage to the optic disc and a corresponding visual field defect. This condition is identified when an open angle is observed during gonioscopy. The most important problem in these eyes is increased IOP due to aqueous humour outflow system dysfunction.

Normal-tension glaucoma (NTG) represents a variant of POAG, distinguished by the presence of glaucomatous optic nerve damage among individuals whose IOP readings remain consistently below 21mmHg [4, 5]. Compared to individuals with POAG, NTG patients often have a higher risk of optic nerve damage at relatively low IOPs. Based on the optic nerve's appearance alone, one would not differenntiiate POAG and NTG. In the realm of ophthalmoscopy, individuals afflicted with NTG exhibit an elevated inclination towards optic disc hemorrhaging and the presence of localized irregularities within the retinal nerve fiber layer, particularly in proximity to the papillomacular bundle.

Secondary open-angle glaucoma pertains to various types of glaucoma wherein there exists a recognizable reason for increased IOP, leading to harm to the optic nerve along with associated VF impairments and the subsequent loss of vision. Unlike, primary open angle glaucoma, secondary open angle glaucoma are diagnosed when glaucoma occurs as a consequence of another systemic problem.

1.3.2 Angele Closure glaucoma

Primary angle closure develops when the drainage angle is closed by a physical obstruction, usually the peripheral iris. This closure prevents access of aqueous humour to the trabecular meshwork (TM) and hence the outflow obstruction [6]. This condition can be associated with very high IOP and can be sudden (acute) or chronic. Due to the severity of the condition, angle closure glaucoma, despite being a less frequent kind of glaucoma compared to POAG, contributes to 50% of all glaucoma-related blindness globally [7].

PACG Primary angle closure is classified as [8] i) Primary angle closure suspects (PACS) characterized by an eye displaying irido-trabecular contact (ITC) spanning ≥180° in an appositional manner, along with normal IOP and devoid of optic disc damage; ii) Primary angle closure (PAC) presenting an eye with ≥180° ITC coupled with features like goniosynechiae, blotchy pigments, peripheral anterior synechiae, elevated IOP, iris whorling, sphincter atrophy, anterior subcapsular lens opacities, and "glaucomfleken," yet without signs of optic disc damage; and iii) Primary angle closure glaucoma (PACG) characterized by the presence of disc damage alongside the clinical characteristics of PAC.

Secondary angle closure glaucoma is defined as glaucomatous optic disc injury with an occludable angle secondary to an evident aetiology.

Clinical Risk Factors for glaucoma

Clinically, certain factors help determine an individual patient's glaucoma risk based on that patient's unique characteristics. Understanding the relative significance and intensity of clinical risk factors is helpful for using this knowledge best. Most information on glaucoma risk factors comes from surveys or case-control research. Clinical risk factors can be categorized into general, ocular, and systemic.

1.4 General risk factors

Age

The correlation between advancing age and an elevated susceptibility to developing the condition has been consistently evidenced [9, 10]. Age stands as the chief demographic trait linked to an elevated susceptibility to glaucoma. The likelihood of contracting glaucoma escalates as one grows older, surging by a significant multiplier between the age range of 40-50 years and that of surpassing 80 years [10, 11]. Except for lineage, which will be discussed in the subsequent section, no other demographic trait has been consistently linked to this heightened risk.

The process of aging is linked to various alterations in both the eye and its corresponding blood circulation, potentially exacerbating the development of glaucoma. Advancing age is correlated with a diverse range of modifications in the eye and its related vascular system. The primary route through which the aqueous humor leaves the eye is the trabecular meshwork [12]. The movement of the aqueous humor, generated within the ciliary body, occurs from the back to the front of the eye, subsequently departing through the trabecular meshwork and the uveoscleral pathways. The aging process is linked to decreased drainage through these pathways for the aqueous humor [13, 14]. As age advances, there is an observed deterioration of trabecular meshwork cells and a buildup of extracellular materials within the meshwork structure [15]. Numerous studies involving population-based samples [16, 17] have consistently indicated that there is an increase in IOP associated with advancing age. It is possible that these alterations are to blame for this trend.

The diverse functions of glial cells within the retina and optic nerve play a crucial role in enhancing the performance of retinal ganglion cells through various mechanisms [18]. These functions encompass the supply of both structural and functional assistance to retinal ganglion cells, which includes regulating the extracellular milieu, synthesizing neurotrophins, and facilitating fundamental metabolic activities. Microglia provide immunological regulatory tasks, assist in the maintenance of perivascular barriers, and make an effort to protect neurons from inflammatory cytokines that are damaging. The ability of microglial cells found in the central nervous system to remain viable and undergo regeneration is negatively impacted by the process of aging [19]. Age-related attrition patterns comparable to the aforementioned

phenomenon can also be observed in retinal and optic nerve glial cells. There is a conjecture that these modifications might exert an adverse impact on the essential neuro-supportive and neuroprotective functions performed by these cells, as indicated by a previous study [20]. Furthermore, a heightened occurrence of apoptosis in conjunction with diminished regenerative capacities becomes evident in the endothelial cells of blood vessels with advancing age. This is due to the fact that our bodies produce less oxygen as we get older. These alterations clearly have deleterious repercussions for the process of autoregulation. Moreover, a decrease in ONH blood flow is seen with advancing age [21].

Ancestry (Population /Ethnicity /Race)

Ancestry stands as a consistent risk element associated with glaucoma. In their study, Rudnicka et al. (2006) executed a comprehensive analysis of variations in the occurrence of POAG concerning age, gender, and ethnic backgrounds. The outcomes revealed a notable pattern: individuals of African-Caribbean ("black") descent exhibit elevated prevalence rates of POAG across all age brackets, surpassing the rates observed in White-Caucasian and Asian populations. In people over the age of 70, the estimated general prevalence is 16% among African Americans, 6% among Caucasians, and 3% among Asians [10]. Individuals of African descent face an elevated risk of POAG and experience a more rapid progression of the disease, culminating in eventual blindness. The underlying reasons for this phenomenon have been ascribed to variances in optic nerve composition, central corneal thickness (CCT), IOP, and the presence of other potential glaucoma-linked risk factors like refractive errors, blood pressure, and diabetes. However, these findings have not exhibited uniformity across various research endeavors [22]. Nevertheless, it is essential to interpret these findings cautiously due to the oversimplification inherent in racial categorizations such as "Black," "White," or "Asian," which neglects the intricate diversity within humanity's populations [59]. Notably, individuals of African heritage constitute one of the most genetically heterogeneous groups globally. Even within "homogeneous racial groups," substantial variations in POAG rates can exist [23]. The prevalence of POAG varies widely, ranging from 2.9% in South Africa and rural Nigeria to 8.3% in Ghana. The prevalence of POAG is notably lower among individuals of African ancestry in South Africa, Nigeria, Tanzania, and Baltimore in comparison to Ghana, St. Lucia,

or Barbados. Furthermore, when contrasted with white Dutch individuals, white Australians exhibit a significantly higher incidence of POAG [23].

Gender

The correlation between gender and glaucoma has yielded inconclusive findings in existing research. Various prevalence studies have yielded differing results: some indicating a higher occurrence in men, while others suggesting an increased frequency in women, and some revealing negligible disparity between genders. For example, research from the Baltimore and Beaver Dam studies found no significant gender-based variance in glaucoma risk [16, 24]. Conversely, the Rotterdam and Barbados Eye studies identified an elevated risk among males[25, 26], whereas the Blue Mountains Eye Study noted a higher prevalence of POAG among women [27]. These inconsistencies are similarly mirrored in numerous incidence studies. For instance, investigations in Rotterdam and Skelleftea concluded that gender had no discernible impact on glaucoma frequency [28, 29], whereas the Dalby study indicated a higher incidence among females [30].

As a result, there appears to be some evidence supporting an augmented risk of POAG in males within specific populations, while in others, females face a higher risk. The underlying causes of these inconsistencies remain unclear. One potential explanation is that these variations are reflective of gender's role in the pathophysiology of glaucoma across diverse populations. In a meta-analysis by Rudnicka et al. (2006) involving 46 studies, it was found that men have a significantly greater likelihood of developing POAG than women[10]. However, certain subgroups of POAG, such as normal tension glaucoma, consistently exhibit higher prevalence among women [31]. Some indications suggest that female sex hormones could potentially have a protective effect against elevated IOP [32, 33]. This could be attributed to estrogen receptors present in the ciliary body and outflow tract, potentially influencing glaucoma through aqueous formation and vascular factors[34, 35]. Nevertheless, population-based studies present conflicting evidence [34, 36, 37].

1.5   Systemic risk factors associated with glaucoma

Systemic conditions like diabetes[38-40], hypertension[41, 42], heart-related ailments[43], thyroid irregularities [44], and migraines[45] have all displayed associations with glaucoma. Regrettably, these correlations haven't always remained uniform across comprehensive

population-based cross-sectional investigations and long-term trials [46]. Certain studies indicate that the influence of these factors might also be subject to the modulation of additional risk determinants. These, along with a plethora of other variables spanning autoimmune, neurodegenerative, endocrine, and vascular factors, have undergone exhaustive evaluation as delineated by Pache and Flammer in 2006 [47], and more recently by Schmidl et al. in 2011 [48].

1.6   Ocular risk factors

Intraocular pressure

A substantial body of evidence establishes a strong correlation between intraocular pressure (IOP) and the occurrence of glaucoma. Numerous cross-sectional studies conducted on populations globally[49] such as the Beaver Dam, Blue Mountains, Tanjong Pagar, Tajimi, and Chennai glaucoma research initiatives, have consistently demonstrated a direct relationship between higher IOP levels and an increased prevalence of glaucoma. Moreover, the severity of this connection between elevated IOP and glaucoma is exacerbated by escalating concentrations of IOP[17]. Clinical interventions including medication, surgical procedures, and argon laser trabeculoplasty, aimed at reducing IOP, have all exhibited their efficacy in curtailing the progression of glaucomatous optic neuropathy and the deterioration of visual field, as evidenced by randomized clinical studies.[50, 51]. It is widely acknowledged that elevated IOP sets in motion a series of sequential events culminating in the apoptotic demise of retinal ganglion cells [52].

Central corneal thickness

While cross-sectional studies have indicated a potential association between CCT and conditions such as POAG, normal-tension glaucoma, and ocular hypertension[53] the initial prospective investigation that demonstrated a causal relationship between thinner central cornea and the progression from ocular hypertension to POAG was the Ocular Hypertension Treatment Study (OHTS). Individuals with a corneal thickness of 555 micrometers or less were found to be at a threefold higher risk of developing POAG compared to those with a corneal thickness exceeding 588 micrometers. Subsequent research has corroborated these

findings[54]. It is widely recognized that a comprehensive assessment of glaucoma risk for potential patients should encompass consideration of CCT[55].

However, the substantiating evidence for CCT as a definitive risk factor for glaucoma is somewhat equivocal. Several investigations, including population-based studies conducted in Barbados and Rotterdam, have established a connection between reduced CCT and the development of glaucoma[56]. Additionally, a thinner central cornea has been linked to both glaucomatous visual field loss and damage to the optic nerve head[57]. Nonetheless, not all studies [58] have managed to replicate these outcomes. As it stands, the confidence in thinner central cornea serving as a decisive risk factor for glaucoma development remains limited [59].

Myopia

Associations between glaucoma and myopia have been observed in cross-sectional and case-control studies conducted on multiple continents [60]. Similar to glaucoma research, comparing studies on refractive error is challenging because there is no universal agreement on how much myopia or hyperopia constitutes. However, not every study has confirmed the link between glaucoma and nearsightedness. Quigley et al. (1994) followed 647 people with ocular hypertension and found that myopia was not associated with an increased risk of glaucomatous visual field loss in this population. Randomized clinical trials have also yielded contradictory findings [61].

Myopia and glaucoma may be linked, however the underlying molecular mechanism is unknown. The increased risk appears to serve as an autonomous risk factor for the development of glaucoma (at least in myopia 4D) rather than a result of an increased correlation with elevated IOP [62]. Myopic eyes differ structurally from emmetropic eyes, which may explain the link between the two [63]. The anterior chambers and axial lengths of myopic eyes are enlarged. In myopia, the lamina cribrosa is less thick and more malleable than in emmetropia [63]. Due to their structural makeup, myopic eyes are higher likelihood to sustain damage to the ONH from variations in intraocular pressure (IOP).

1.7　The Role of Family History and Genetic Factors in Glaucoma Susceptibility

The significant role of hereditary or genetic inheritance in the development of glaucoma is now widely recognized. While it's important to note that a substantial portion of cases might go unreported, cross-sectional epidemiological investigations reveal that the prevalence of a family history of glaucoma among POAG patients ranges from 10% to 50% [64]. This risk is notably higher, around 2-10 times, for first-degree relatives [65]. Moreover, the incidence of glaucoma displays variation among different self-reported ethnic groups, as evidenced by several comprehensive population-based studies. A recent meta-analysis [10] highlights that the estimated occurrence of POAG among individuals aged 70 and above varies from 3% in Asian populations to 6% in white Caucasian communities, and up to 16% in black African and Caribbean populations. Notably, black individuals are prone to developing POAG at an earlier age and a faster progression rate compared to white Caucasians, which also translates to a heightened risk of POAG-related irreversible blindness [66]. Research conducted on population samples suggests that PACG is more prevalent among East Asian and Eskimo populations, whereas white Caucasians and individuals with Afro-Caribbean heritage show a higher prevalence of POAG [57].

Additional support for the genetic basis of glaucoma emerges from twin studies [67]. A prospective study conducted by Gottfredsdottir et al. (1999) on 50 pairs of monozygotic twins and their spouses reveals higher concordance rates of open-angle glaucoma within twin pairs (98%) compared to spouse-twin couples (72%). This observation emphasizes the genetic contribution to the condition. Subsequent to this, numerous genes have been identified as potential factors contributing to POAG, along with the discovery of multiple genetic loci associated with the condition.

1.8   Detecting glaucoma in a clinical environment

The three fundamental pillars of glaucoma detection, namely tonometry, visual field testing, and ophthalmoscopy, have been applied for screening purposes[68]. Intriguingly, there exists a surprising gap in the comprehensive exploration of the collective effectiveness of these three tests in accurately identifying glaucoma. While Harper & Reeves (1999) conducted an assessment of the sensitivity and specificity of diverse glaucoma screening measures, positioning visual field screening, optic disc cupping, and intraocular pressure (IOP) as the most dependable indicators of glaucoma in that sequence, the holistic synergy of these tests remains insufficiently probed [69]. As an added bonus, the combined sensitivity of the tests

was shown to be much higher than that of either test done separately. A visual field screening test was also required to give sensitivity and specificity more than 0.90 [70] . Visual field testing takes longer, and clear guidelines for automated mass screening of the visual field are yet to be established [71]. When it comes to sensitivity, it's possible that analysing the ONH is doable, but doing so requires specialised personnel and/or costly equipment. The need for trained persons to document the ONH and retinal nerve fibre layer (RNFL) thickness during mass screening [72] has been eliminated thanks to photography that captures ONH images and examines the RNFL thickness using various equipment like OCT [73].

1.8.1 Tonometry

In the realm of glaucoma diagnosis and management, monitoring and measuring intraocular pressure (IOP) stands as the predominant clinical tool. While an elevated IOP remains linked to a heightened glaucoma risk [74], it is imperative to acknowledge that no level of IOP exists that can be deemed entirely "safe," devoid of risk. The influence of age on IOP remains unsubstantiated, and the established average IOP for the general population hovers around 16 mmHg [75]. Amidst the process of IOP recording, various potential sources of error emerge, some tethered to the tonometer type in use, while others encompass broader factors that affect measurements universally, including inherent patient biological traits.

Even the widely acclaimed Goldmann tonometry, heralded as the "Gold Standard," is not without its notable deficiencies. An in-depth exploration into inaccuracies inherent to the Goldmann tonometer exposed six primary categories, each harboring a cascade of potential error triggers [76]. The unveiling of the OHTS [74] ushered in a spotlight on corneal thickness as an influential predictor in POAG onset. This revelation spurred extensive discourse on corneal thickness as a conceivable error source. The OHTS illuminated the misclassification of individuals as "at risk" due to erroneous IOP readings, stemming from augmented corneal thickness, despite their actual levels likely falling within the normal range [77]. In essence, there are several vulnerabilities that can infiltrate IOP measurements, which practitioners might choose to overlook. Misattributions of IOP status extend to patients with glaucoma, those under suspicion of glaucoma, and even those classified as normal based solely on tonometry outcomes, lacking accurate procedures and IOP adjustments grounded in corneal thickness evaluations.

1.8.2 Visual field

The region of space that becomes perceptible during steady fixation with one or both eyes open is referred to as the visual field of the fixating eye. The visual field of one eye can be analogously likened to a "Hill of Vision" (HoV) encircled by a "sea of blindness" (SoB)[78]. The highest point of this HoV is centered on the fovea, where retinal sensitivity reaches its peak, measured in terms of elevation. Interestingly, this peak also corresponds to the area of highest retinal sensitivity. Remarkably, the blind spot on the retina corresponds to an unfathomable hollow within the HoV. Should there be an absence of photoreceptors in this blind spot, it would lead to an absolute scotoma or complete visual deficiency. Conversely, a relative scotoma results from localized depressions within the HoV. More generally, a reduction in the overall height of the HoV, without altering its shape, constitutes a generalized loss. The conventional extent of the visual field comprises approximately 60 degrees nasally, 60 degrees superiorly, 70 to 75 degrees inferiorly, and 100 to 110 degrees temporally. In cases of visual field abnormalities linked to glaucoma, the patterns typically trace the course of impaired retinal nerve fibers. These glaucomatous visual field deficits are recognizable through widespread depression, nasal step, paracentral scotoma, and arcuate abnormalities [78, 79].

1.9 Analyzing the Optic Nerve Head and evaluating the Retinal Nerve Fiber Layer

The primary glaucoma, a progressive multifactorial optic neuropathy, is marked by a hereditary, gradual decline in retinal ganglion cells and their associated axons [80]. Typically, injury to the RNFLleads to specific visual field irregularities and modifications in the structure of the ONH. A comprehensive comprehension of these transformations is imperative for diagnosing and monitoring glaucoma patients.

Distinguishing between glaucomatous, nonglaucomatous, and normal optic discs can be achieved through appropriate clinical examination techniques. Key features to observe encompass the size and shape of the optic disc, the ratio of cup-to-disc (C:D) concerning disc size, the depth of the optic cup, configuration of the neuroretinal rim, location of the central

retinal vessel trunk exit, presence and site of disc haemorrhage, defects in the RNFL, and the pattern and position of parapapillary chorioretinal atrophy.

Indications of ONH damage due to glaucoma encompass focal (notches) or diffuse narrowing of the neuroretinal rim along with concentric expansion of the optic cup, or a combination of these features[81]. Although glaucoma manifests diverse patterns of RNFL damage, such as wedge defects and diffuse loss [81, 82], they all share a commonality of reduced RNFL thickness [83].

Evaluation of the ONH can be conducted using fundus images, scanning laser devices, and direct or indirect ophthalmoscopy (e.g., scanning laser ophthalmoscopy, OCT, scanning laser polarimetry). Objective technological tools are progressively replacing traditional assessments of the RNFL and optic disc. However, it is prudent for clinicians to initiate with fundamental examinations that do not necessitate advanced equipment or methods that may not be universally accessible across eye clinics.

To appraise the baseline condition of the optic disc, color fundus images, or even better, stereo-photographs, are valuable, while monochromatic fundus photographs (specifically the blue channel profile in RGB) are optimal for assessing the RNFL. Repeated acquisition of these images during subsequent visits is a recommended practice to establish a longitudinal visual record for monitoring changes over time. Factors such as the state of the retinal pigment epithelium, image alignment, and optical clarity should also be taken into account, especially when considering therapeutic interventions.

1.9.1 Typical appearance of the Optic Disc

A standard indication of a healthy ONH structure is the presence of a well-proportioned, vibrant neuroretinal rim within an appropriately colored and sized cup, as depicted in Figure 1-2. Nevertheless, the task of identifying impairments in the optic disc is complex due to the

extensive range of natural disc morphologies in humans.

Figure 1-2 The optic nerve head in its the normal condition Maintaining its natural state, the optic nerve head

Notably, the Central India Eye and Medical Study (CIEMS) reported an average optic disc area of 2.25 ± 0.51 mm² for Indian eyes, while the Vellore Eye Study (VES), conducted among the South Indian population, observed a mean area of 2.58 mm². In contrast, the Andhra Pradesh study documented a higher mean optic disc area of 3.37 mm². These variations in measurements could potentially be attributed to divergent measurement methods employed.

1.9.2 Conventional Criteria for Assessing Optic Disc Morphology

Conducting thorough systematic examinations of the optic disc is imperative to avoid the oversight of crucial information that could differentiate between an individual with good ocular health and someone with early-stage glaucoma. Conventional assessment techniques involve employing five principles:

1. Discerning the scleral ring's location to demarcate the optic disc's boundaries and dimensions.
2. Appraising the constraints of the neuroretinal rim.
3. Examining the integrity of the peripapillary retinal nerve fiber layers (RNFLs) and the arrangement of RNFL bundles within the retinal structure.
4. Detecting indicators of peripapillary atrophy and its associated manifestations.
5. Assessment for optic disc hemorrhage
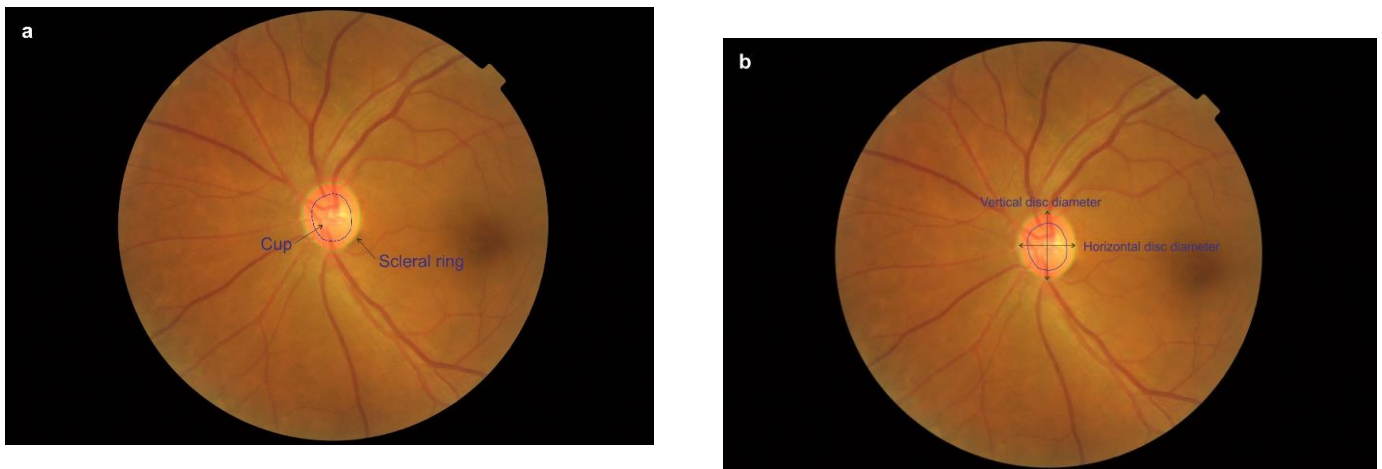
Figure 1-3. (a) Determining Scleral Ring and Optic Disc Boundaries in an Average Individual (b) Utilizing the length of the beam from the slit lamp, it is possible to compute the dimensions of both the vertical and horizontal optic discs.
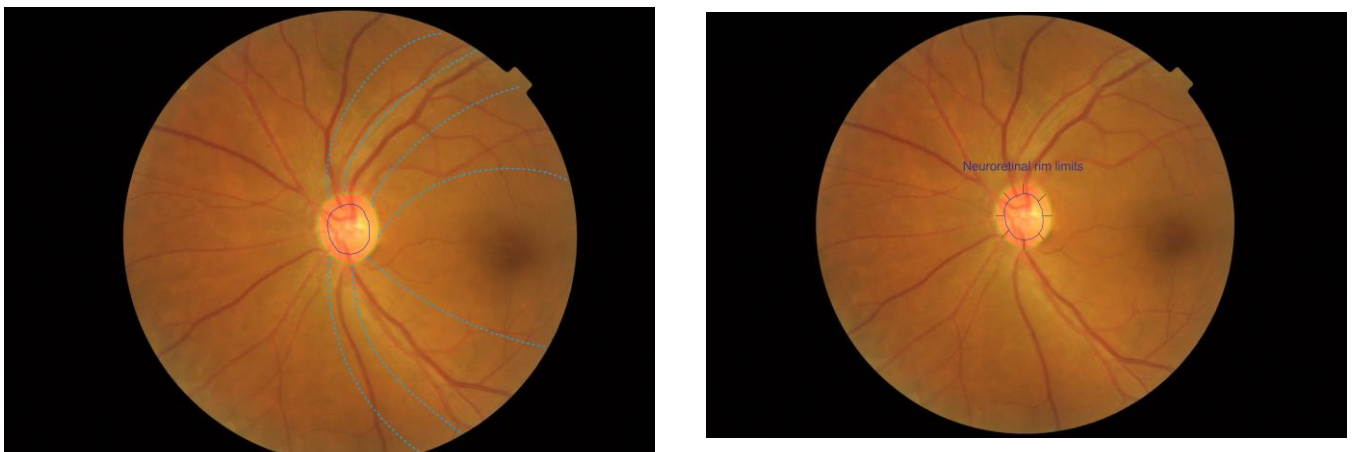
Figure 1-4(Left) The area enclosed by the scleral ring and the border of the cup may be used to identify the neuroretinal rim. (Right) High-quality colour fundus pictures show striations on the retina's surface as the retinal nerve fibre layer bundles.

1.9.2.1 Identify the Scleral Ring

The initial step in the analysis of ONH morphology involves determining the limits of the scleral ring and evaluating the optic disc's dimensions to classify it as small, medium, or large. The size of the optic disc holds notable importance in influencing the vertical cup-to-disc ratio[84, 85]. When dealing with small glaucomatous discs, observable cupping or neuroretinal rim atrophy might be absent. Conversely, sizeable optic discs might exhibit physiologically amplified cupping, often misinterpreted as glaucomatous. Thus, it is crucial to recognize that a large disc in healthy eyes could possess a substantial cup, while a diminutive disc in eyes affected by glaucoma might display a proportionally small cup. Precisely demarcating the ONH boundaries can pose challenges in instances of highly myopic eyes, discs characterized by significant chorioretinal atrophy, and atypical optic disc configurations. However, the curvature or orientation alterations of blood vessels might provide insights into the optic disc's contour under such conditions.

1.9.2.2 Neuroretinal Rim Limits Evaluation

The variation in the distribution of RNFLbundles across different regions of the optic disc leads to varying thicknesses of the neuroretinal rim. Specifically, the inferior pole of the optic disc exhibits a greater abundance of RNFL bundles, resulting in the thickest rim. Following this, the superior region possesses a moderately thick rim, succeeded by the nasal region, and finally the temporal region, which features the thinnest rim among these areas (inferior ≥ superior ≥ nasal ≥ temporal). This distinct arrangement of neuroretinal rim thickness is encapsulated by the ISNT rule, denoting the sequence of inferior, superior, nasal, and temporal.[86, 87]. Nevertheless, this ISNT pattern is often disregarded in cases of glaucoma, leading to localized loss or narrowing of the rim, recognized as notching or thinning. To comprehensively evaluate the condition of the optic disc, attention must be given to both the thickness of the neuroretinal rim and its coloration. While a healthy rim typically exhibits hues ranging from orange to pink, a rim appearing white or yellowish signifies the presence of neuro-ophthalmological conditions

other than glaucoma. Generally, a pale neuroretinal rim indicates the occurrence of
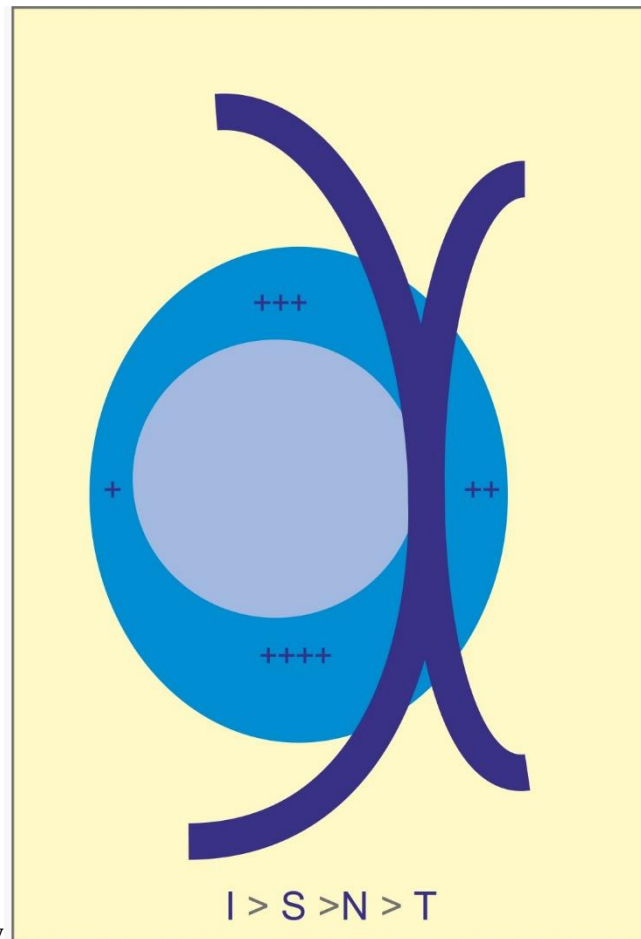
nonglaucomatous optic neuropathy.

Figure 1-5 In the context of ocular health, the neuroretinal rim's lower region typically exhibits greater thickness compared to the upper, nasal, and temporal regions (Inferior ≥ Superior ≥ Nasal ≥ Temporal). This pattern is in accordance with the guidelines outlined by the ISNT rule.

1.9.2.3 Investigation into the Distribution of Retinal Nerve Fiber Layer Bundles

Monochromatic fundus images, utilizing shorter wavelength illumination like blue light, facilitate enhanced visualization of RNFLbundles compared to color photos. Optimal image quality generally necessitates pupil dilation and clear ocular media to counteract the limited blue light penetration through the retina. The RNFL correlates with ganglion cell axons. Within healthy individuals, the RNFL displays distinct, bright striations with a curving pattern along

31

the   inner   retina,   as   illustrated   in   Figure   1-6.
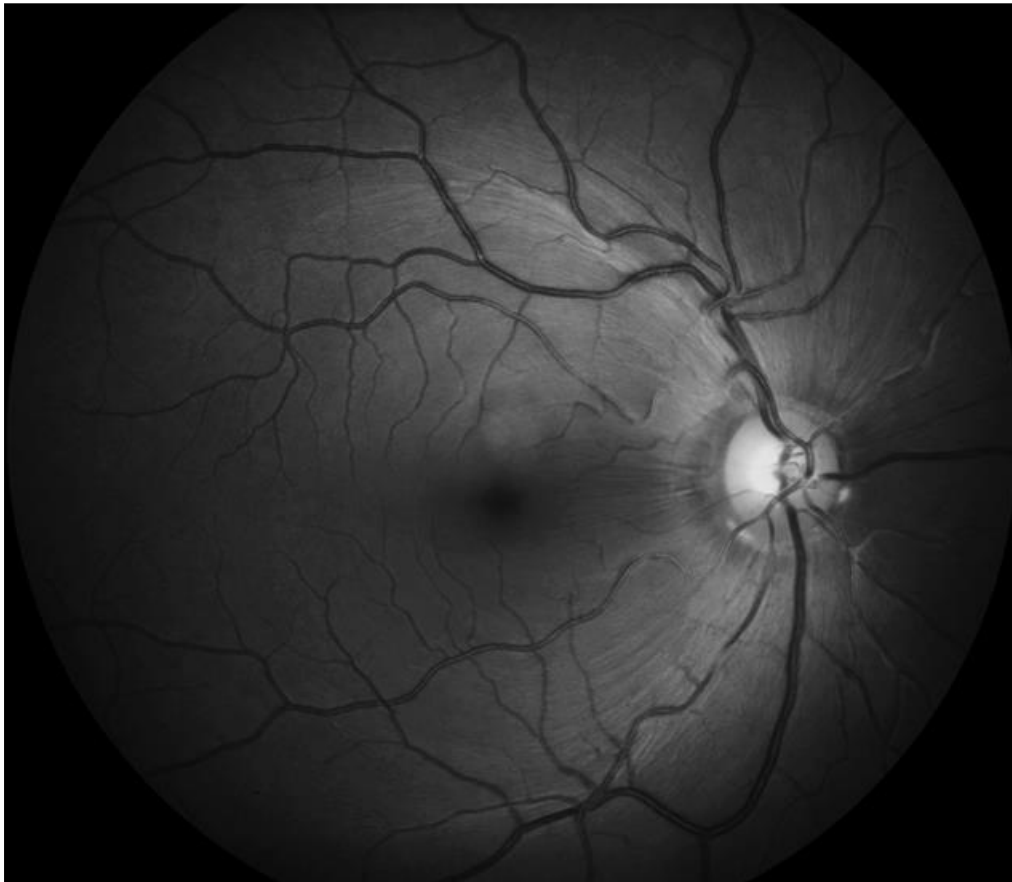
Figure 1-6 Monochrome fundus image displaying the vivid RNFL bundle striation pattern.

The regions of the fundus including the inferior temporal, superior temporal, and inferior nasal areas are the primary locations for observing the retinal nerve fiber layer (RNFL), as noted in references[88-91]. Notably, the optic disc vicinity offers a favorable site for RNFL identification. Conversely, peripheral retina exhibits the lowest RNFL visibility. The process of aging results in diminished RNFL visibility due to the inherent decline of ganglion cells (approximately 3000–5000 annually) from the initial count of 1.4 million cells at birth, as highlighted in references [90, 91].

Patients with glaucoma exhibit two distinct patterns of RNFL reduction [82].

1. The presence of localized RNFLabnormalities or wedge-shaped defects (illustrated in Figure 1-7) manifests as darker focal regions exhibiting reduced or diminished visibility of the characteristic striated pattern. These wedge defects may become apparent in individuals with mild to moderate glaucoma, while also serving as an early indicator of glaucomatous damage in individuals who exhibit normal scores on standard automated

perimetry tests. Notably, since age-related RNFL loss tends to be widespread, the occurrence of specific localized losses cannot be attributed to patient age.
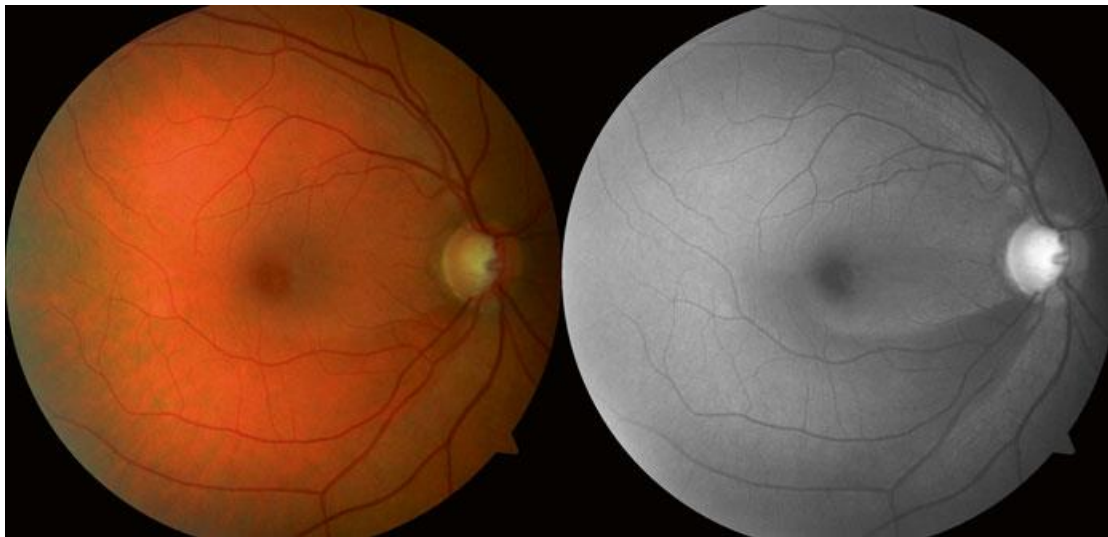
Figure 1-7 Wedge imperfections are discernible within color fundus images (depicted in the left image), yet they become more pronounced in monochrome fundus pictures (illustrated in the right image). The visual depicts a case of a patient afflicted by primary open-angle glaucoma, with the inferior wedge anomaly manifested as a distinct dark area.

2. Diffuse losses encompass a wide attenuation of the usual RNFL striation arrangement, which seems to merge into a uniform or slightly granulated light-gray region. In contrast to wedge-shaped defects, diffuse loss commonly signifies more progressed stages of glaucoma. The identification of this pattern frequently necessitates a thorough scrutiny of the photographs.

1.9.2.4 Peripapillary Atrophy Assessment

The correlation between glaucoma and peripapillary atrophy dates back to the early 20th century. Peripapillary atrophy is clinically categorized into two zones: the alpha zone (α-zone) and the beta zone (β-zone) [92]. The alpha zone is characterized by irregular hypo- and hyperpigmentation along the peripapillary atrophy's border. Situated between the peripapillary scleral ring and the alpha zone, the beta zone encompasses visible sclera and prominent choroidal veins [93]. Positioned closer to the ONH, the beta zone exhibits a substantial loss of photoreceptors and closure of the choriocapillaris, indicating retinal pigment epithelium deficiency or atrophy [94, 95].

Interestingly, the alpha zone is present not only in healthy eyes but also in those with glaucoma. Notably, the beta zone's presence is closely linked to glaucoma [92, 93]. Glaucoma patients tend to exhibit a more extensive and prevalent beta zone compared to

individuals without glaucoma [93]. Moreover, the location of glaucomatous optic disc degeneration and visual field abnormalities correlates with the site of the beta zone [92]. In individuals with glaucoma, the beta zone serves as a significant risk factor in the development of visual field deficits [96, 97]. However, it's worth noting that the size of the beta zone does not seem to correlate with the progression of glaucoma from ocular hypertension [98].

Figure 1-8 The green patch depicts the beta zone of the glaucomatous optic disc, while the peripheral alpha zone (indicated in purple) is unrelated to the development of glaucoma..

1.9.2.5 Examining the Eye for Optic Disc Hemorrhage

Thorough examination of the optic disc is imperative for the detection of disc hemorrhage, which can often be minuscule and evade detection. Hemorrhagic occurrences adjacent to the optic ONH border, characterized by splintered or flame-shaped appearances, typically signify the presence of glaucoma. Although optic disc bleeding is infrequent among individuals in good health, it doesn't invariably denote the presence of glaucomatous optic neuropathy. Hemorrhaging within the optic disc predominantly materializes within the inferotemporal or supero-temporal sectors. Comparatively, individuals afflicted with normal-tension glaucoma display a heightened frequency of disc hemorrhage in contrast to

those with high-tension glaucoma, with incidence ratios ranging from 1.3-2.8 compared to prevalence ratios ranging from 1.9-5[99-101]. Hemorrhagic manifestations exhibiting a flame-like configuration are indicative of advanced glaucoma[100].
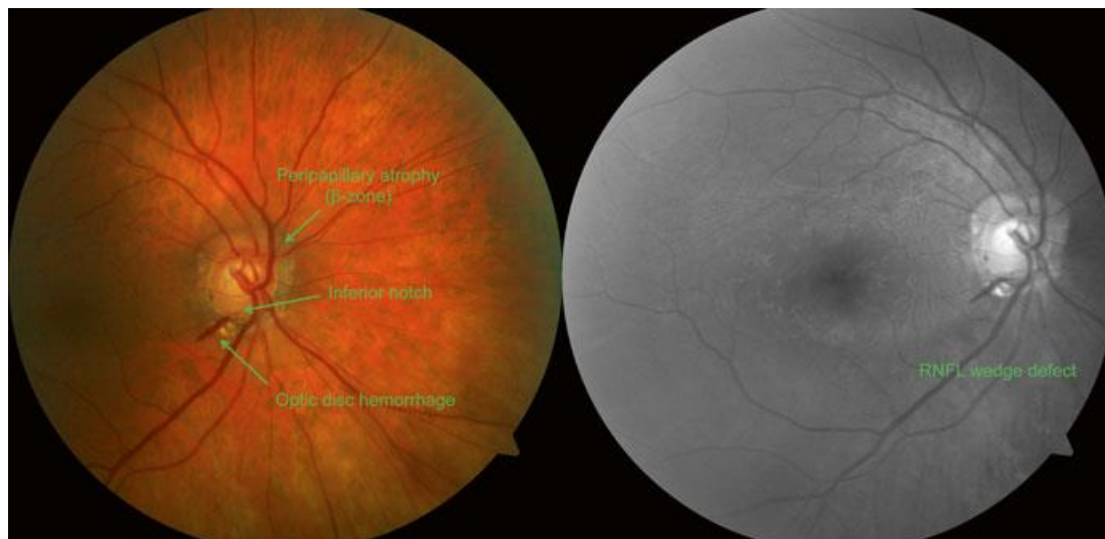
Figure 1-9 In color retinography (depicted in the left image), a flame-like hemorrhage might be observed close to the lower aspect of the optic disc. The monochromatic fundus image, on the other hand, reveals the hemorrhage situated in proximity to both a wedge-shaped RNFL defect and a notch present in the neuroretinal rim (as shown in the right picture).

1.10   Introduction, Evolution, and Current State OCT

1.10.1 Introduction:

In the field of ophthalmology, OCT has made significant advancements over the past two decades. One major area that has benefitted from OCT is the diagnosis and monitoring of macular disorders by retina doctors. The effectiveness and widespread use of modern anti-VEGF therapies have been greatly aided by OCT imaging. Additionally, OCT has had a major impact on the field of glaucoma. It has become the primary method for early diagnosis and monitoring of the disease, specifically in assessing the ONH, retinal nerve fiber layer (RNFL), and inner macular health. This technology has replaced previous methods such as confocal scanning laser ophthalmoscopy and scanning laser polarimetry due to its reproducibility and ability to provide various metrics and biomarkers. OCT has the capability to detect glaucoma at an early stage, even before visual field abnormalities develop. While OCT is widely used by ophthalmologists around the world, interpreting its data for glaucoma requires expertise in both operating the device and interpreting the images. Lack of familiarity with the technology can lead to diagnostic errors caused by artifacts and inter-individual variances.

OCT, an imaging technique founded on optical reflectometry, which captures high-resolution, in-vivo images of tissues that are semi-transparent or translucent. These images possess a resolution similar to that of a low-power microscope and can penetrate tissues at a depth of 2-4 mm [102, 103]. The sole divergence between ultrasonic imaging and its application lies in the utilization of laser light. While the depth of tissue penetration remains limited to a few millimeters, the substitution of light for ultrasound results in a significant enhancement in imaging resolution. OCT facilitates non-invasive, in vivo cross-sectional imaging by gauging both the intensity and the time delay of echoed light scattered from biological tissues possessing transparency or semi-transparency. Addressing the challenge of directly quantifying echo delay time from diverse tissues, interferometry emerges as the solution due to the technical impracticality [102].

Ophthalmology extensively utilizes OCT due to the proliferation of numerous commercial OCT devices available for diagnostic applications. Contemporary OCT systems enable the swift acquisition of high-resolution 3D images of the desired tissues. As a result of these inherent qualities, OCT has emerged as an indispensable instrument for diagnosing and monitoring glaucoma and retinal disorders.

1.10.2 What is the actual functioning process of OCT?

Detecting tissue attributes and their relative positions within a biological imaging system is crucial in creating the final image [102, 103]. Different biological tissues exhibit distinct light-reflecting characteristics, resulting in varied intensity echoes. The OCT technology can leverage these differences to distinguish layers within the target tissue. To reconstruct tissue architecture, the system must accurately determine the relative positioning of these layers, a task influenced by echo delay time. As the distance increases, the delay in light returning to the detector also increases. Achieving a 10 m resolution for distance measurement in the retina necessitates an echo delay of 30 femtoseconds (30×10^{-15} seconds), a challenging feat with existing detectors [102]. Interferometry techniques are employed to detect such minuscule delays in reflected signals, initially with Time Domain-OCT and later evolving into Fourier Domain-OCTs.

1.10.3 Architectures for OCT Systems

Time-domain and Fourier-domain analyses represent the fundamental techniques for the discrimination and juxtaposition of the reflected light originating from target tissues. Within the realm of Fourier-domain analysis, spectral domain and swept-source techniques emerge as the primary sub-divisions. The research inquiry herein amassed OCT data from both healthy and glaucomatous ocular specimens, employing the Zeiss Cirrus HD-OCT—an apparatus crafted by Carl Zeiss Meditec and situated in Dublin, California. Consequently, this thesis is exclusively dedicated to an exploration of the Cirrus HD-OCT.

Spectral Domain OCT

Spectral Domain OCT (SD-OCT), an OCT technique harnessing Fourier domain transformation, is characterized by its utilization of the spatially encoded frequency domain. In 2000, Leitgeb et al. introduced an adaptation of Fourier domain-OCT spectroscopic measurements within this context [104]. In contrast to Time Domain OCTs (TD-OCTs), SD-OCTs dispense with the need for a moving reference mirror, resulting in a significant enhancement of scanning speed [105]. Commercially available SD-OCT systems operate at scanning rates of 18,000–70,000 A scans per second, outpacing TD-OCT devices by a factor of 200–400. The spectral domain technique relies on broad-bandwidth light sources and employs a fixed mirror, a spectrometer, and a linear CCD to exploit Fourier transformation for interference analysis between the reference and sample beams. This stands in contrast to TD-OCT, which relies on an interferometer equipped with a scanning reference arm and a mechanically movable reference mirror for echo delay time detection. The absence of a mechanically movable mirror translates to a picture acquisition speed increase of 200–400 times in SD-OCT systems. By dispersing and concurrently collecting spectral interference patterns, SD-OCT devices utilize a spectrometer to capture all light echoes in a single instance. This instantaneous data collection enables swift conversion of the gathered spectra into depth information (A-scans) [106]. Augmenting scanning speed and minimizing artifacts from eye movement further heighten resolution [105]. Presently, SD-OCT systems offer an axial resolution of approximately 5 microns, a feature that contributed to their swift adoption as the industry standard in the realm of OCT systems. Notably, Heidelberg Engineering Spectralis OCT and Zeiss Cirrus HD-OCT emerged as prominent SD-OCT platforms following their inception, alongside globally recognized systems like Optovue RT-Vue and Topcon 3D OCT

2000,   which   have   garnered   widespread   usage.
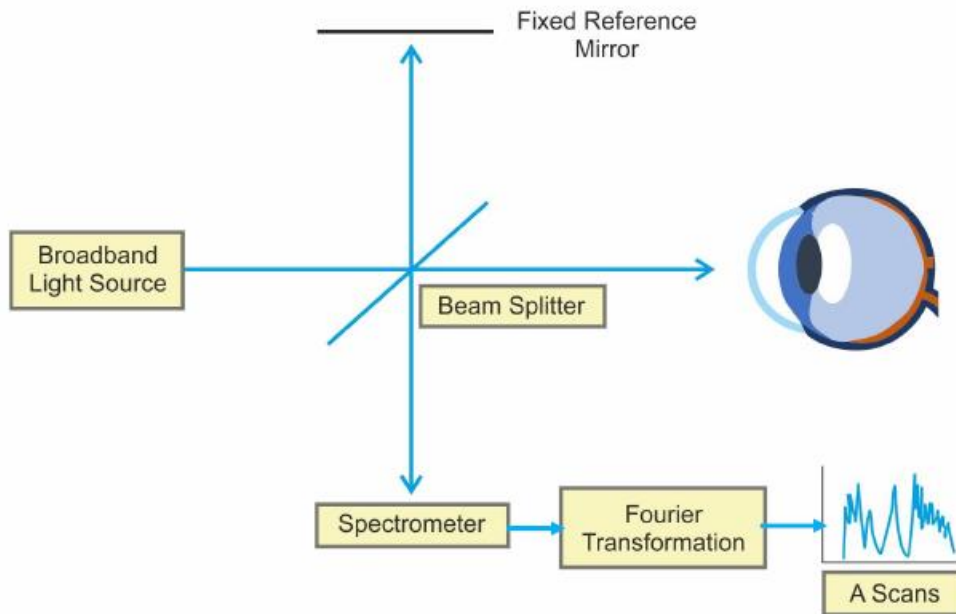
Figure 1-10 Fundamental Principle of Spectral Domain Optical Coherence Tomography (SD-OCT)

1.10.4 Utilization of OCT in Diagnosing and Treating Glaucoma: Exploring its Functionalities

Glaucoma is distinguished by the impairment of retinal ganglion cells (RGCs), which are intricate and substantial neurons. These RGCs extend dendrites that establish connections with bipolar and amacrine cells within the inner plexiform layer (IPL) of the retina. The ganglion cell layer (GCL) is formed by these RGCs, while their axons collectively constitute the retinal nerve fiber layer (RNFL). The convergence of these axons at the ONH gives rise to the neuro-retinal rim. In the lateral geniculate body, RGC axons synapse with the third neuron of the visual pathway.

The introduction of OCT has significantly revolutionized the diagnosis and monitoring of glaucoma due to its ability to objectively and quantifiably detect RGC deterioration [107]. Early glaucoma diagnosis greatly benefits from methods capable of recognizing structural damage [108, 109], which often precedes functional impairment. In the past, diagnosing glaucoma using structural techniques solely relied on decades of clinical observation of changes in ONH images. However, advancements in digital imaging techniques, including scanning laser polarimetry and confocal scanning laser systems, have enabled objective and

quantitative evaluation of ONH and RNFL [110, 111]. OCT, having replaced these methods over the last decade, has emerged as the foremost standard for early detection of structural glaucomatous damage, evaluating RNFL, macular ganglion cells, and ONH changes with high reliability and reproducibility [111, 112]. According to Kuang et al., OCT based on average RNFL thickness measurements might identify glaucomatous damage up to 5 years before initial visual field deficits become apparent in about one third of affected individuals [113].

The 2016 10th World Glaucoma Association Consensus report emphasized that progressive glaucomatous RNFL thinning and neuroretinal rim narrowing are the most reliable criteria for diagnosis, rendering the detection of VF abnormalities unnecessary. Nonetheless, OCT stands as the forefront digital imaging technique for detecting structural damage in glaucoma [114].

1.10.5 RNFL Analysis

Quigley et al.[115] demonstrated the precedence of RNFL alterations over ONH changes. Nevertheless, detecting RNFL thinning through conventional fundus examination poses challenges. To surmount this issue, various imaging modalities such as red-free photography, scanning laser polarimetry, and confocal scanning laser ophthalmoscopy have historically been employed to detect peripapillary RNFL loss. Among these methods, OCT swiftly emerged as the preferred choice for assessing RNFL in glaucomatous eyes, whether in suspected or confirmed cases [103]. Analyzing peripapillary RNFL via OCT involves two main approaches. The first entails conducting a scan to generate a comprehensive three-dimensional map of the peripapillary RNFL in the vicinity of the ONH. Noteworthy manufacturers of these maps include Carl Zeiss Meditec, Topcon Medical Systems, and Optovue Inc., situated in Dublin, California, Oakland, New Jersey, and Fremont, California, respectively. The second approach, garnering significant attention from researchers, encompasses quantifying peripapillary RNFL thickness along a 3.46-millimeter-diameter circular scan centered on the ONH or Bruch's membrane opening (BMO). This circle, referred to as the "calculation circle," stems from the 6 mm by 6 mm optic disc cube and is a feature of Zeiss Cirrus HD-OCT. Data acquisition transpires through a single circular scan of 3.46 mm diameter encircling the ONH, denoted as the "scan circle."

1.10.6 RNFL Thickness Map

SD-OCT instruments can rapidly construct highly precise three-dimensional maps of the RNFLsurrounding the ONH within mere seconds of scanning the ONH vicinity. Illustrative

data (Figure 1-11) showcases scans captured via Cirrus HD-OCT apparatus, presenting the thickness of the RNFL. Utilizing 200 horizontal scans, each comprising 200 A-scans, the Cirrus HD-OCT generates a three-dimensional map of RNFL thickness with a resolution of 200 X 200 A-scans, post scanning a 6 mm by 6 mm area. In the context of an emmetropic eye, the per A-scan coverage on the retina measures 30 microns per side. Comparatively, monitoring RNFL alterations across a 6 mm by 6 mm expanse (equivalent to 200 pixels by 200 pixels) encircling the ONH, as opposed to a single 3.46 mm circumpapillary RNFL scan, could potentially enhance the precision of identifying RNFL degradation.[116, 117]. Leung et al. conducted a study wherein they juxtaposed data from the circumpapillary RNFL calculation circle with the RNFL thickness map. Their observations led them to deduce that the RNFL thickness map notably amplified diagnostic sensitivity by furnishing supplementary spatial and morphological insights into RNFL impairment [117]. Another inquiry by the same researchers unveiled that the most frequently affected region, where RNFL thinning occurred, was along the infero-temporal meridians, roughly 2 mm from the disc center. Given that the radius of the 3.46 mm calculation circle is 1.73 mm, discerning alterations in this locale, beyond the confines of the 3.46 mm circle, mandates an RNFL thickness map scan [118].

1.10.7 The RNFL Calculation Circle

The predominant method for evaluating the loss of retinal nerve fiber layer (RNFL), known as peripapillary RNFL measurement, relies on the calculation of a circular path. During the initial stages of OCT, the resolution and scanning speed of TD-OCT devices were inadequate to produce maps of RNFL thickness. Consequently, a solitary circumpapillary RNFL scan centered on the ONH was employed. In a preceding study by Schuman et al. [119], the diameter of this scanning circle was arbitrarily set at 3.4 millimeters. Subsequently, all manufacturers of OCT devices adopted a standardized scanning circle size of 3.46 mm, which swiftly became the established norm for glaucoma diagnostic investigations documented in medical literature.
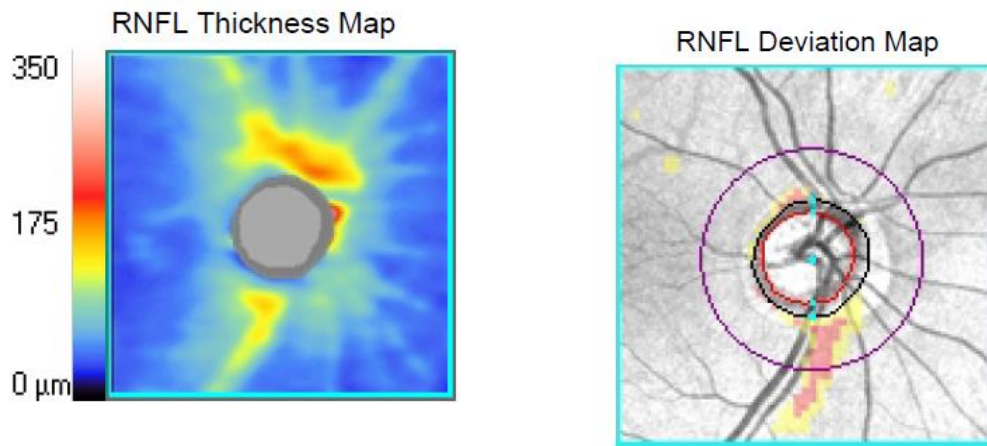
Figure 1-11 On the left side, there is the Zeiss Cirrus HD-OCT displaying a map indicating the thickness of the RNFLusing pseudo-colors. On the right side, an RNFL deviation map is presented.

1.10.8 The computation of the RNFLCalculation Circle and the Temporal-Superior-Nasal-Inferior-Temporal (TSNIT) plots

Various forms can be employed to represent calculations of circumpapillary RNFL or data derived from circular scans. Each manufacturer has their own recommended file format to achieve this goal. Most OCT devices have the capability to measure thickness across different sectors, including average, hemifield, quadrant, and clock hour. Additionally, numerical data are commonly visualized using pseudocolor maps or pie graphs.

The peripapillary circular scan employs an OCT program's segmentation algorithm to measure RNFL thickness. This algorithm also identifies the RNFL. Subsequently, RNFL measurements along the circular calculation or scan path are depicted in a TSNIT plot. This plot starts at the temporal quadrant (9 o'clock for the right eye, 3 o'clock for the left eye), proceeds clockwise for the right eye, and counterclockwise for the left eye. Temporal-Superior-Nasal-Inferior Temporal (TSNIT) designations are employed (see Figure 1-12). RNFL thickness data is presented along the calculation circle, beginning temporally, advancing superiorly, nasally, inferiorly, and concluding temporally. The right eye follows a clockwise direction, while the left eye follows a counterclockwise path. Typically, a double-hump pattern is observable on the TSNIT plot, with the highest RNFL thickness regions in the superior and inferior quadrants. These plots are utilized for data analysis. Probability values indicating abnormal RNFL

thickness are displayed on the TSNIT plot using a four-color scale – white represents the lowest probability, while green, yellow, and red signify progressively lower probabilities.

RNFL measures falling within the red region, indicating values at or below the lowest one percentile from the normative database, are considered abnormally low. Measures between the 1st and 5th percentiles are deemed potentially abnormal and marked with a yellow warning signal. RNFL measures within the middle 90th percentile of the normative database are considered normal and shown in green. RNFL thickness readings exceeding the 95th percentile of the normative database are considered abnormally high and indicated by a white flag [120].

The majority of publications addressing OCT's diagnostic potential in glaucoma primarily utilize average and sector-specific RNFL thickness data from the calculation circle. Among these parameters, superior quadrant RNFL thickness demonstrates the highest diagnostic accuracy [80, 93, 98, 120]. Overall, average and inferior quadrant peripapillary RNFL thickness values exhibit the best diagnostic accuracy. Due to test-retest variability being less than 5 micrometers in current SD-OCT systems for average RNFL thickness, this parameter is highly reproducible [99-101]. While quadrant and sector-specific RNFL thickness data can enhance sensitivity, they may reduce specificity when utilized for glaucoma diagnosis. Abnormalities in average RNFL thickness can detect glaucoma with 95% specificity in up to 35% of glaucoma suspect eyes four years before detectable VF loss and in up to 19% of eyes eight years before detectable VF loss [113]. This level of accuracy can be achieved in up to 35% of cases. RNFL thickness naturally decreases with age. Patient RNFL thickness findings are compared to age-matched normative values, with commercial OCT systems offering normative databases across various age groups. Distribution of normal RNFL thickness measurements is influenced not only by age but also by axial length and race. To address this, several OCT devices incorporate normative data from individuals of diverse ethnicities and varying degrees of myopia.
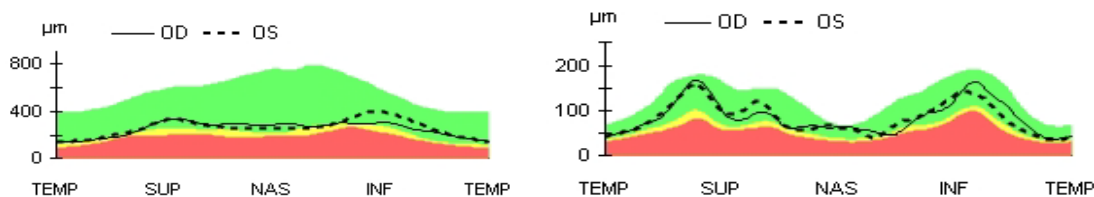
Figure 1-12 (Left) Neuro-retinal Rim Thickness, (Right) RNFL Thickness

1.10.9 Optic Nerve Head Analysis

Every OCT apparatus conducts an examination of the ONH, yielding insights into its condition. Variations in scan dimensions and quality arise across OCT devices. Multiple characteristics of the ONH are provided, including the cup-to-disc ratio, neuroretinal rim area, and neuroretinal rim volume, dictated by software algorithms and technological capabilities. Establishing the ONH boundary is an initial step in ONH analysis, where the Bruch's membrane is often used as a proxy, facilitated by its visibility on OCT images.

Determining the BMOis achieved through segmentation, a relatively uncomplicated process across available B-scans. The ONH centroid is computed based on the accurately identified BMO, enabling precise centering of the calculation ring for RNFLanalysis. This method is universally employed, exhibiting exceptional repeatability and reproducibility [121, 122]. Moreover, BMO serves as a reference point for defining the optic disc boundary, in contrast to earlier OCT devices that utilized the optic disc border ring positioned by the internal limiting membrane (ILM) [123, 124]. A revised approach to assessing the neuroretinal rim area measures the minimum distance between BMO and ILM, termed BMO-MRW (minimum rim width) (Figure 1-13). Zeiss' Cirrus HD-OCT, since its 2010 software release, employs a BMO-MRW-like approach to extract ONH data. Though often referred to as MRW, Cirrus software reduces areas rather than distances to locate the neuroretinal rim border. Normative limits for ONH in Cirrus are customized for disc size and tilt. Results are depicted through TSNIT-type graphs and summary tables, with the BMO-MRW method excelling in diagnostic accuracy and structure-function correlation [124, 125].

The debate persists regarding the superior diagnostic efficacy between ONH parameters and RNFL findings for early glaucoma detection. Some studies assert ONH parameters' superiority, while others highlight RNFL and macular measures' effectiveness [126, 127]. ONH attributes aid in distinguishing non-glaucomatous optic neuropathies, crucial when OCT shows RNFL and GCC/GCL + IPL damage alongside normal ONH parameters.

Enhancing early glaucoma diagnosis involves combining RNFL, ONH, and macular measures due to their variable vulnerability over time. Accurate early diagnosis hinges on comprehensive analysis of these parameters [107].
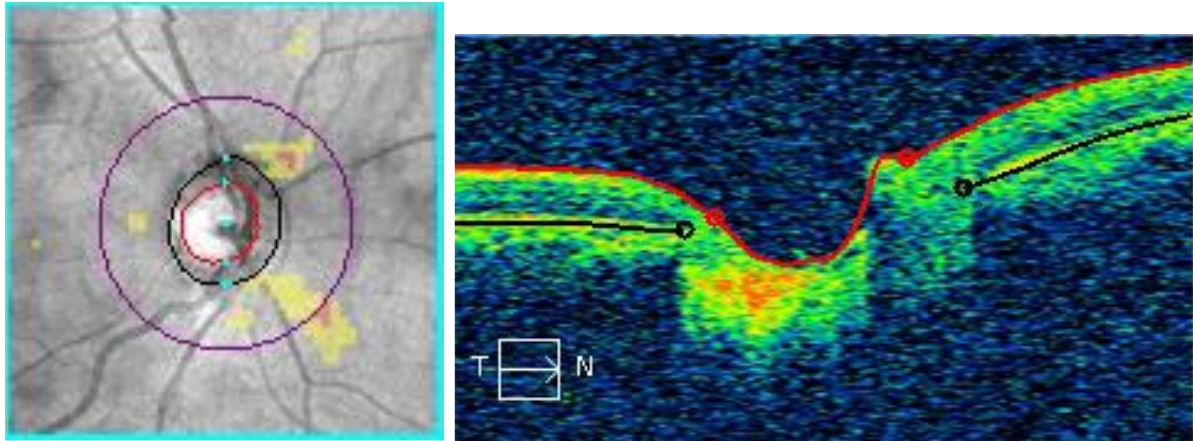
Figure 1-13 Detection of BMOand the BMO-Minimum Rim Width (BMO-MRW) Identification

1.10.10 The Normative Databases of Cirrus HD-OCT

The comparative analysis between patient-specific data and the instrument's normative database is a fundamental element of every OCT system. These comparisons significantly influence clinical decision-making for healthcare professionals. It's imperative for clinicians to acknowledge the limitations that govern such comparisons. Within the Cirrus HD-OCT normative database, data from 282 individuals without ocular pathologies and within specific refractive error limits are encompassed, including RNFL, GCL+IPL, and ONH measurements. This database employs age-matched information for RNFL and GCL+IPL evaluations, along with age and disc size-matched data for ONH assessments, specifically for patients aged 18 and above. Given the absence of participants under 18 during data collection, normative database evaluations for individuals below 18 years old are unfeasible. The presumption is that measurements from the included eyes follow a normal distribution.

1.11 Thesis aims

Title: Statistical and Computational Analysis Of Retinal Images to Understand Ocular Disorders

The main objectives of the present study are to:

(1) generate OCT NRR data in the form of 180 circular measurements of NRR thickness in a given eye.

(2) introduce CIFU, a computational pipeline for **CI**rcular **FU**nctional data modeling and analysis that is demonstrated using the OCT NRR dataset, and

(3) address the disparity of representation of the Asian Indian population in normative OCT databases.

(4) design of an analytical framework that is capable of defining and using a *directional measure* of RNFL decay based on OCT data and then applying it for focused detection of structural characteristics in glaucomatous eyes from an Indian population.

(5) we developed an analytical framework based on novel applications of circular statistics to OCT data analysis. This includes specifying a directional representation of RNFL decay in a given glaucomatous eye relative to its corresponding normal population. New indices were introduced to measure eye-specific global and local RNFL decay.

1.12 Data and Inclusion Criteria

Participants for this thesis were drawn from two distinct studies: the LVPEI Glaucoma Epidemiology and Molecular Genetic Study (LVPEI-GLEAMS), which is a population-based study, and the Longitudinal Glaucoma Evaluation Study (LOGES), a cross-sectional study. These studies were both conducted by the L.V. Prasad Eye Institute (LVPEI) located in Hyderabad, India. Ethical approval for data protocols was granted by the LVPEI Institutional Ethics Committee for both LVPEI-GLEAMS (LE-08131) and LOGES (LEC 11-252) prior to their implementation. The methodologies were rigorously reviewed and endorsed by the LVPEI ethics and review committee, adhering closely to the principles outlined in the Declaration of Helsinki. All participants provided informed consent prior to their involvement in the research. The data encompassed 3973 healthy eyes, consisting of 1981 right eyes (OD) and 1992 left eyes (OS), derived exclusively from the LVPEI-GLEAMS. Additionally, a total of 331 Glaucoma-affected eyes were included, comprising 65 eyes from LVPEI-GLEAMS (OD-36, OS-29) sourced from 55 individuals, and 266 eyes from LOGES (OD-143, OS-123) derived from 155 individuals.

Inclusion criteria mandated an age of ≥ 40 years, regardless of gender, with a best-corrected visual acuity of 20/40 or better, a spherical equivalent (SE) falling within ± 6 diopters, exceptional quality stereo optic disc images, and the absence of any media opacities. Exclusion criteria involved intraocular surgery within the previous six months and any retinal or other neurological conditions that might impact structural measurements obtained through SD-OCT.

To ensure accurate analysis, OCT scan images of inferior quality, defined by a signal strength < 6, motion artifacts, blinking artifacts, misidentification of retinal layers, and off-center artifacts, were excluded from the study. Past research, including work by some of the authors of this study, has demonstrated a connection between diminished signal strength and reduced accuracy in nerve fiber layer thickness measurement using OCT, which could potentially lead to incorrect interpretation of glaucomatous damage. For this study, the Cirrus SD-OCT platform's manufacturer-defined criterion for acceptable scan quality was a signal strength of 6 or above on a scale of 0 to 10.

Healthy eyes were characterized by the absence of both anterior and posterior pathology. Each digital optic disc photograph underwent independent evaluation by three glaucoma specialists, with masking to prevent influence from other clinical data or imaging outcomes. Any cases of disagreement among specialists resulted in the exclusion of the respective eyes from the study. Comprehensive ophthalmic examinations were conducted for all participants, encompassing thorough medical and systemic history, measurements of best-corrected visual acuity, slit-lamp photography, Goldmann applanation tonometry, gonioscopy using a Sussman four-mirror gonioscope, dilated fundus examination, assessment of central corneal thickness, and Humphrey visual field testing with the 24-2 Swedish Interactive Threshold Algorithm. Visual fields were considered valid if the rates of false positives, false negatives, and fixation losses were all below 20%, and if stereophotographs of the optic disc demonstrated satisfactory quality.

Chapter 2

Introduction to Circular and Functional Data Analysis

2.1 Circular Statistics: Introduction

The field of statistics that deals with datasets represented as points around the unit circle is referred to as "circular statistics." This term is employed to distinguish such data from the more familiar linear data. Circular data emerges across diverse scientific domains, from studying animal migration patterns in ecology (Rivest et al., 2015)[128] to investigating protein structure (Mardia et al., 2006 [129] and neuronal activity (Rutishauser et al., 2010) [130].

Measurements of two-dimensional directions are quantified as angles relative to a chosen "zero direction" or reference point. The concept of positive direction is determined by the sense of rotation, either clockwise or counterclockwise. As directions lack magnitude, they can be conveniently represented as points on a unit circle centered at the origin or as unit vectors connecting these points to the origin. These observations in two-dimensional space are commonly referred to as circular data due to their circular arrangement. The prevalence of circular data spans diverse scientific fields like physics, biology, geology, and medicine.

Circular data are visualized on the unit circle, where any angle θ between 0 and 360 degrees corresponds to specific observations. The angle θ represents the deviation from the reference direction, forming the angle between the unit vector and the starting point. Given the periodic nature of these measurements and the circular continuity, the measurement range wraps around, i.e., $0 = 2\pi$ and $\theta = \theta + p * 2\pi$, where p is an integer. Consequently, methodologies addressing directional data must meticulously consider distance estimation between pairs of points due to these distinctive attributes.

Due to these inherent properties, the analysis of circular or directional data significantly diverges from conventional "linear" statistical analysis of univariate or multivariate data. Common linear techniques and metrics often prove misleading or entirely inapplicable owing to the necessity for "invariance" of statistical methods and metrics concerning the choice of

arbitrary zero-direction and rotational sense. Established measures like sample mean, variance, and higher moments are erroneous in this context. Likewise, analytical tools such as moment-generating functions and other generating functions lose their efficacy. Fundamental concepts like correlation, regression, and their associated statistical indices require redefinition to suit directional data..

2.2   Descriptive Statistics

Circular data can be depicted in either angular or point forms along the perimeter of a unit circle. A distinct directional value can be attributed using a pair of coordinates. This purpose can be fulfilled by employing the Cartesian coordinate system, centered at origin O, consisting of two (or three) perpendicular axes, namely X and Y passing through O. The representation of any point P on the plane involves either the use of rectangular coordinates (X, Y) or polar coordinates (r, α), where 'r' signifies the distance from the origin and 'α' indicates the direction. It's important to note that the exceptional point O, denoting the origin, possesses r = 0 and lacks a specific direction, rendering it undefined, implying an absence of α. The conversion between polar and rectangular coordinates, and vice versa, is straightforward, facilitated by trigonometric functions such as sine and cosine. Taking into consideration the polar coordinates of a given point P: (r, α) [146].
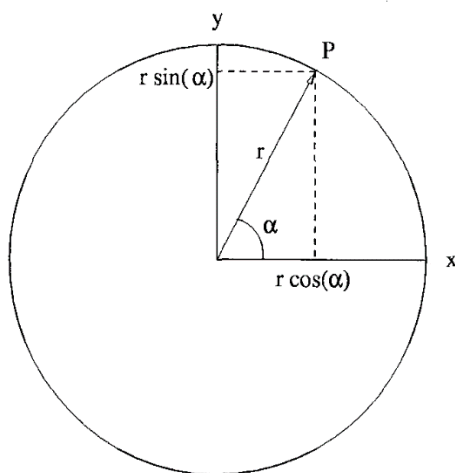
Point P's rectangular coordinates are given by

x = r cosα, y= r sinα.

In the context of directional analysis, our focus is solely on the vector's orientation rather than its magnitude. To facilitate this, we designate these vectors to possess a unit length, denoted as r=1.

Consequently, every cardinal point on the circumference of the unit circle is denoted by a point P. An alternative approach involves utilizing the angle itself to pinpoint a position along the unit circle's perimeter. This conversion from polar to rectangular coordinates becomes uncomplicated when point P is positioned precisely along the unit circle's circumference, leading to the following relationship:

(1, α) is equivalent to (x = cosα, y = sinα).

(2.1)

2.2.1 Measure of Center

A common measure of center for circular data is the circular mean, often represented by the term "circular average". This measure takes into account the circular nature of the data and provides a sense of where the data tends to cluster around the circle.

Here's an example of calculating the circular mean for a set of compass directions:

Let's say we have a dataset of wind directions recorded in degrees:

{30°,45°,350°,10°,20°}

To calculate the circular mean of these wind directions, we need to account for the circular nature of angles. We can't simply calculate the arithmetic mean like we do with linear data. Instead, we use the concept of vector addition to find the mean direction. One common way to calculate the circular mean is by using the circular mean vector:

- Convert each wind direction angle to its corresponding unit vector on the unit circle
- Calculate the mean of these unit vectors.
- Convert the mean unit vector back to an angle.

Suppose we have a set of circular observations denoted as α_1, α_2, ..., α_n, which are expressed as angles. The process involves the conversion of each observation from polar coordinates to rectangular coordinates., i.e.[146]

$$(\cos\alpha_i, \sin\alpha_i), i=1,2,\ldots,n.$$

(2.2)

The resultant vector can be acquired through the component-wise addition of these n-unit vectors

$$R=(\sum_{i=1}^{n} \cos\alpha_i, \sum_{i=1}^{n} \sin\alpha_i)=(C,S)$$

(2.3)

Let,

$$R=\lVert R \rVert=\sqrt{C^2 + S^2}$$

(2.4)

The direction of the resultant vector (R), representing the length of the final vector, is indicated as $\bar{\alpha}_0$. This $\bar{\alpha}_0$ signifies the circular mean direction in this context.

$$\bar{\alpha}_0=\arg\left\{\sum_{j=1}^{n} \cos\alpha_j + i \sum_{j=1}^{n} \sin\alpha_j\right\}$$

Alternatively using equations

$$\cos \bar{\alpha}_0=\frac{C}{R}, \sin \bar{\alpha}_0=\frac{S}{R}$$

(2.5)

More specifically, it is given by the tangent's "quadrant-specific" inverse, which will refer to as,

$$\bar{\alpha}_0= \arctan^*(S/C),$$

(2.6)

Where,

$$\bar{\alpha}_0= \text{arctan*(S/C)}=\begin{cases} \arctan\left(\frac{S}{C}\right), if\ C > 0, S \geq 0, \\ \frac{\pi}{2}, if\ C = 0, S > 0, \\ \arctan\left(\frac{S}{C}\right) + \pi,\ if\ C < 0, \\ \arctan\left(\frac{S}{C}\right) + 2\pi, if\ C \geq 0, S < 0, \\ Undefined, if\ C = 0, S = 0 \end{cases}$$

(2.7)

It is necessary to define a quadrant-specific inverse of the tangent in a reasonably convoluted manner because $\tan(\theta) = \tan(\theta + \pi)$. Therefore, there are two inverses for any given θ. Considering that arctan is often defined to take values in $(-\frac{\pi}{2}, +\frac{\pi}{2})$, gives us a unique inverse on $[0, 2\pi)$ that is accurate and considers the signs of cosine and sine functions.

To calculate the cosine and sine of every angle, we first convert a given set of angles into rectangular coordinates. Equation (2.3) is used to get the resultant vector by component-wise adding them. Using Equation(2.7), we can then determine the direction of this resultant vector, $\bar{\alpha}_0$. This represents the circular mean direction for the specified collection of angles.

The scenario in which data is uniformly distributed across a circle, displaying no bias towards any particular direction, is indicative of a situation where the resultant vector holds a magnitude of zero. In such cases, it is clear that the statistical observations do not exhibit any preferred or distinct orientation.

2.2.2 The Median Direction

The symbol $\bar{\theta}$ represents a sample median direction derived from the set of angles $\theta_1, \ldots \theta_n$. Let $\emptyset$ be an arbitrary angle meeting two conditions: (i) approximately 50% of the data points fall within the arc$[\emptyset, \emptyset + \pi)$, and (ii) a majority of the data points are positioned closer to to $\emptyset$ than to $\emptyset + \pi$. Notably, when the sample size n is odd, the sample median coincides with one of the data points. In situations where n is even, utilizing the sample median as an intermediary between two suitably adjacent data points proves advantageous.

2.2.3 The Mean Resultant Length (MRL) and Circular Variance (CV)

The value of the mean resultant length R can be expressed as follows

$$\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}.$$

<div align="right">(2.8)</div>

R will be 1 if the directions $\alpha_1, \alpha_2, \ldots \alpha_n$ are closely grouped. On the other hand, R will be almost 0 if $\alpha_1, \alpha_2, \ldots \alpha_n$ are far dispersed.

The average resultant length, denoted as $\bar{R}$, holds greater significance compared to any dispersion measurement in the context of most descriptive analyses and inferential applications. However, it might occasionally be helpful to take measurements of the dispersion of circular data into account for purposes of comparison with data on the line. The sample **circular variance**, which is the simplest of them, is defined as

$$V = 1 - \bar{R}$$

<div align="right">(2.9)</div>

2.2.4 Circular Standard Deviation:

The standard deviation of data on the line can occasionally serve as an analogue for circular data. A transformation of the sample variance V is one method for getting such a summary statistic.

$$\vartheta = \{-2\log(1-V)\}^{1/2} = \{-2\log \bar{R}\}^{1/2}$$

<div align="right">(2.10)</div>

ϑ takes values in $[0, \infty]$, whereas V takes values in $[0,1]$.

For small V the Equation reduce to

$$\vartheta \approx (2V)^{1/2} = \{2(1-\bar{R})\}^{1/2}$$

<div align="right">(2.11)</div>

2.3   Circular Probability Distributions

The central notion revolves around the unit circle's perimeter, encompassing the predominant share of the total probability within a circular distribution. This kind of distribution serves the purpose of assigning probabilities to diverse orientations or establishing a directional pattern due to the unique correspondence between each point along the circle's boundary and a specific

direction. The measurement unit for this circular context is represented by radians, denoting the scope of a circular random variable (RV) denoted as θ. This variable's domain can be defined as [0, 2π) or alternatively as [-π, π). Circular distributions essentially bifurcate into two categories: discrete, which allots probabilities discretely to a countable array of directions, and continuous, which pertains to the Lebesgue measure along the circle's circumference. In cases of continuous data, a probability density function (PDF) denoted as f(θ) comes into play, accompanied by the subsequent fundamental characteristics.

i. $f(\theta)$ is greater than or equal to 0;
ii.   The integral of f(θ) with respect to θ over the interval [0, 2π) is equal to 1.
iii. f(θ) maintains the property of periodicity, specifically f(θ) = f(θ + k * 2π) for any integer k.

2.3.1 Wrapped Distributions

By applying a modulo 2π reduction, it becomes possible to transform any linear random variable (r.v.) X that is situated on the real number line into a circular random variable, thereby establishing its definition as a circular r.v.

$$\theta = X(\text{mod } 2\pi)$$

This procedure involves the transformation of the real number line onto a unit-radius circle, concurrently summing the probabilities of all coinciding points such as x = θ, θ ±2π, θ ±4π, and so forth. Evidently, this corresponds to a many-to-one mapping. In the context where g(θ) signifies the density on the circle and f(x) denotes the density of the real-valued random variable, the relationship can be expressed as follows.

$$g(\theta) = \sum_{m=-\infty}^{\infty} f(\theta + 2\pi m), \qquad 0 \le \theta < 2\pi.$$

This technique allows for the generation of both discrete and continuous wrapped distributions utilizing the described approach..

2.3.2 Circular Distributions

Uniform Distribution

The Circular Uniform (CU) distribution with a consistent density arises when uniform probability is evenly distributed around an entire circle.

$$f(\theta) = \frac{1}{2\pi}, \quad 0 \leq \theta < 2\pi$$

This probability distribution is referred to as the isotropic or random distribution, as it equally favors all directions. With the sole exception of $\alpha 0$, which holds a value of 1, all trigonometric moments of orders p, namely αp and βp, are rendered as 0. The absence of a well-defined mean or favored direction in this distribution stems from the fact that the length of the initial trigonometric moment, denoted as ρ in this context, is quantified as 0.

2.3.3 Cardioid Distribution

The cardioid curve is the basis for this distribution and has the probability density function.

$$f(\theta; \mu, \rho) = \frac{1}{2\pi}\{1 + 2\rho \cos(\theta - \mu)\},$$

$$0 \leq \mu < 2\pi, -\frac{1}{2} < \rho < \frac{1}{2}$$

This distribution is symmetric and unimodal about p. Given by are its trigonometric moments.

$$\varphi_1 = |\rho|. \, e^{i\mu} \text{ and } \varphi_p = 0 \text{ for } p \geq 2,$$

so that the mean direction is represented by μ and the concentration by ρ.

2.3.4 The Circular Normal (CN) (von Mises Distribution)

A circular random variable denoted as θ is characterized by a Von Mises distribution or a Circular Normal (CN) distribution when its probability density function is expressed as follows:

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0 \, (k)} e^{k\cos(\theta - \mu)}, \qquad 0 \leq \theta < 2\pi,$$

For the purpose of the thesis, let $0 \leq \mu < 2\pi$, and $k \geq 0$ be considered as defining parameters. Within this context, the modified Bessel function of the first kind and order zero, denoted as

$I_0(k)$, assumes significance as an integral component within the normalizing constant. This function can be expressed as follows:

$$I_0(k) = \frac{1}{2\pi} \int_0^{2\pi} \exp(k\,cos\theta)\,d\theta = \sum_{r=0}^{\infty} \left(\frac{k}{2}\right)^{2r} \left(\frac{1}{r!}\right)^2$$

2.3.5 Characteristics of the von Mises Probability Density Function

Symmetry and Modal Orientation of Circular Normal Density:

The symmetry inherent in the cosine function indicates a symmetrical distribution of density around the axis represented by μ, along with its counterpart μ+π.

Modal Orientation at μ:

Given that the circular normal density reaches its peak at θ=μ, the direction corresponding to the highest value aligns with μ, as the cosine function achieves its maximum value at zero.

$$f(\mu) = \frac{e^k}{2\pi I_0(k)}$$

Antimode at $(\mu \pm \pi)$: Again, since $\cos \pi = -1$ is the minimum value, the minimum density is found at the other end of the spectrum when $\theta = (\mu \pm \pi)$.

$$f(\mu \pm \pi) = \frac{e^{-k}}{2\pi I_0(k)}$$

As a result, $(\mu \pm \pi)$ is the anti-modal direction.

Part of k:

$$\frac{f(\mu)}{f(\mu \pm \pi)} = e^{2k}$$

Because of this, the ratio of $f(\mu)$ to $f(\mu \pm \pi)$ will increase as k increases, indicating a more potent concentration in the direction of the population's mean, or μ. k, then, is a quantity that gauges the concentration in the direction of the mean vector μ (see the figure).
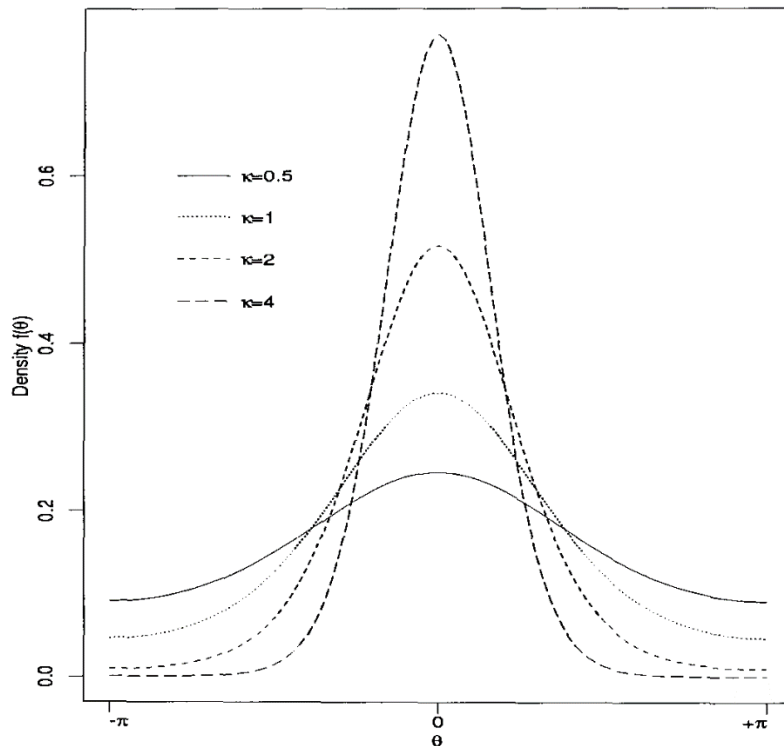
Figure 2-1 Circular (von Mises) normal distributions characterized by k = 0.5, 1, 2, and 4 with $\mu = 0^0$.

2.4   Functional data analysis

The concept of "functional data analysis" (FDA) pertains to the examination and conceptualization of data represented in the form of functions, images, pictures, or even more generalized objects. At its core, functional data comprises functions, which serve as the fundamental components. In the context of a random sample, each individual participant's data often encompasses one or more functions. While the terminology "functional data analysis" was initially coined by Ramsay (1982) [131], and further developed by Ramsay & Dalzell (1991) [132], the origins of this field can be traced back even earlier to the works of Grenander (1950) [133] and Rao (1958) [134]. Notably, functional data possess an essentially infinite-dimensional nature. This inherent high dimensionality poses both theoretical and computational challenges, which vary depending on the method of collecting the functional data. However, it's important to recognize that the elevated or infinite-dimensional structure of the data provides a valuable source of knowledge, offering a diverse range of opportunities for exploration.

The behavior of an underlying smooth function that is believed to be the source of the observations is frequently reflected in data sampled along a continuum, typically across time.

Functional data analysis is one of the newest subfields of statistics, which involves representing and analyzing sequences of discretely observed data as separate entities using functions. The FDA has an advantage over normal statistical approaches in that it can extract more information from underlying functions, derivatives, and primitives. The objectives of the FDA are essentially the same as those of standard statistical methods.

Converting the individually observed data to functions or curves is one of the procedures of the FDA. Therefore, many curves estimating techniques are crucial in FDA. Today, when very accurate monitoring equipment is frequently employed in many fields, and measurement mistakes can be disregarded, curve fitting frequently involves interpolation. However, the modeling may involve smoothing if the observed data contains errors that need to be eliminated.

Important aspects of each curve can be lined up so that they roughly occur at the same time by registering or aligning the functions. The common pattern among the curves is thus frequently represented in a much more realistic manner when we average the curves.

2.4.1 The objectives of functional data analysis encompass

Functional data analysis shares similar objectives with other branches of statistics, encompassing the subsequent aims:

I. The objective involves presenting the data in formats conducive to facilitating further analytical investigations.

II. To effectively convey information, it is essential to present the facts in a manner that highlights diverse characteristics.

III. To explore notable origins of diversity and trends within the dataset

IV. Utilizing information derived from input or autonomous factors to explain the fluctuations observed in an outcome or reliant variable.

V. to analyze multiple datasets, potentially containing distinct sets of repeated instances of identical functions or different functions within a comparable collection of repeated instances, all in the context of particular variations.

2.4.2 Functional data: Summary statistics

Functional Measures of Means and Variances

Functional data can be analysed using the same traditional summary statistics for univariate data. The mean function with numerical values

$$\bar{x}(t) = N^{-1} \sum_{i=1}^{N} x_i(t)$$

is the average of the functions point-wise across replications The variance function var has values in a similar manner.

$$\mathrm{var}_X(t) = (N-1)^{-1} \sum_{i=1}^{N} [x_i(t) - \bar{x}(t)]^2$$

The variance function's square root yields the standard deviation function.

Functional Measures of Covariance and correlation

The covariance function, calculated across all instances of t_1 and t_2, serves as a comprehensive summary and is subjected to computation.

$$\mathrm{COV}_X(t_1,t_2) = (N-1)^{-1} \sum_{i=1}^{N} \{x_i(t_1) - \bar{x}(t_1 \{x_i(t_2) - \bar{x}(t_2)\}$$

The connected correlation function pertains to

$$\mathrm{corr}_X(t_1,t_2) = \frac{cov_X(t1,t2)}{\sqrt{var_X(t_1)var_X(t_2)}}$$

2.5 Techniques employed for the modeling of functional data

The core principle of functional data analysis postulates that observed data functions should be treated as distinct entities rather than mere collections of individual observations. In the context of observed data, the term "functional" pertains to the inherent structure of the data as opposed to its stated format. In practical terms, functional data is often discretely recorded and observed as pairs (t_j, y_j), where y_j captures a snapshot of the function at a given time point t_j, possibly influenced by measurement inaccuracies. While we commonly refer to time (t_j) as the continuous domain over which functional data is recorded, it's important to acknowledge the potential presence of other continua, such as spatial position, frequency, weight, and similar factors.

Functional data usually are seen discretely. A functional observation of a single function x(t) is said to consist of n pairs (y_j, t_j), where the scalar y_j is an observation of the function $x(t)$ at the time t_j. In writing, we say,

$$y_j = x(t_j) + \epsilon_{j,} \quad j=1, \ldots n.$$

(2.11)

where ϵ_j, the error term is presumed to be independent, identical, and distributed normally with a mean of zero and a constant variance. Additionally, we assume that the function $x(t)$ is reasonably smooth (continuous) on $\mathcal{T}$ and that the range of t is bounded on the interval $\mathcal{T} = [t_1, t_n]$.

2.6   Creating the best-fit representation of data through the least squares method using a basis system.

Letting the function be a linear combination of K known basis functions $(\emptyset_1, \emptyset_2, \dots, \emptyset_K)$ is one of the most popular methods of predicting $x(t)$.

$$x(t) = \sum_{k=1}^{K} c_k \emptyset_k(t)$$

(2.12)

If we calculate the coefficients of the expansion c_k by minimizing the least squares criterion, we can produce a straightforward linear smoother.

$$\text{SSE} = \sum_{j=1}^{n}\left[y_j - \sum_{k=1}^{K} c_k \emptyset_k(t_j)\right]^2$$

(2.13)

Introducing the following vector and matrix definitions:

$$c = (c_1, c_2, \dots c_K)^T$$

(2.14)

$$y = (y, y_2, \dots y_n)^T$$

(2.15)

$$\Phi = \begin{bmatrix} \phi_1(t_1) & \phi_2(t_1) & \cdots & \phi_K(t_1) \\ \phi_1(t_2) & \phi_2(t_2) & \cdots & \phi_K(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(t_n) & \phi_2(t_n) & & \phi_K(t_n) \end{bmatrix}$$

(2.16)

Then Equation (2.3) can be expressed as,

$$SSE = (y - \Phi c)^T (y - \Phi c)$$

(2.17)

By differentiating (2.7) with respect to the coefficient vector c, setting the Equation to zero, and then solving it for c, the least squares estimate of the coefficient vector c is achieved. This results in the estimation below:

$$\hat{c} = (\Phi^T \Phi)^{-1} \Phi^T y$$

(2.18)

The x(t) estimator is now calculated using the expected values of c.

2.7   Representing functions by basis functions

A basis function system comprises a set of distinct functions, denoted as k, which are mathematically orthogonal and have the capacity to effectively approximate any given function through the utilization of a weighted sum or linear combination of a considerable quantity represented as K, of these functions.

The basis functions differ in their strengths and limitations and should be chosen based on the characteristics of the function being estimated. Many good function estimate bases may have undesired high-frequency oscillations between observation sites that impact the derivatives. As a result, one should typically select basis functions such that one or more of the approximation's derivatives behave sensibly. It is also preferable for computational reasons to select as few K basis functions as possible.

It is ideal for basis functions to exhibit similarities with the functions being approximated. This quality streamlines the process of achieving a robust estimation using a limited set of basis functions, denoted as K. As K decreases, the capacity of the basis functions to faithfully represent certain attributes of the data enhances. This relationship between a smaller K value and an improved alignment of basis functions with distinct data features becomes more pronounced.

- The extent of available degrees of freedom for hypothesis testing and the calculation of accurate confidence intervals directly impacts the scope and precision of our analytical capabilities.
- requiring reduced computation effort, and
- the greater the chance that the coefficients themselves will develop into meaningful descriptions of the data.

Although there are numerous different basis systems, Fourier, Polynomial, Splines, and Wavelets are the most often utilized basis functions. There aren't any universal basis functions that are ideal in every circumstance. The many basis functions are briefly described in the paragraphs that follow.

2.7.1 The Fourier basis system for periodic data

A mathematical tool called the Fourier basis system is used to analyse periodic data, or data that repeats itself after a certain amount of time. The Fourier basis system is based on the Fourier series, which is a means to aggregate sine and cosine waves of various frequencies to describe a periodic function.

A periodic function may be decomposed into a collection of smaller periodic functions using the Fourier series. Any periodic function has the potential to be expressed as a linear amalgamation of these base functions because these simpler functions—also referred to as the Fourier basis functions or Fourier basis vectors—form a complete orthonormal system.

The Fourier basis functions are defined by the bases on an interval $\mathcal{T} = [t_1, t_n]$.

$$\emptyset_0(t) = \frac{1}{\sqrt{T}}$$

$$\emptyset_{2r-1}(t) = \sqrt{\frac{2}{T}} \sin(r\omega t)$$

$$\emptyset_{2r}(t) = \sqrt{\frac{2}{T}} \cos(r\omega t).$$

The relationship $\omega = 2\pi/T$ connects the constant ω to the duration of the interval $T = t_n - t_1$. The first five Fourier basis functions on the range $[0, 1]$ are displayed in Figure 2-5.
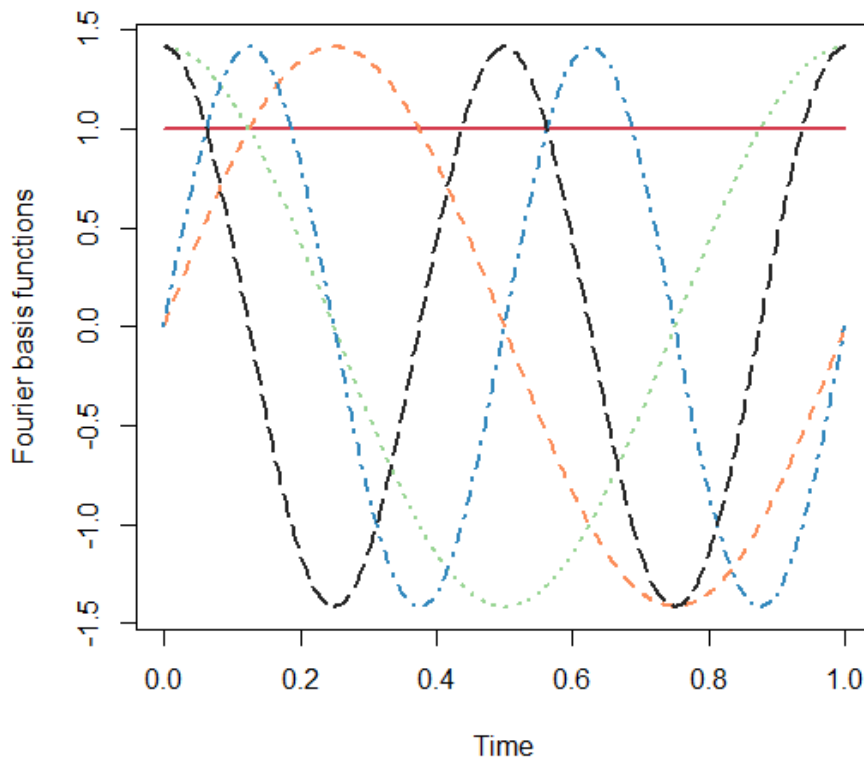
Figure 2-2 Shows the first five Fourier basis functions on the interval [0; 1], constructed in R

The Fast Fourier Transform (FFT) technique allows for rapid computation of the coefficient values c_k. Normally, it requires $\mathcal{O}(n^2)$ operations, however the FFT can reduce the computation time to $\mathcal{O}(n\,log_2 n)$ operations for all the coefficients if the time values t_j are evenly spaced on the interval $\mathcal{T}$ and if n is a power of two. The Fourier series is one of the most famous mathematical tools due to its versatility and ease of use. The Fast Fourier Transform (FFT) has made it even more accessible by allowing for the rapid computation of derivatives. This makes it an excellent choice for solving various mathematics, engineering, and physics problems. Moreover, the fact that many derivatives can easily be derived from the Fourier series makes it a highly versatile tool that can solve many problems [135].

Fourier series is beneficial for periodic data that lacks prominent characteristics and has a consistent order of curvature throughout. The degree of smoothness in the approximation is ascertained by the quantity K of basis functions employed. A low number of K typically leaves out significant characteristics that are known to be part of the estimated function, whereas a

high value produces a function that more closely matches the data. The calculated function, however, might show undesirable high-frequency oscillations between observation points if numerous basis functions are used.

2.7.2 Polynomial:

The monomial basis is one of the first K polynomial basis functions.

$$\emptyset_k(t) = (t - \omega)^k, k = 0, \dots, K - 1$$

where a shift parameter, ω, is used. The first five polynomial basis functions on the range [0,1] are displayed in Figure 2-6 with the shift parameter set to 0. According to[136], these basis functions frequently produce a nearly singular cross-product matrix $\Phi^T\Phi$ and the parameter ω therefore must be picked with care.
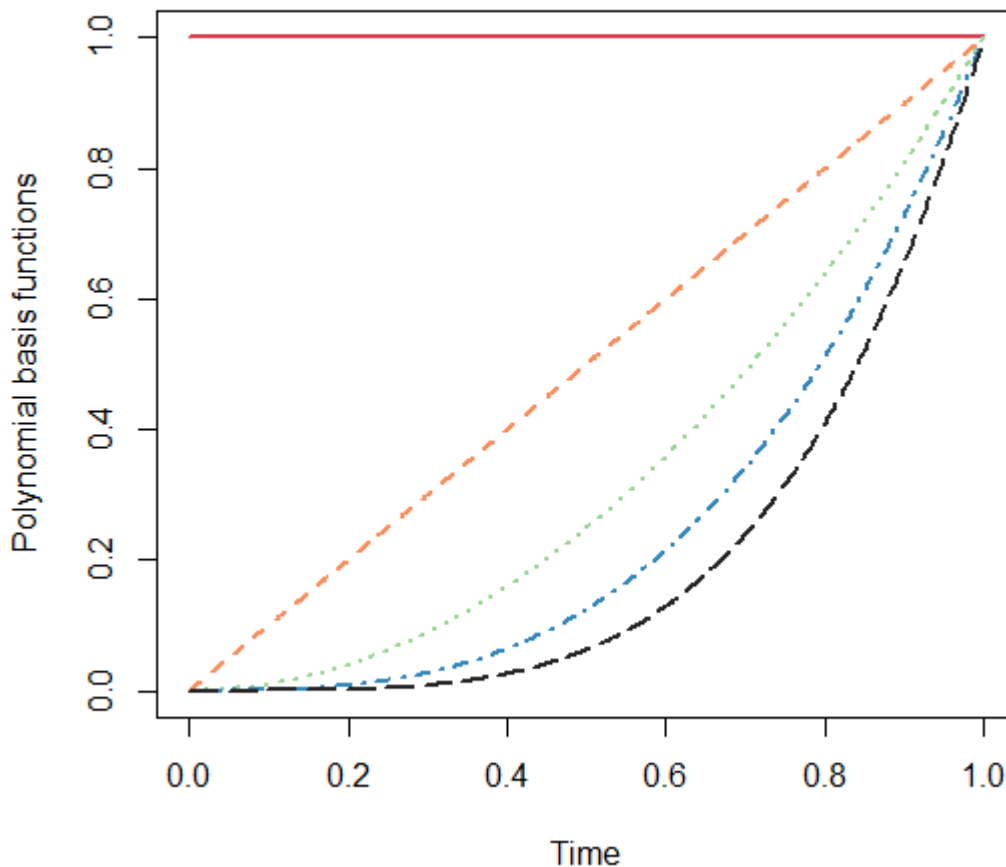
Figure 2-3 The first eight polynomial basis functions on the interval [0; 1] and where ω = 0, created in R

However, if the time values t_j are evenly spaced on the interval T, then it is possible to determine all of the coefficients c_k in $\mathcal{O}(n(K-1))$ operations rather than $\mathcal{O}(n(K-1^2))$ operations. This is because the equal spacing of the time values t_j simplifies the problem.

As a consequence of the quick localised oscillation, the polynomial basis functions are unable to demonstrate strong local behaviour unless they make use of a big K, which leads to inaccurate derivative estimations. The polynomials have a propensity to fit the data satisfactorily in the middle of the interval, but their behaviour close to the endpoints is unsatisfactory.

2.7.3 Splines and B-splines

Both splines and B-splines are basis function types utilised in numerical analysis, computer graphics, and other disciplines that call for function approximation.

Splines are a kind of polynomial function with a piecewise-defined that is used to approximate complicated curves or surfaces [137]. They are made by putting together polynomial segments that have degrees that are either less than or equal to the spline's total degree. These segments' joining locations are known as knots and the values of the spline there are typically provided by the user.

Splines are piecewise polynomials. Spline bases are more complex than finite Fourier series because they are more flexible. They are determined by the order, knots, and domain of validity. Splines come in a variety of forms. We solely take into account B-splines in this part.

Splines are built by segmenting the observation period into smaller segments, with borders at locations known as break points or just breaks. The spline function is a polynomial of constant degree or order across any subinterval, but the nature of the polynomial changes as one enters the next subinterval. The highest power in the polynomial is referred to as the degree. A polynomial has an order that is one higher than its degree. For instance, a polynomial of degree one, with one as its greatest power, defines a straight line; nonetheless, this polynomial is of order two since it also contains a constant component.

In reality, a spline foundation is described in terms of a collection of knots. These have a connection to break points in that each knot has the same value as a break point, though there might be more than one knot at a given break point.

The spline function is characterized by the arrangement of the knot sequence τ and the polynomial segment order. In the conventional scenario where there is one knot per breakpoint, the spline function is uniquely determined by the sum of the polynomial segment order and the count of internal knots, denoted as m + L-1. These parameters collectively establish the spline function.

We define a set of basis functions $\emptyset_k(t)$, which must possess the following characteristics:

- Every fundamental basis function, represented as $\emptyset_k(t)$, is inherently a spline function. This is defined by its order, denoted as m, and a sequence of knots τ.

- A spline function can be represented as the result of combining basis functions linearly. This is supported by the fact that multiplying a spline function by a constant retains its spline nature, and the addition and subtraction of spline functions also yield splines.

- One can express any spline function characterized by parameters m and τ as a linear combination of the aforementioned basis functions

The relation determines the number K of basis functions in a spline basis system

number of basis functions = order + number of interior knots:

Figure 2-7 displays the 13 order four B-splines created in R and correspond to nine inner knots evenly separated from one another across the interval [0;10].
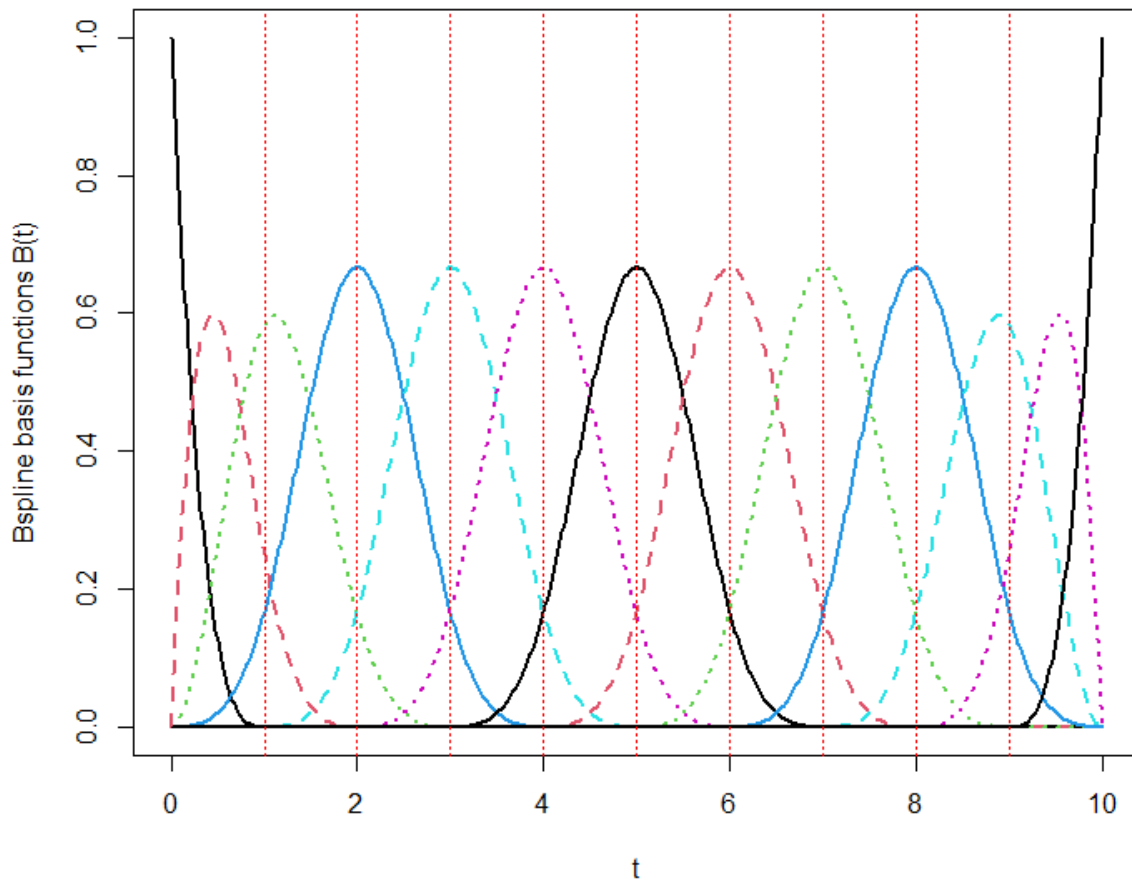
Figure 2-4 The 13 order four B-splines created in R and correspond to nine inner knots evenly separated from one another across the interval [0;10].

Fig The thirteen spline basis functions are defined across the interval [0,10] by a total of nine knots that are considered to be interior boundaries. Cubic polynomials, also known as order four polynomials, make up the polynomial segments, and the polynomial values and their first two derivatives are expected to correspond at each knot.

2.7.4 Wavelets

Often compared to the Fourier series [138], a wavelet is defined as a wave with a short duration and an average value of zero. The primary distinction between the two is that although Fourier series are only locally localized in time, wavelets are also locally localized in frequency. Furthermore, the discrete wavelet transform (DWT) can quickly compute coefficients and estimators. The DWT is even faster than the FFT (taking $\mathcal{O}(n)$ computational time compared to $\mathcal{O}(n\log_2 n)$ when the time values t_j are evenly spaced at $n = 2^M$ points).

By choosing a suitable mother wavelet function ψ and then examining all dilations and translations of the form, we may establish a basis for all square-integrable functions on $(-\infty, \infty)$.

$$\psi_{jk}(t) = 2^{\frac{j}{2}}\psi(2^j t - k)$$

Where j and k are integers.

Chapter 3

Circular functional analysis of OCT data for precise identification of structural phenotypes in the eye

3.1   Abstract:

Progressive optic neuropathies such as glaucoma are major causes of blindness globally. Multiple sources of subjectivity and analytical challenges are often encountered by clinicians in the process of early diagnosis and clinical management of these diseases. In glaucoma, the structural damage is often characterized by neuroretinal rim (NRR) thinning of the optic nerve head, and other clinical parameters. Baseline structural heterogeneity in the eyes can play a key role in the progression of optic neuropathies, and present challenges to clinical decision-making. We generated a dataset of OCT based high-resolution circular measurements on NRR phenotypes, along with other clinical covariates, of 3,973 healthy eyes as part of an established clinical cohort of Asian Indian participants. We introduced CIFU, a new computational pipeline for CIrcular FUnctional data modeling and analysis. We demonstrated CIFU by unsupervised circular functional clustering of the OCT NRR data, followed by meta-clustering to characterize the clusters using clinical covariates, and presented a circular visualization of the results. Upon stratification by age, we identified a healthy NRR phenotype cluster in the age group1 40-49 years with predictive potential for glaucoma. Our dataset also addresses the disparity of representation of this particular population in normative OCT databases.

3.2 Introduction

Progressive optic neuropathies such as glaucoma can cause irreversible blindness, especially when left untreated or diagnosed late. Indeed, early detection and management hold the key to slowing the progressive loss of vision and preventing blindness due to many chronic and age-related degenerative eye diseases. Glaucoma, for instance, is the second-leading cause of blindness worldwide[139]. In 2020, an estimated 80 million individuals worldwide had glaucoma and this number is expected to increase to over 111 million by 2040[3].

There are multiple sources of subjectivity and analytical challenges that are often encountered by the clinicians in the process of early diagnosis and clinical management of degenerative eye diseases. In glaucoma, the functional damage is established most commonly by the occurrence of visual field loss whereas the structural damage is often characterized by neuroretinal rim (NRR) thinning of the ONH, and loss of retinal nerve fibers, which are the axons of RGC. Such thinning could be measured in terms of reduction in either NRR area or NRR thickness. On the functional side, while standard automated perimetry (SAP) has been the gold standard for detection of VF loss, often 30% of RGC loss may have already occurred before VF defects could be detected by SAP[140].

On the structural side, biological heterogeneity of ONH phenotypes, with or without any neuropathy, can present challenges to clinical decision-making. For instance, the NRR area has been found to normally decline at the rate of 0.28%-0.39% per year[141]. There is no single, specific management guidance for patients with diverse morphology of ONH[142]. For instance, to assess the progression of glaucoma, one of the parameters assessed is the optic cup to optic disc ratio (CDR) which is calculated by comparing the diameter of the "cup" portion of the optic disc with the total diameter of the latter. Yet, while a large CDR may indicate glaucoma or other pathology, deep yet stable (over age) cupping, i.e., a normal physiologically large optic disc cup, can occur due to genetic factors in the absence of any disease or associated clinical covariates (e.g., high intraocular pressure)[143]. It is of great importance that sources of natural variation are rigorously understood thereby controlling for subjectivity in diagnosis.

In the clinic, non-invasive, high-resolution eye imaging platforms such as SD-OCT provide excellent glaucoma diagnostic performance, especially during early stages of the disease[117, 118]. The quantitative and reproducible OCT data provide objective measurements of ONH parameters, NRR area, retinal nerve fiber layer (RNFL), macular thickness, etc., which inform

clinicians about structural damage. For instance, Zeiss Cirrus HD-OCT platform uses the clinically invisible but OCT detectable BMOas the landmark to measure the amount of NRR tissue in the optic nerve. It has been reported that NRR thickness calculation by Cirrus HD-OCT has high reproducibility and glaucoma diagnostic ability, and a low rate of incorrect optic disc margin detection [127, 144, 145]. The platform generates a comparative report of a patient's data based on its normative database.

The OCT platform performs circular scans of the eye which, as in many biomedical technologies, are examples of measurements that are recorded or indexed at different directions, say, at given angular positions around a central point. Unlike the analyses of "linear" data points that reside on the real line or Euclidean spaces, directional data requires special and altogether different treatment. For instance, a direction in two-dimensional plane can be represented as a point on the circumference of a unit circle, or simply as an angle, but neither representation is unique, as both depend on the selection of some appropriate "zero-direction" from which to start measuring, as well as the sense of rotation, viz., clockwise or anti-clockwise. The unique properties of circular data – for instance, if one wishes to compare two such scans with a distance measure – are appropriately addressed by the field of circular statistics[146].

Traditional OCT analysis may involve division of the circle around ONH into 4 fixed quadrants, or 12 clock-hours, to record measurements at these sectors. In this study, we extended such data collection to divide the same circle (of total 360 degrees) into much finer segments of 2 degrees each. Thus, we generated 180 circular data points measuring NRR thickness for each clinical sample (human eye). These rich and evenly spaced high-resolution circular data allow for natural application of FDA where the data are not viewed as points but as curves or mathematical functions. Not to be confused with alternate usage of the term "function" (such as in vision), FDA is increasingly popular in biomedical informatics due to the emergence of new monitoring technologies that can record data as curves[147-149]. The high-resolution OCT data for each sample can be modeled as a circular function or curve, with angle and magnitude being the independent and dependent variables, respectively. The circular nature of the OCT NRR data is captured rigorously by the use of Fourier basis functions for their representation as functional or curve data[150]. In this study, we describe a novel method for functional clustering of OCT NRR curves, and apply it for unsupervised identification of

NRR phenotypes in healthy Asian Indian eyes. We also address the disparity of representation of Asian Indian eye phenotypes in normative OCT databases.

Progressive and degenerative eye diseases benefit from pre-existing knowledge and cumulative collection and description of normal phenotypes such as these, may help to identify early and characterize precisely the new phenotypes that emerge over time. While some normative OCT databases do exist, they are generally limited in their size and diversity. Thus, the breakup of their ethnic representation may not reflect the actual epidemiologic distribution of the disease. For instance, only 1% of the popular Cirrus HD-OCT platform's normative database is of Asian Indian origin[120], although India contributes to more than 12% of the global cases of both primary open-angle and primary angle-closure glaucomas[151]. Towards this, we generated a new and large high-resolution OCT dataset on NRR phenotypes, along with other clinical covariates, of 3973 healthy eyes as part of a well-established clinical cohort (LVPEI-GLEAMS) at the L.V. Prasad Eye Institute, Hyderabad, India.

The main objectives of the present study are to: (1) generate OCT NRR data in the form of 180 circular measurements of NRR thickness in a given eye, (2) introduce CIFU, a computational pipeline for CIrcular FUnctional data modeling and analysis that is demonstrated using the OCT NRR dataset, and (3) address the disparity of representation of the Asian Indian population in normative OCT databases. In the next section, we describe the clinical cohort and the protocol for data generation as well as the algorithm of CIFU for unsupervised circular functional clustering of the NRR thickness data, followed by meta-clustering to characterize the clustering output using clinical covariates of glaucoma. In the following section, the results of CIFU analysis are described with help of circular visualization. In particular, upon stratification of the samples by age, we identified a healthy NRR phenotype cluster in the age group1 40-49 years, and having the highest mean values of cup volume and average CDR among all clusters, with predictive potential for glaucoma. We end with discussion of the CIFU approach and its potential applications to future work.

3.3   Data and Methods
3.3.1 Data

All participants were selected from a population-based study conducted by the L.V. Prasad Eye Institute (LVPEI), Hyderabad, India. It is denoted by LVPEI Glaucoma Epidemiology And Molecular Genetic Study [152] The LVPEI-GLEAMS data protocol was approved by the LVPEI Institutional Ethics Committee (LEC 08131). The data on a total of 3973 healthy eyes of which 1981 right eyes (OD) and 1992 left eyes (OS) were collected from 2222 participants from the southern Indian state of Andhra Pradesh, India. Written informed consent was obtained from all participants to participate in the study, and the ethics and review committee of the LVPEI reviewed and approved the methodology and was conducted in strict adherence to the tenets of the Declaration of Helsinki. The inclusion criteria used were age >= 40 years, male or female, best-corrected visual acuity of 20/40 or better, SE of ±6 diopters, good quality stereo optic disc photographs, and no media opacities. The exclusion criteria used were intraocular surgery within the previous 6 months, and any retinal (including macular) or neurologic diseases other that could confound the structural measurements with SD-OCT.

The performance of the OCT layer segmentation algorithms can be affected by poor image quality, leading to erroneous demarcation of the retinal layers and inaccurate measurements. We excluded any OCT scan image sample from our study with poor image quality, signal strength <6, motion artifacts, blinking artifacts, misidentification of inner and outer retinal layers, and off-center artifacts. Several independent studies, including by some of the authors of the present study, have shown that signal strength reduction is associated with decreased accuracy of nerve fiber layer thickness measurement by OCT, which may be erroneously interpreted as presence of glaucomatous damage on a cross-sectional evaluation or when multiple scans are compared [153-155]. For the Cirrus SD-OCT platform used in the present study, we followed the manufacturer's definition of scans of adequate quality to be those of signal strength 6 or above (within a range from 0 to 10)[120].

Healthy eyes were defined by the absence of anterior and posterior pathology. Each digital optic disc photograph was evaluated by three glaucoma specialists independently. The specialists were masked to the other clinical findings and the other imaging outcomes of the subjects. Eyes were excluded from the study in case of any disagreements among the specialists. All participants underwent a comprehensive ophthalmic examination which included detailed medical and systemic history. The means of clinical determination included best-corrected visual acuity measurement, slit-lamp photographs (Topcon, Bauer Drive,

Oakland, NJ), Goldmann applanation tonometry (Hagg-Streit AT 900, Hagg-Streit AG, Switzerland), gonioscopy with a Sussman four mirror gonioscope (Volk Optical Inc, Mentor, Ohio, USA), dilated fundus examination, CCT assessment, Humphrey visual fields (HVF) with 24-2 Swedish Interactive threshold algorithm (Carl Zeiss Meditec Inc. Dublin, CA). Visual fields were considered if false positive, false negative, and fixation losses were less than 20%, and all the stereophotographs of the optic disc had good quality.

In addition, digital optic disc photography and SD-OCT imaging with Cirrus HD-OCT (software version 9.0.0.281; Carl Zeiss Meditec, Dublin, CA, USA) were used. This is a computerized instrument that acquires and analyzes cross-sectional and three-dimensional tomograms of the eye using SD-OCT technology. The instrument's algorithm automatically identifies the optic disc margin as the termination of Bruch's membrane (BM). BM opening (BMO) is used as the landmark to measure the amount of NRR tissue in the optic nerve. Optic Disc Cube 200x200 protocol was used to scan the ONH and peripapillary area through a 6 mm square grid, which consists of 200 horizontal linear B-scans and each composed of 200 A-scans. First, the Cirrus HD-OCT algorithm identifies the center of ONH and then automatically places a calculation circle of 3.46 mm diameter evenly around it. The circular scan starts at an extreme temporal point and moves around the ring in the superior direction, then nasal, then inferior, then back to temporal (TSNIT). The circular measurements are made clockwise for the right eye and counter-clockwise for the left eye. NRR thickness is measured by the amount of neuro-retinal tissue in the optic nerve around the entire edge of the optic disc. Zeiss Cirrus HD-OCT used Bruch's membrane opening–minimum rim width (BMO-MRW) to measure the rim area. The BMO-MRW is the shortest distance from BMO to the retinal internal limiting membrane. The advanced export functionality was used to record the NRR thickness values at 180 points in TSNIT order spaced evenly by 2 degrees (from 2°–360°) around the circle. We refer to this as our NRR OCT high-resolution circular data. The data were stratified into 3 age groups: (1) 40-49 years, (2) 50-59 years, and (3) 60+ years.

3.3.2 Methods

We describe CIFU pipeline for circular functional modeling and clustering of OCT NRR data, followed by metaclustering-based clinical characterization of the clusters identified by CIFU. The steps of the CIFU pipeline are graphically illustrated in Figure 3-1.
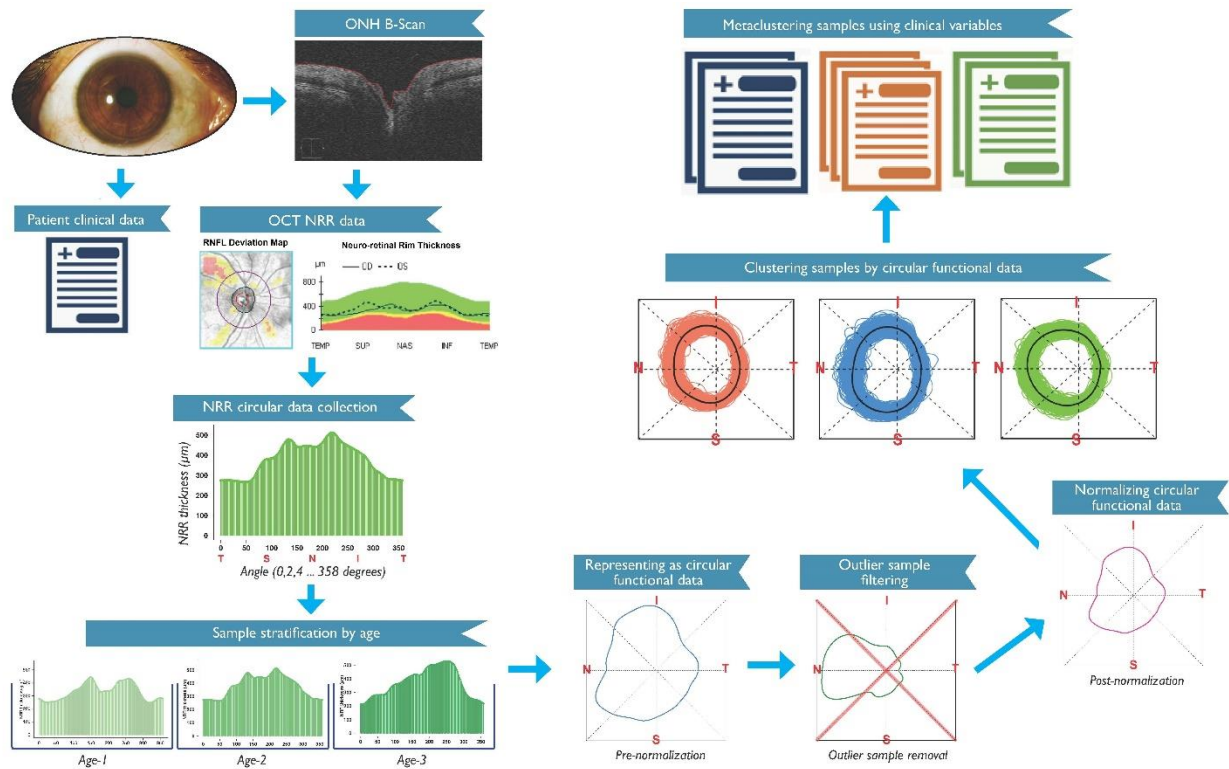
Figure 3-1 CIFU pipeline gives an illustration of each step including the collection of OCT NRR data and clinical data, representation as and clustering of NRR functional data, and metaclustering of clusters using clinical variables.

3.3.2.1 Circular Functional Modeling and Clustering

First, we introduce a method for clustering OCT NRR data into K homogeneous groups of samples, i.e., eyes. As an unsupervised approach, the clustering is based only on NRR data as input, and not any clinical variables of the samples. While the actual OCT measurements are taken on a discrete grid of angles around the center of ONH, in principle, these finely indexed measurements are assumed to vary continuously around a circular scale ranging from 0 to 360 degrees, and wrapped around. Thus, for statistical modeling, we will refer to the OCT data as a collection of n circular curves $X_1(t), X_2(t) \dots, X_n(t)$, where $X_i(t)$ represents the NRR thickness in the ith sample $(i = 1, 2, \dots n)$ measured at angle $t \in [0,360]$ and around a common center. The angular indices t_l, where $l = 1, \dots 180$, at which the curves are actually measured are aligned for all samples, and spaced 2 degrees apart. Specifically, the l-th measurement angle is $t_l = 2(l - 1)$.

Our modeling begins with the use of p basis functions for capturing the functional nature of data. If ψ_1, ψ_2, . . ., ψ_p are the basis functions with the associated basis expansion coefficients γ_{ij}, where $i = 1, 2, \ldots n$ and $j = 1, 2, \ldots p$, then the functional approximation for the ith curve at point t, $X_i(t)$, is given by

$$X_i(t) \approx \sum_{j=1}^{p} \gamma_{ij}\psi_j(t) = X_i^p(t). \qquad (1)$$

Different sources of artifacts could be present in OCT data[156]. As mentioned above, we exclude samples with such artifacts, yet OCT NRR curves with highly unusual shapes might appear rarely. While precise characterization of an outlier curve may be difficult[24], we used an intuitive criterion for detecting outlier NRR curves using equation (2) below, in which the factor 3.5 is based on our exploratory analysis. A given curve X_i is considered an outlier if there exists a point t_l for which $X_i(t_l)$ exceeds the following threshold

$$X_{out} = \max_l \bar{X}_i(t_l) + 3.5 \times \max_l sd(X_i(t_l)). \qquad (2)$$

After the outlier curves were removed, to allow for comparison of the curves based on their shapes rather than the magnitude, we normalize each curve $X_i(t)$ by dividing it by $\int_0^{360} X_i(t)dt$, where the integral is approximated numerically. Note, the normalization step is optional in the CIFU pipeline.

As the NRR thickness values are measured radially around a circle (Figure 3-2), they are nonnegative and have the natural periodicity of 360 degrees, so that the value corresponding to 5 degrees is the same as that at 365 degrees. After performing the normalization described above, they also contain a unit area on [0, 360), and thus possess the properties of a probability density around the circle. Such "circular densities" can be expressed in terms of an infinite series of Fourier coefficients[12] and then approximated by the first p terms, for a sufficiently large p, of the expression in equation (1). The idea is similar to approximating a continuous function by a polynomial of high enough degree, except that the Fourier basis also retains the periodicity that is critical to interpretation of the normalized OCT NRR curves.
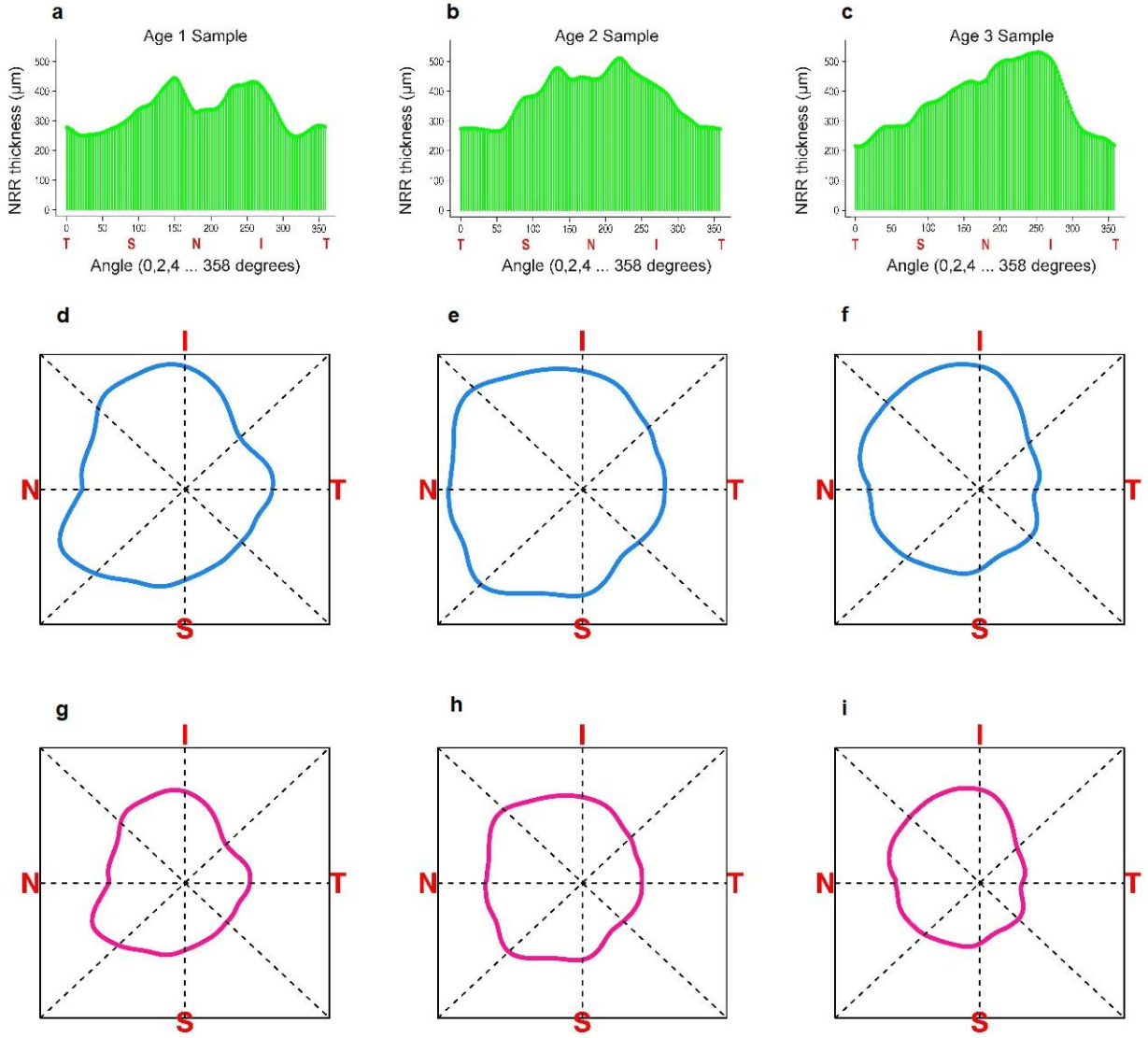
Figure 3-2 For 3 real samples, selected from each age group (40-49 years, 50-59 years, 60+ from left to right), the NRR thickness data measured for 180 evenly-spaced points around the circle is shown in the top panel. The corresponding circular functional approximations are shown as pre - and post-normalized NRR curves in the middle and the bottom panels respectively. The direction of the NRR curves is given by TSNIT (clockwise)

Our objective is to cluster each of the curves described above into a pre-specified number (K) of clusters. Then, we used a discriminative functional mixture (DFM) model given by Bouveyron et al. (2015)[157] in which $\gamma_i = \left(\gamma_{i1}, \dots, \gamma_{ip}\right)^t$ of curve X_i follows a finite mixture model of K Gaussian components with density function

$$f(\gamma) = \sum_{k=1}^{K} \pi_k \emptyset\big(\gamma; U\mu_k, U^t \Sigma_k U + \varXi\big), \qquad (3)$$

where $\pi_k \geq 0$ is the mixing proportion of the kth component (i.e., the cluster k) such that $\sum_{k=1}^{K} \pi_k = 1$, and $\emptyset$ is the standard Gaussian density function. Here, U is a $p \times d$ orthogonal matrix mapping the basis coefficients γ into the discriminative subspace (of dimension $d < p$) through a linear transformation. Similarly, μ_k and Σ_k are the mean vector and covariance matrix (for cluster k) of γ mapped into the discriminative subspace, and the noise of the above transformation is normally distributed with mean zero and covariance Ξ.

The DFM model was fit with an Expectation-Maximization (EM) algorithm implemented in the R package funFEM[158]. EM is a popular iterative method used for optimal fitting of a model and estimation of the model parameters[159]. Following the idea of constraining variance parameters in Fraley and Raftery (1999) [160], the funFEM package allows 12 different choices of DFM models. As an initial step, using different restrictions on the noise covariance matrix Ξ, a preliminary model search was run with NRR data, with outlier curves included, over all 12 models and a flexible set of (61) basis functions. Based on the results of the initial run, and after removing the outlier curves as described below, the clustering algorithm was run using a smaller selective set of (21) basis functions.

To avoid model overfitting, we determined the smallest number of basis functions (p) that recover the input curves sufficiently well, as determined by the fraction of variation explained (FVE) as described below. Let the sample mean of n given curves be

$$\bar{X}(t_l) = \frac{1}{n}\sum_{i=1}^{n} X_i(t_l), \ \ l = 1, 2, \dots 180. \tag{4}$$

Then the total variation (TV) is given by

$$TV = \frac{1}{n-1}\sum_{i=1}^{n} \int \left(X_i(t) - \bar{X}(t)\right)^2 dt \tag{5}$$

and the fraction of variation explained (FVE) by

$$FVE = \frac{TV - \frac{1}{n-1}\sum_{i=1}^{n} \int \left(X_i^p(t) - X_i(t)\right)^2 dt}{TV}, \tag{6}$$

where the integrals are again approximated numerically using the discrete observations $X_i(t_l)$.

We used FVE as the criterion for selecting an optimal number of basis functions (p) by identifying the smallest value p for which FVE exceeds 0.99, i.e., 99% (Figure 3-3).
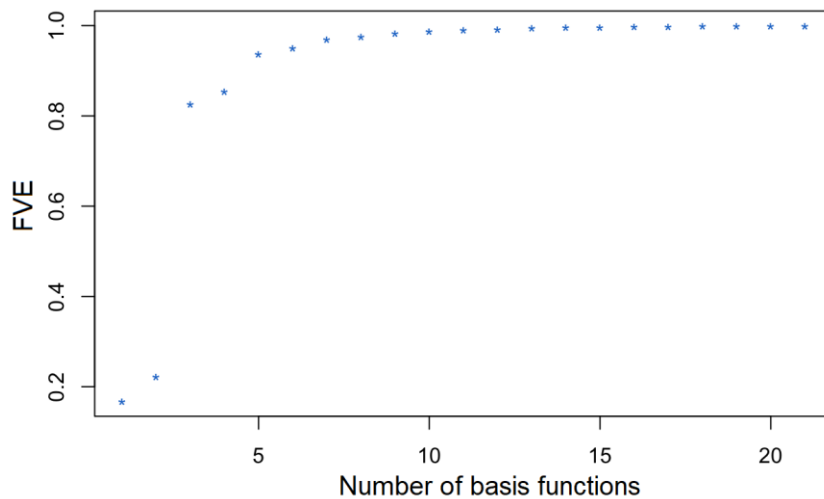
Figure 3-3 The fraction of variation explained (FVE) by models with different choices of the number (p) of basis functions used for the functional representation of OCT data.

Finally, the number of clusters identified by the DFM models for each age group was determined by 3 popular model selection criteria: AIC (Akaike Information Criterion) [161], BIC (Bayesian Information Criterion) [162], and ICL (Integrated Complete Likelihood) [163].

For intuitive visualization of the clustering results, we plotted the curves of every cluster in a distinct color using a circular scale. The mean curve of each cluster, as computed by (4), is included as a bold black curve, which serves as a cluster-specific template.

3.3.2.2 Comparative Clustering Analysis

We performed comparative analysis of our circular functional clustering method against three popular non-functional data clustering methods, namely, k-means[164], Partitioning Around Medoids (PAM [164]) which is similar to kmeans but uses higher dimensional medoids in place of means, and Gaussian mixture model (as implemented in Mclust [160]). Each of these methods were run with the same OCT NRR data points, i.e., not as NRR curves, of each age group. Each method was run with a possible choice of fitting $K = 2$ through 10 clusters for each age group. Thus, we ran the 3 alternative clustering methods for 3 age groups for 9 possible values of K. The optimal number of clusters identified by each method was determined with the Average Silhouette Width (ASW), while Dunn Index was calculated as a measure of inter-cluster variation. We used the R packages 'optCluster' [165] and 'factoextra'[166] for clustering, validation and visualization.

3.3.2.3 Metaclustering and Clinical Characterization of Clusters

In the metaclustering step, the clusters identified by circular functional data were grouped based on their samples' similarity in terms of a selected set of clinical variables that are known covariates of glaucoma. A feature selection step was performed simultaneously to detect the covariates that were the most distinctive across the metaclusters. The metaclustering workflow consists of the following steps:

1. In each age group, we performed agglomerative hierarchical clustering of the clusters given by their mean covariate data with complete linkage, while simultaneously doing feature selection to select a sparse set of covariates that are the most distinctive across the metaclusters.

2. We plotted the metaclusters (identified in Step 1) with age group-specific dendrograms. A flat cut of the dendrograms at a common height threshold was used to distinguish the metaclusters in each age group. The metaclusters that correspond across the age groups are shown as subtrees of matched colors.

3. We visualized using contour plots the corresponding metaclusters of each age group to compare the distributions of the selected covariates across the metaclusters as well as the age groups.

In step 1, a set of 9 covariates were used based on their clinical relevance. The R package 'sparcl' [167]was used for agglomerative and sparse hierarchical metaclustering in step 2; the feature selection in this package is done by varying the values of its 'wbound' parameter from 2 to 5.

3.4   Results

The CIFU pipeline was run with OCT NRR thickness data and clinical assessment data of a normal cohort consisting of 3973 healthy eyes. The steps of the pipeline began with stratification of the OCT and clinical data by age into 3 age groups with (1) 1841, (2) 1351, and (3) 781 samples respectively. The list of clinical variables is summarized in Table 3-1. An identical sequence of steps of analysis was followed by CIFU within each age group.

Clinical variables	40-49 years	50-59 years	60+
Number of eyes, N (OD/OS)	1841(917/924)	1351(677/674)	781(387/394)
Age (years)*	43.95±2.83	53.18±2.76	64.63±5.39
Gender, N (Female/Male)	1256/585	755/596	372/409
BCVA LogMAR*	0.01±0.04	0.03±0.08	0.09±0.11
SE (diopter)*	0.02±0.75	0.16±1.05	-0.17±1.3
IOP (mmHg)*	12.58±2.34	12.49±2.41	12.13±2.38
CCT (µm)*	525.7±32.22	524.28±31.96	517.14±31.62
Axial length (mm)*	22.59±0.74	22.63±0.71	22.6±0.8
Family history of Glaucoma, n, (no/yes)	1835/6	1345/6	781/0
Diabetes mellitus, n, (no/yes)	1720/121	1163/188	666/115
Hypertension, n, (no/yes)	1696/145	1083/268	575/206
RIM AREA (mm2)*	1.36±0.22	1.34±0.23	1.33±0.25
DISC AREA (mm2)*	1.96±0.35	1.97±0.35	2.01±0.37
AVERAGE CDR*	0.52±0.15	0.53±0.14	0.55±0.14
AVERAGE THICKNESS (µm)*	94.36±9.25	92.3±9.57	90.73±9.9
VERTICAL CDR*	0.49±0.15	0.5±0.13	0.52±0.13
CUP VOLUME (mm3)*	0.18±0.16	0.18±0.15	0.2±0.18
DISC DIAMETER (mm)*	1.49±0.15	1.51±0.15	1.53±0.16

Table 3-1: The clinical variables of the study participants in the three age groups

The units are given in parentheses. The asterisk (*) denotes that a variable is described as mean±sd. (N: Number of samples; OD: Oculus Dexter; OS: Oculus Sinister; BCVA LogMAR: Best Corrected Visual Acuity Logarithm of the Minimum Angle of Resolution; IOP: IntraOcular Pressure; CCT: Central Corneal Thickness; CDR: Cup-to-Disc Ratio; sd: standard deviation.)

The 180-point data for each sample (eye) were modeled using $p = 11$ Fourier basis functions. We chose $p = 11$ since it was the smallest value of p for which FVE, as given in equation (6), exceeded 99% (Figure 3-3). The curves were normalized and aligned to a common starting angle of 0 degrees to allow for comparison of their shapes around the center of ONH. Using equation (2), the outlier OCT curves were removed: 6 samples from age group 1, 1 from age group 2, and 5 from age group 3. Then, within each age group, the curves were clustered by a DFM model as described in equation (3). The optimal number of clusters (K) for each age group was determined by 3 different well-known criteria: AIC, BIC, and ICL (described in Methods). These criteria showed overall strong agreement attesting to optimal model selection as seen in Figure 3-4. Based on the value of K beyond which no significant gain was noted in

these criteria, we determined the number of clusters, for age group 40-49 years, 50-59 years, 60+ as 7, 8, and 6 respectively.

(c)
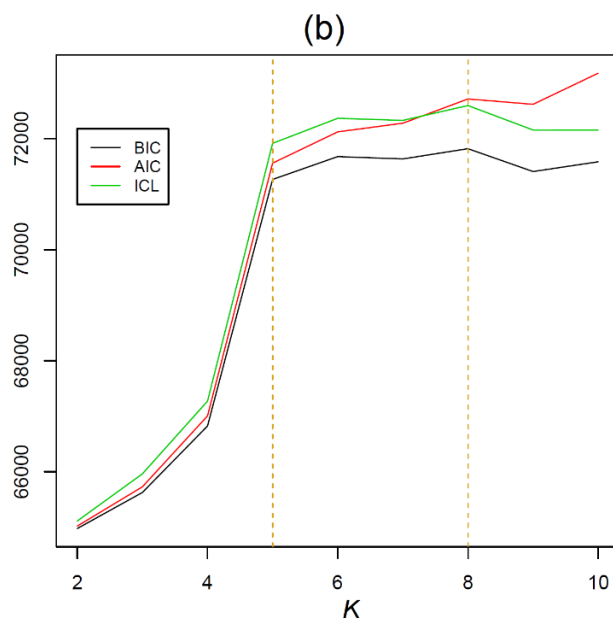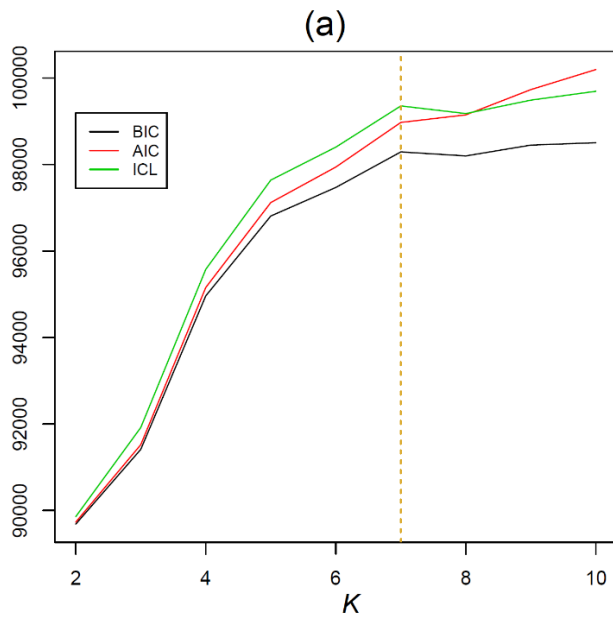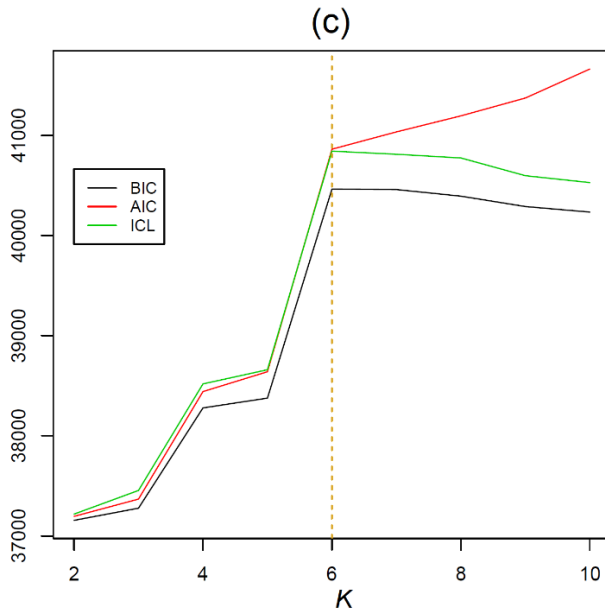
Figure 3-4 The values of model selection criteria AIC, BIC, and ICL corresponding to fitting of a DFM model of K clusters to OCT NRR samples of age group 40-49 years shown in (a), 50-59 years in (b) and 60+ in (c). The optimal DFM models for the age groups 40-49 years, 50-59 years and 60+ were selected for K=7, 8 and 6 respectively, beyond which no significant gain was noted.

The results of our circular functional clustering are shown in Figure 3-5 (a)-(c) as a panel of K clusters for each age group. Each cluster C within a panel consists of the circular curves for the samples that belong to C (all shown in a common color specific to C) based on the similarity of their functional representation. To gain an intuitive understanding of the 180-point OCT data on NRR phenotypes, we used a visualization of curves as represented on a common circular scale. Unsupervised clustering of the circular functions revealed various NRR patterns in the identified clusters, some of which were distinctive whereas others have subtle differences. Notably, the visualization reveals the unique mean shape (or NRR "template") of each cluster as shown by a bold black circular curve in each plot of Figure 3-5.
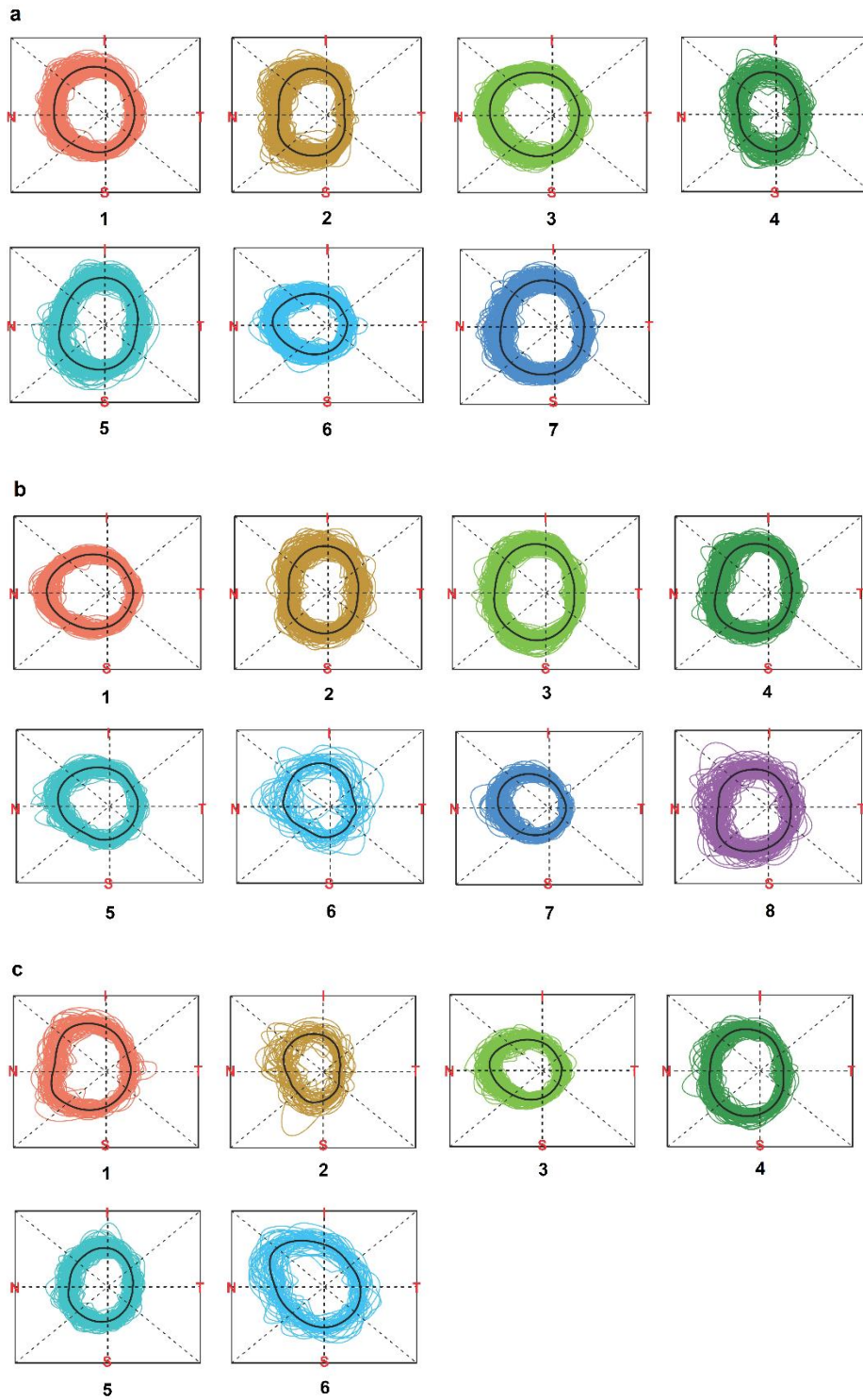
Figure 3-5 The clusters of the OCT NRR functional data are shown. For each of the age groups (a) 40-49 years, (b) 50-59 years, and (c) 60+, the normalized NRR curves that belong to the same cluster are shown together using a common color. In age groups 40-49 years, 50-59 years, and 60+, CIFU identified 7, 8, and 6 clusters respectively. For each cluster, its mean NRR curve is shown in black. The direction of the NRR curves is given by TSNIT (clockwise).

The circular curve visualization allows several interesting observations. We note the consistent dip at the temporal (T) region near 0 degrees, which is a characteristic feature shared by the templates of all clusters. This is supported by the well-known ISNT rule[87] according to which, from the center of ONH, the rim is the thinnest at the temporal (T) region. Interestingly, we also observed various shapes and features in the cluster templates (such as distinctive protrusions, notches, tilts, etc.) that appear as well as vary continuously in different (non-T) regions around the circle. The clustering solution allows us to record the intra-cluster variation which could be used to quantitatively compare the dynamics (say, the rates of focal change) of corresponding clusters across age groups. In this regard, we note that the popular methods of traditional clustering of the same OCT NRR data failed to capture the distinctive shapes and other spatial features of the NRR curves. These methods yielded only 2 clusters of NRR data points (not curves) each in every age group (Figure 3-6). Further, the traditional clusters (Figure 3-7) had low inter-cluster structural variation as measured by their small values of Dunn Index for age group 1(40-49 years) kmeans: 0.063, PAM: 0.055, Mclust: 0.043; 2(50-59 years) kmeans: 0.049, PAM: 0.044, Mclust: 0.034; and 3(60+) kmeans: 0.054, PAM: 0.056, Mclust: 0.042.
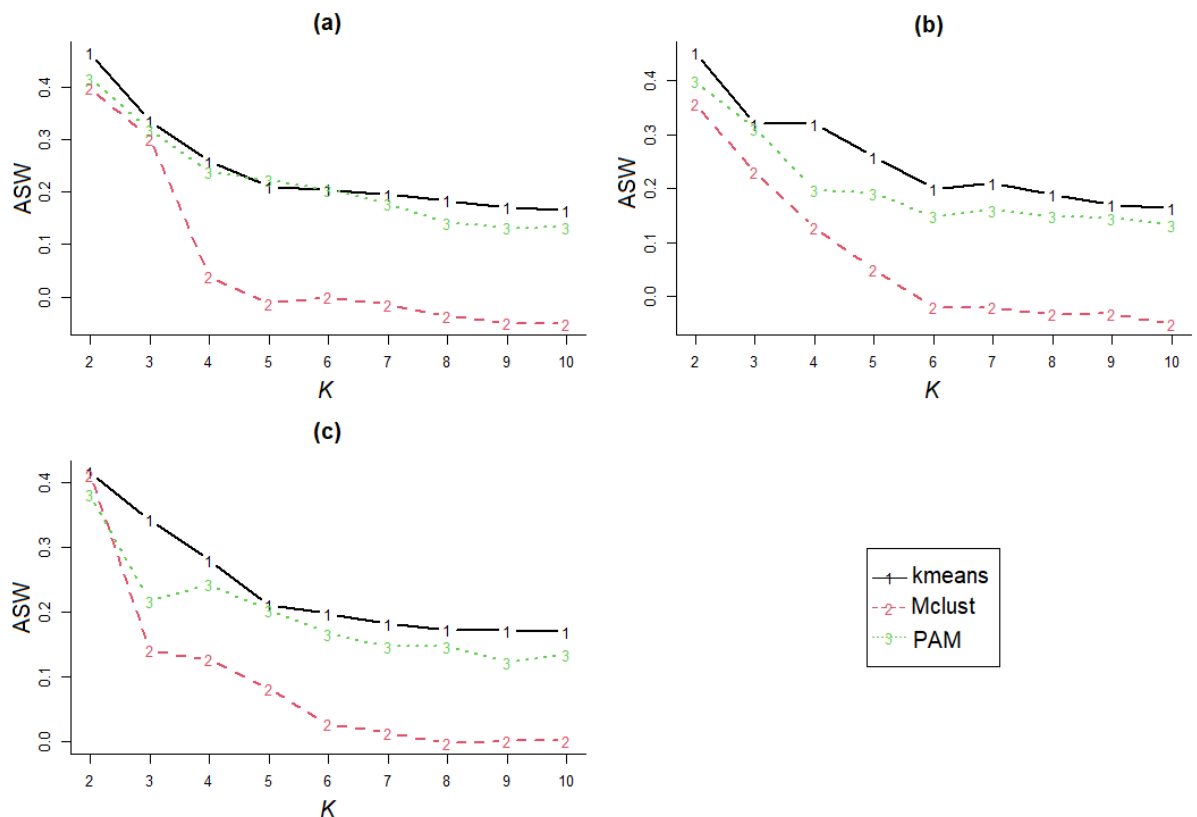
Figure 3-6 Non-functional clustering of the normalized OCT NRR data was conducted using 3 popular methods: k-means, PAM, and Gaussian mixture model by Mclust. For age groups (a) 40-49 years, (b) 50-59 years, and (c) 60+, the Average Silhouette Width (ASW in y-axis) for different choice of the number of clusters (K in x-axis) is shown. ASW is maximized for $K=2$, thus giving an optimal number of 2 clusters for all 3 clustering methods and for all 3 age groups.
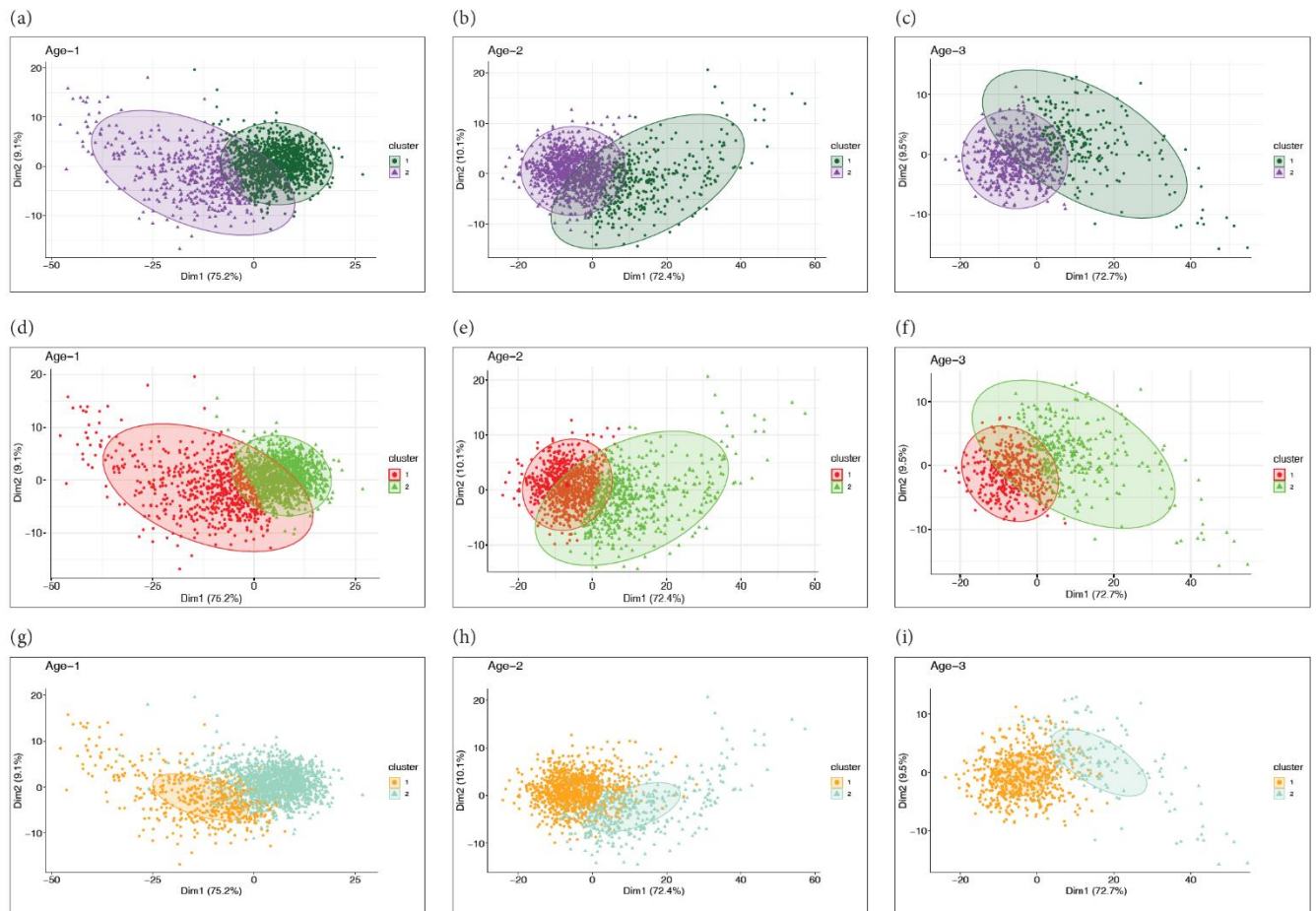
Figure 3-7 Non-functional traditional clustering of the normalized OCT NRR data was conducted using 3 popular methods: k-means (a)-(c), PAM (d)-(f), and Gaussian mixture model by Mclust (g)-(i). The age groups are noted on top of each plot. Based on the Average Silhouette Width for each method and each age, the optimal number of clusters is 2. The samples belonging to the 2 clusters are shown in each plot in different colors. Given the high-dimensionality of the data, we used as axes the first two principal components as to visualize the clustering results.

In order to establish a correspondence among the clusters in different age groups as well as to characterize the samples that belong to each cluster, we conducted a metaclustering analysis. In this step, we clustered the clusters based on a set of 9 clinical variables (Table 3-2) of the samples in each. These are known covariates of glaucoma, and no NRR data from the previous clustering step was used. The results of sparse hierarchical metaclustering are shown in Figure

3-8. The dendrograms reveal the similarities among the clusters in terms of their mean sample covariates as well as the counts of metaclusters identified at different levels of each dendrogram. Based on flat cuts of all the dendrograms (at a common height threshold of 0.1), we identified 3 metaclusters $\{M_1^1, M_2^1, M_3^1\}$ for age group 1; 2 metaclusters $\{M_1^2, M_2^2\}$ for age group 2; and 2 metaclusters $\{M_1^3, M_2^3\}$ for age group 3. Notably, all the dendrograms show the metacluster $M_2^\bullet$ (pink subtree) to be more heterogeneous in every age group than the metacluster $M_1^\bullet$ (blue subtree). Among the youngest participants, i.e., in age group 1, the metacluster M_3^1 (consisting of the original cluster 4) is distinct from the metacluster M_2^1.
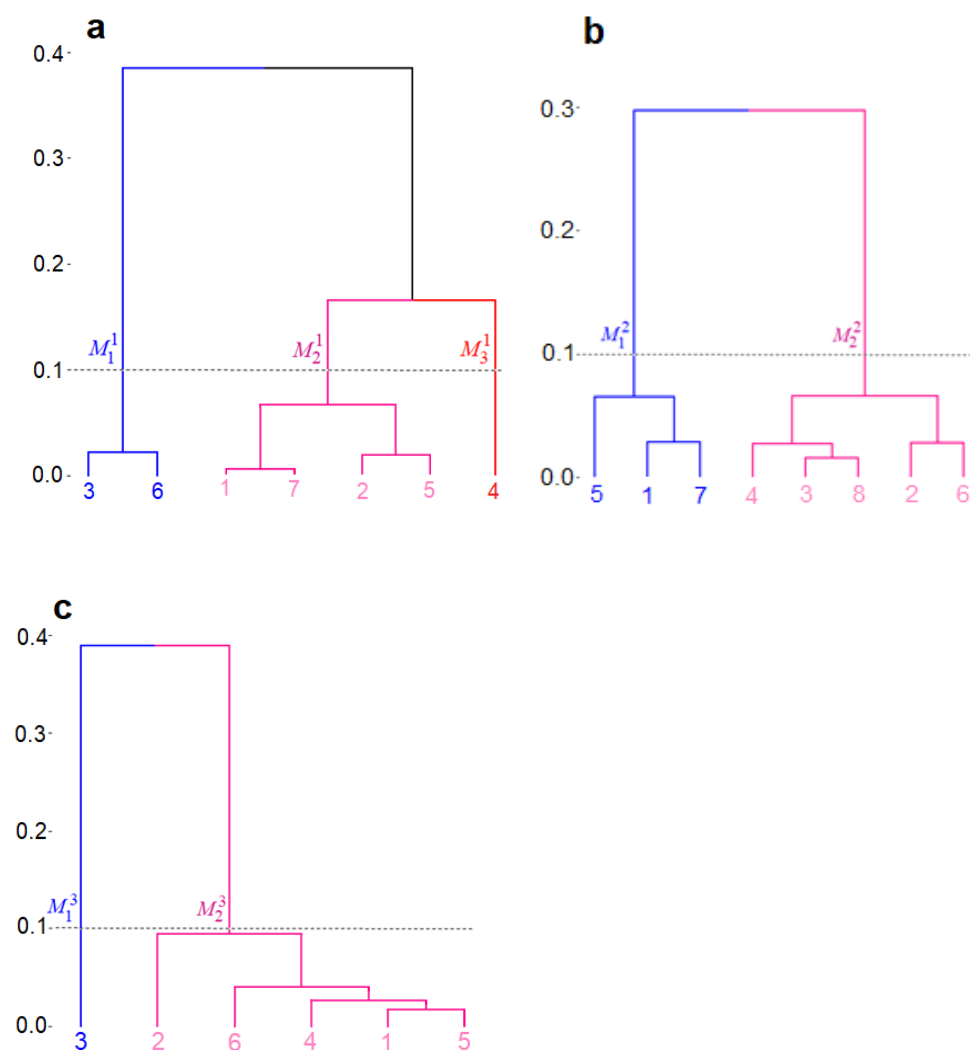
Figure 3-8 Metaclustering of the clusters was performed using the clinical covariates of the samples in each cluster. The results are shown using dendrograms for age groups (a) 40-49 years, (b) 50-59 years, and (c) 60+. The y-axis shows the distance between metaclusters. The leaves of a dendrogram denote the Id-s of the clusters identified by the previous clustering step. The metaclusters are obtained by a flat cut of each dendrogram at the common height of 0.1, and the labels and subtrees representing them are shown in different colors (blue, pink, and red).

A feature selection step, performed along with metaclustering, identified the optic disc cup volume and the average cup-to-disk ratio (CDR) of an eye as the most significant features in terms of the contributions of the different covariates to the metaclustering. These distinctive covariates allow us to register the correspondence of the metaclusters across the different age groups in Figure 3-9, which shows the contour plots of the metaclusters in their matched colors. The 3 metaclusters $M_1^{\bullet}$ shown in blue have the smallest mean values of cup volume and CDR, the 3 metaclusters $M_2^{\bullet}$ shown in pink have higher mean values of these covariates, and the single unmatched metacluster M_3^1 shown in red (Figure 3-9c) which has the highest mean values (Table 3-2).

	Age Group 40-49 years			Age Group 50-59 years		Age Group 60+	
Metacluster {Clusters}	M_1 {3,6}	M_2 {1,2,5,7}	M_3 {4}	M_1 {1,5,7}	M_2 {2,3,4,6,8}	M_1 {3}	M_2 {1,2,4,5,6}
Metacluster N (percent)	549 (29.8)	1170 (63.6)	122 (6.6)	547 (40.5)	804 (59.5)	212 (27.1)	569 (72.9)
IOP (mmHg)	12.68±2.31	12.54±2.33	12.54±2.61	12.34±2.31	12.6±2.48	12.12±2.48	12.13±2.34
CCT (µm)	525.93±31.08	526.24±32.86	519.53±30.68	523.06±31.53	525.09±32.23	518.71±33.58	516.56±30.87
Axial Length (mm)	22.55±0.75	22.59±0.74	22.7±0.67	22.63±0.69	22.63±0.72	22.8±0.72	22.53±0.82
RIM AREA (mm2)	1.44±0.23	1.32±0.21	1.25±0.19	1.42±0.24	1.29±0.21	1.42±0.25	1.3±0.24
DISC AREA (mm2)	1.86±0.29	1.99±0.35	2.17±0.4	1.89±0.32	2.02±0.36	1.88±0.32	2.06±0.38
DISC DIAMETER (mm)	1.47±0.13	1.5±0.15	1.59±0.17	1.48±0.14	1.52±0.15	1.48±0.14	1.55±0.16
VERTICAL CDR	0.45±0.13	0.49±0.15	0.56±0.14	0.47±0.14	0.52±0.13	0.48±0.12	0.53±0.14
AVERAGE CDR	**0.44±0.13**	**0.54±0.15**	**0.62±0.12**	**0.46±0.14**	**0.57±0.13**	**0.46±0.12**	**0.58±0.14**
CUP VOLUME (mm3)	**0.1±0.1**	**0.2±0.16**	**0.3±0.19**	**0.12±0.12**	**0.22±0.16**	**0.11±0.11**	**0.23±0.19**

Table 3-2: The clinical covariates used for metaclustering in the three age groups.

The two most significant covariates due to feature selection are shown in bold. The values of each variable in a metacluster are described as mean±sd. The units are given in parentheses. (*N*: the number of samples; IOP: IntraOcular Pressure; CCT: Central Corneal Thickness; CDR: Cup-to-Disc Ratio; sd: standard deviation.)
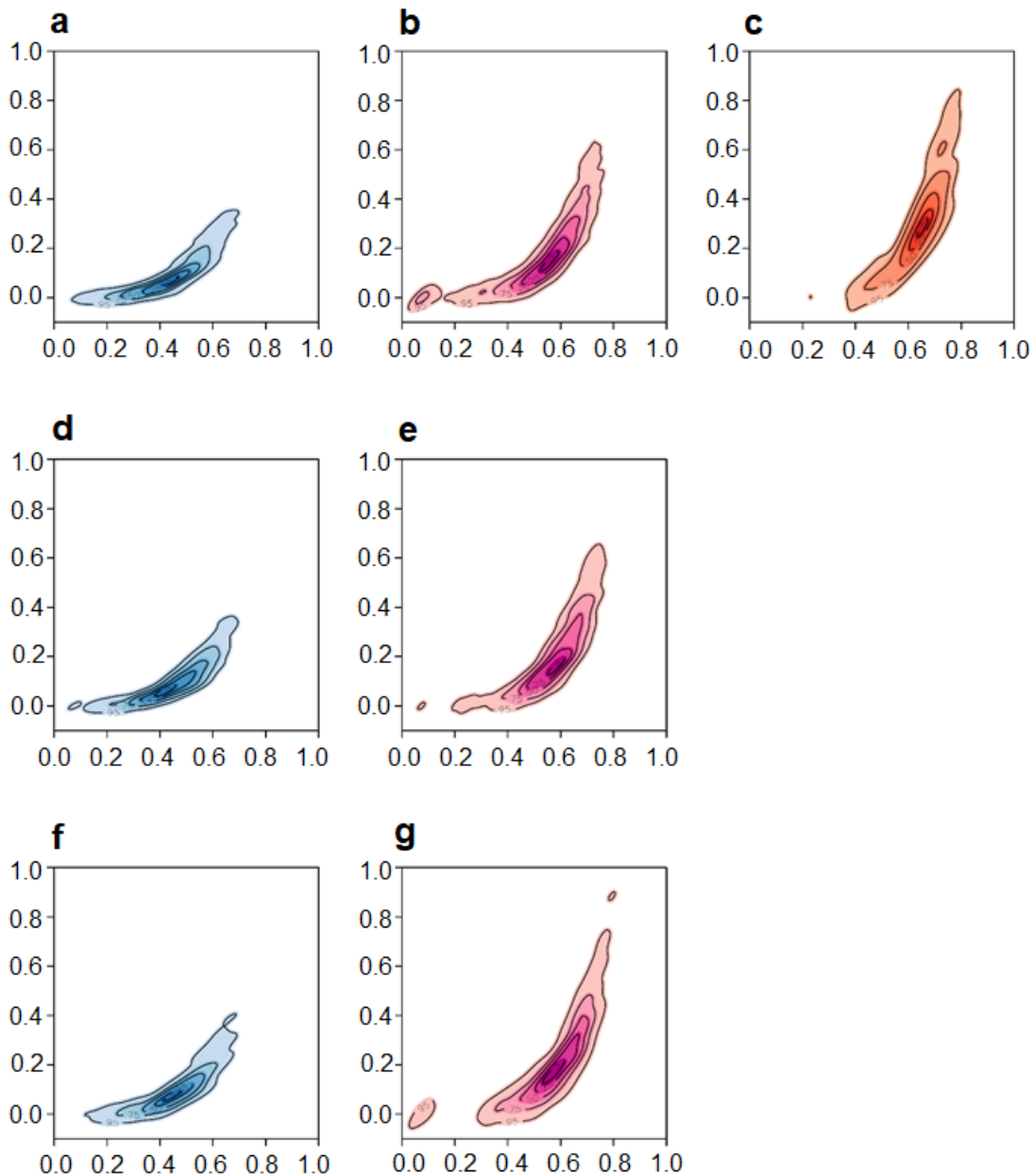
Figure 3-9 Contour plots of the distributions of clinical covariates optic cup volume (y-axis) and average CDR (x-axis) of the samples belonging to metaclusters of age groups 40-49 years: (a) M_1^1, (b) M_2^1, (c) M_3^1; 50-59 years: (d) M_1^2, (e) M_2^2; and 60+: (f) M_1^3, (g) M_2^3. The metaclusters that correspond across the 3 age groups are shown in matched colors (blue and pink) while the distinct metacluster (c) is shown in red.

It is interesting to consider the unmatched metacluster M_3^1 which not only has the highest mean values of the covariates (cup volume and average CDR) but is, in fact, comprised of a single, distinct cluster based on the OCT NRR phenotype data (cluster 4 in Figure 3-5a). Here we note that notwithstanding a large value of CDR (especially > 0.5), cupping by itself is not indicative of glaucoma. In fact, it is known that deep but stable cupping can occur due to hereditary reasons without glaucoma (see Discussion). Rather, it is a change in these ONH parameters with age of the participants that is a clinical indicator of glaucoma. Since the samples included

in the present study contain only healthy eyes as determined clinically by agreement of multiple glaucoma specialists, the presence of this unmatched cluster only in the youngest age group serves as a signature of healthy NRR phenotype with predictive potential for glaucoma. That is, the corresponding metacluster with such high values of these covariates among the older age groups would have the likelihood of progressing to glaucoma, and thus, is unlikely to be represented in a cohort of only healthy eyes, as we have in the present study.

3.5   Discussion

Unsupervised learning of the heterogeneity of normative ONH phenotypes in a given population can provide a more comprehensive understanding of the diversity of baselines that exist for degenerative neuropathies. Such knowledge is particularly useful in glaucomas for which different ONH parameters play a combined role in early detection. For example, in a non-glaucoma multiethnic cohort of Asian individuals, the inter-eye RNFL profile was found by OCT to be less symmetric in Malays and Indians than that in Chinese eyes [168]. Not only are the structural characteristics of individual eyes known to vary racially, even their rates of change over time could be different across population groups. For instance, the rate of change of BMO-MRW was recorded as −1.82 μm/year and −2.20 μm/year in glaucoma suspect eyes of European and African descents respectively[169]. In another multi-centered normal population study, both age-related decline and between-subject variability in BMO-MRW were observed[123]. Indeed, even the manufacturers of OCT technology noted racial differences in optic disc area, CDR, cup volume, and RNFL thickness when measured using their platform[170].

The presence of phenotypic heterogeneity makes it less justified to apply common, universal thresholds for clinical determination of glaucomatous damage in different population groups using OCT measurements, particularly in the early stages of the disease when the baselines could have stronger initial effects. To account for the effects of normal variation in ONH parameters, large and racially representative normative databases of healthy eye OCT phenotypes should be created. However, often such collections of healthy samples tend to be small or moderately sized, e.g., the normative database of Cirrus HD-OCT platform included just 284 subjects [120]. In that cohort, Caucasians represented 43%, Chinese 24%, African Americans 18%, Hispanics 12%, and others 6%. The representation of Asian Indians, in

contrast, was about 1% of the Cirrus HD-OCT cohort, which does not adequately reflect the 2020 projections about India to become the second in global glaucoma numbers, surpassing Europe [7]. Thus, more than 16 million Indians could be affected by glaucomas, and nearly 1.2 million could be blinded from the disease. Some resources such as the HRT3 Normative Database, while including 104 Indian individuals, did not improve the diagnostic sensitivity or specificity for glaucoma in that group for the potential reasons of limited sample size and intra-racial variation of ocular topography [171].

In this study, we leveraged the large population-based LVPEI-GLEAMS study to generate new and a relatively large OCT dataset based on nearly 4000 samples from normal Asian Indian participants. In fact, given that the recruitment of all the study participants was from a single geographic region (namely, the state of Andhra Pradesh), the scope of intra-racial variation to affect this dataset is limited. Moreover, the relatively large sample size of the data allowed us to identify a variety of clusters of NRR phenotypes, including the signature (cluster 4) with predictive potential for glaucoma consisting of 6.6% of all samples in the youngest age group (40-49 years). The absence of its corresponding cluster in the older healthy age groups despite their considerable sample sizes (total of 2132 samples of age 50 years and above) leads to a reasonable supposition of its potential pathological progression with increase in age, thereby resulting in lack of subsequent representation in a healthy cohort such as in the present study. Such phenotypic decline is consistent with the findings from a prospective longitudinal study that found the rate of age-related, glaucomatous global (or global percentage) rim area loss to be 3.7 (5.4) times faster as compared to healthy eyes [172].

Importantly, independent support for the identified signature relies on its clinical characterization in terms of covariates such as cup volume and CDR, which are useful parameters for diagnosis of glaucoma suspect[85]. Despite its normal mean value of intraocular pressure (IOP) as is expected of healthy eyes, the signature cluster has the highest mean values of average CDR and cup volume of all metaclusters across all 3 age groups (Table 3-2). In a recently published longitudinal study that started from baseline values and was run over a 5-year period, the covariates which had statistically significant increase in glaucomatous progression included CDR and cup volume [173]. We understand that it would perhaps be ideal to follow-up healthy individuals and measure the changes in their clinical covariates as they age in order to classify the ONH phenotypes via supervised learning. The approach of CIFU,

in comparison, involves unsupervised learning of different high-resolution phenotypes in age-stratified data and their characterization using key covariates, which is far less time-consuming and yet has the potential to produce a clinically insightful database for diverse populations with high phenotypic heterogeneity.

In addition to its sample size and racial representation, perhaps the most noteworthy feature of the present OCT dataset is its unique high-resolution measurements of NRR thickness around a circle. These samples, collected at every 2 degrees, extend the typical use of such measurements recorded at either 4 quadrants or 12 clock-hours, and even 48 angular positions[174], to higher-dimensional analysis. As clustering with curves show, the focal variations could be more nuanced than that suggested by a general rule, e.g., ISNT, and a capability to "zoom" into finer angular divisions can reveal further patterns[175]. In low-resolution data, it may be difficult to detect focal changes within the confines of pre-determined inflexible sectors. Moreover, the templates of the clusters could also be compared using known tests in shape analysis [176]. Be it circular data from OCT, or optic phenotypes in general, they seem suitable as candidate applications of circular statistics, and yet, we are unaware of any major previous studies in this regard. Further, the high-resolution also allowed the data to be closely approximated by continuous curves, and thus, specified by the corresponding functional representation. While clustering of circular point data [146, 177, 178]as well as clustering of curves[179-181] and (non-circular) functional data [147, 148, 157, 182-184]have been addressed by past studies, the present clustering of curves in the form of circular functions is possibly a novel application.

We understand that the present study has certain limitations. As we noted above, a prospective cohort study would be better suited to validate the predictive glaucomatous potential of the identified NRR phenotypic signature. We plan to address this in our future work. While high-resolution data could be accessed from the OCT scans, it is not commonly done by the clinical protocols. We hope that by adding user-friendly interfaces to CIFU, we may promote such data acquisition and analysis, especially since the functional representation is independent of the number of observations per sample. Indeed, there are several distinct advantages of our approach which could be built upon further in future studies. The estimated parameters of the fitted functional mixture model could be used to test the similarity of ONH phenotypes in normal versus disease conditions, thus allowing us to characterize any changes with precision

and rigor.

After a database of phenotypic parameters is developed, known measures of shapes and distances between curves could be used for objective clinical classification of new OCT samples. Applied to longitudinal analyses, our high-resolution modeling could identify intermediate, or previously uncharacterized, stages of disease progression, especially by focusing on variations within fine angular sections. Focused analysis of angular sections of OCT NRR and RNFL data have revealed interesting differences between healthy and glaucoma subjects, and we plan to apply CIFU for mining locally distinctive features in higher resolution [185]. Indeed, straightforward extensions are feasible for similar circular data such as RNFL phenotypes and other optic neuropathies as well as related eye imaging platforms, e.g., OCT-Angiography (OCTA). As we have demonstrated for other biomedical platforms [181, 186-189], the new pipeline CIFU could be enhanced incrementally with different functionalities, say, to increase computational efficiency or capture the perspective of the clinical experts. The circular curve visualization introduced in the present study may lead to a more user-friendly tool for clinical purposes as we plan to make it interactive, with advanced capabilities to jointly handle data and metadata, in our future work.

Chapter 4

Focused Analysis of RNFL Decay in Glaucomatous Eyes using Circular Statistics on High-resolution OCT data

4.1 Abstract

Using the OCT platform, we can generate high-resolution data on RNFLthickness in a given eye at a large number of angular-points defined on a circular coordinate system. However, insightful characteristics of OCT RNFL data such as its circularity and granularity generally remain unutilized in its common downstream statistical analyses. We present a new circular statistical framework that defines an Angular Decay function and thereby provides a directional representation of an eye in terms of its focused RNFL loss. By applying to a clinical cohort of Asian Indian eyes, the generated circular data was modeled with a finite mixture of von Mises distributions, which led to unsupervised identification in different age-groups of 2 recurrent clusters of glaucomatous eyes with distinct directional signatures of RNFL decay. New indices of global and local RNFL loss were computed for comparing the structural differences between these glaucoma clusters across the age-groups.

4.2 Introduction

Progressive optic neuropathies such as glaucoma can cause irreversible blindness, especially when left untreated or diagnosed late. Glaucoma is a chronic ophthalmic disorder characterized by irreversible damage of ganglion cell and retinal nerve fiber layer (RNFL), progressive neuroretinal rim (NRR) thinning, and excavation of the ONH [190]. It is the second-leading cause of blindness worldwide [139]. In 2020, an estimated 80 million individuals worldwide had glaucoma and this number is expected to increase to over 111 million by 2040 [3, 7]. While early detection and management are essential for slowing the progression of glaucoma, substantial retinal ganglion cell loss could occur before the functional visual field loss from glaucoma is detected [109]. Standard automated perimetry (SAP) provides valuable information about the stage of glaucoma, but it is not useful for detecting small changes, particularly in early glaucoma. More than 30% of ganglion cell loss is needed before SAP shows the visual field loss [140].

The platform of SD-OCT provides excellent glaucoma-diagnostic performance and reproducibility [117, 191]. SD-OCT is a non-invasive scanning technology that provides precise, quantitative measurements of the retinal nerve fiver layer (RNFL), ONH parameters, neuro-retinal rim (NRR) area, and macular thickness – objective evaluation of these parameters have been used to identify glaucomatous structural damage. The RNFL data of the present study are generated with the Cirrus HD-OCT platform (software version 9.0.0.281; Carl Zeiss Meditec, Dublin, CA, USA) that uses Optic Disc Cube 200x200 protocol to scan the ONH and peripapillary area through a 6 mm square grid, which consists of 200 horizontal linear B-scans and each composed of 200 A-scans.

The predominant method for assessing RNFLdegradation using OCT data involves measuring the peripapillary RNFL through a calculation circle approach. In their study, Schuman et al. [119] established an arbitrary 3.4 mm diameter for the scanning circle.
(Figure 4-1 a).

(a)   (b)
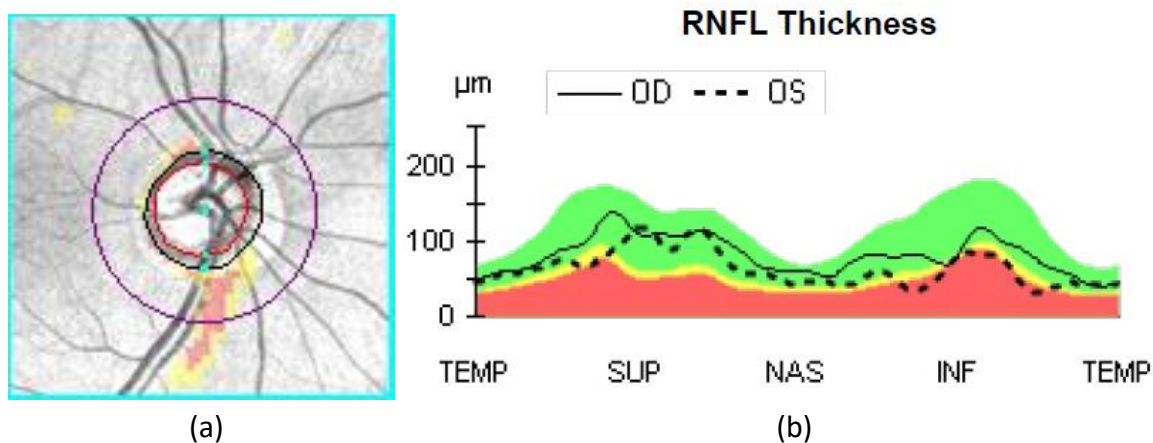
Figure 4-1 (a) RNFL Deviation Map: The OCT device measures the thickness of the RNFLacross the 200 × 200 data cube, wherein individual A-scans denote pixels that correspond to 30 μm wide squares. This enface infrared image of the OCT displays the overlaid BMOcircle in black, the cup border in red, and the calculation circle for RNFL measurement in purple. (b) RNFL Thickness TSNIT Plot: The software of Cirrus HD-OCT retrieves a set of N=256 data points (displayed as black curves for OD/OS) along the y-axis, derived from the 200 × 200 area of the Disc Cube. These data points are extracted for analysis within the circular calculation zone marked on the x-axis.The color-coded background represents the ranges of RNFL thickness based on the observations in the manufacturer's age-matched normative database. Note, as the *y*-axis does not depict relative frequency of the observations, it is not a circular density function.

Subsequently, a standardized 3.46 mm scanning circle was established for the diagnosis of glaucoma, which gained universal acceptance among OCT manufacturers. The initial step of the Cirrus HD-OCT algorithm involves the identification of the optic nerve head's center. Following this, an automated placement of a calculation circle with a diameter of 3.46 mm occurs around this center in a uniform manner. Subsequently, the system retrieves 256 A-scan samples from the data cube and proceeds to measure the thickness of the RNFLat each specific point along the trajectory of the calculation circle. This measurement sequence initiates from the temporal quadrant at the 9 o'clock position in a clockwise direction for the right eye, and for the left eye, it commences at the 3 o'clock position in a counter-clockwise fashion. This sequence follows the Temporal-Superior-Nasal-Inferior-Temporal (TSNIT) order. (Figure 4-1 b).

Notably, the above approach makes the OCT platform, along with certain other biomedical scanning technologies, an ideal example of such generators of data-points that are indexed along multiple well-defined *directions* – here, at given angular positions equally-spaced around

a central point. Unlike the analyses of "linear" data that reside as points on the real line or Euclidean spaces, directional data requires special and altogether different treatment. For instance, a particular direction in two-dimensional plane can be represented as a point on the circumference of a unit circle, or simply as an angle, but neither representation is unique, as both depend on the selection of some appropriate "zero-direction" as the reference starting point of measurement, as well as the sense of rotation, viz., clockwise or anti-clockwise. The unique properties of directional or circular data – for instance, if one wishes to compare or cluster such scans with a suitable distance measure – are appropriately addressed by the fields of directional and circular statistics[146, 192, 193] .

Traditional OCT analysis often involves a simple division of the circle around ONH into 4 quadrants (of 90 degrees each) or 12 clock-hours (of 30 degrees each) to index and record measurements at these angular sections. Past studies on focused analysis of angular sections of OCT RNFL and NRR data have revealed interesting differences between healthy and glaucoma subjects[185]. In this study, we extend such approaches by allowing for analysis of data collected at arbitrarily higher angular resolution that divide the circle (comprising of a total of 360 degrees) into a much larger number (N) of fine sections of $360/N$ degrees each. While our analytical framework holds in general for any value of N, here, we generated a circular sequence of $N = 256$ data points, which measures RNFL thickness at every 1.41 degrees, for each OCT sample from a human eye. Such high-resolution directional measurements, generated via automated processing of standard SD-OCT images, are ideally suited for application of new and insightful techniques from circular statistics for focused OCT RNFL data, as we introduce in this study.

In addition to age, on the structural side, biological heterogeneity of ONH phenotypes, with or without any neuropathy, can also present challenges to clinical decision-making. Unsupervised learning of the heterogeneity of normative ONH phenotypes in a given population can provide key insights into the diversity of baselines that may exist for degenerative neuropathies, and thus, reduce subjectivity in diagnosis [194]. Such knowledge is particularly useful in glaucoma for which different ONH parameters play a combined role in early detection. For instance, in a non-glaucoma multiethnic cohort of Asian individuals, the inter-eye RNFL profile was found by OCT to be less symmetric in Malays and Indians than that in Chinese eyes [168]. Even the manufacturers of OCT technology noted racial differences in optic disc area, RNFL thickness,

etc., when measured using their platform[170]. Thus, representation of diverse populations in OCT databases is a matter of great significance to glaucoma patients, e.g., more than 16 million in India of whom nearly 1.2 million could be blinded by the disease[7] .

The present study addresses multiple important aims that include design of an analytical framework that is capable of defining and using a *directional measure* of RNFL decay based on OCT data and then applying it for focused detection of structural characteristics in glaucomatous eyes from an Indian population. In this direction, our study makes several new contributions. First, we used a large clinical cohort of normal Indian eyes to generate high-resolution OCT RNFL data (as noted above) for each eye. Second, we built a catalog of empirical cumulative distribution functions of RNFL loss at each of the *N* angular-points based on the above data for normal eyes stratified by their age and optic disc size. Third, we developed an analytical framework based on novel applications of circular statistics to OCT data analysis. This includes specifying a directional representation of RNFL decay in a given glaucomatous eye relative to its corresponding normal population. New indices were introduced to measure eye-specific global and local RNFL decay. Fourth, we fit a finite mixture of von Mises distributions to identify for glaucomatous eyes two separate clusters of directional representations of RNFL decay, which were characterized using their clinical covariates. Finally, we demonstrated the utility of the clusters in the improvement of glaucoma classification. In the next section, we describe the Data and Methods, followed by the section on the Results, and end with Discussion.

4.3   Data

The study's participants were sourced from two distinct studies: the LVPEI Glaucoma Epidemiology and Molecular Genetic Study (LVPEI-GLEAMS), which is a population-based research endeavor, and the Longitudinal Glaucoma Evaluation Study (LOGES), conducted as a cross-sectional analysis. These investigations were carried out under the auspices of the L.V. Prasad Eye Institute (LVPEI) situated in Hyderabad, India. Prior to their participation, all subjects provided written informed consent, and the methodologies employed in both studies (identified as LEC 08131 for LVPEI-GLEAMS and LEC 11-252 for LOGES) were subject to review and approval by the LVPEI Institutional Ethics Committee. The execution of these studies adhered rigorously to the principles outlined in the Declaration of Helsinki. The present thesis incorporates a dataset encompassing 3985 healthy eyes (derived from 2225 individuals

without known ocular conditions) and 270 eyes affected by glaucoma (obtained from 210 individuals). The clinical covariates of the cohort are shown in Table 4-1.

Table 4-1: Clinical covariates of the glaucomatous eyes. (Abbreviations: CD = cup to disc, s.d. = standard deviation, min. = minimum, max. = maximum)

Variable (unit)	mean	median	s.d.	min.	max.
RIM AREA (mm^2)	1.09	1.06	0.25	0.49	1.89
DISC AREA (mm^2)	2.28	2.22	0.59	1.23	4.69
AVERAGE CD RATIO	0.7	0.7	0.1	0.4	0.9
AVERAGE THICKNESS (μm)	86.84	87	9.62	54	112
VERTICAL CD RATIO	0.68	0.69	0.1	0.41	0.88
CUP VOLUME (mm^3)	0.52	0.34	0.49	0.04	3.38
DISK DIAMETER (mm)	1.62	1.6	0.22	1.1	2.4

4.4   Method:

All participants underwent a comprehensive ophthalmic assessment. This assessment encompassed a detailed review of their medical and systemic history, measurement of best-corrected visual acuity, capturing of slit-lamp photographs using a Topcon device situated in Oakland, NJ, determination of intraocular pressure through Goldmann applanation tonometry using the Hagg-streit AT 900 from Hagg-streit AG in Switzerland, gonioscopy performed with a Sussman four mirror gonioscope sourced from Volk Optical Inc. in Mentor, Ohio, USA, a thorough examination of the dilated fundus, evaluation of CCT, analysis of Humphrey visual fields (HVF) utilizing the 24-2 Swedish Interactive threshold algorithm (SITA) from Carl Zeiss Meditec Inc. in Dublin, CA, as well as digital optic disc photography and SD-OCT imaging facilitated by the Cirrus HD-OCT device from Carl Zeiss Meditec located in Dublin, CA, USA.[120].

Visual field assessments were included if they exhibited false positive, false negative, and fixation loss rates of less than 20%. Additionally, all stereophotographs of the optic disc were required to demonstrate good quality. The eligibility criteria encompassed individuals aged ≥ 40 years with best-corrected visual acuity of 20/40 or better, spherical equivalent (SE) within ± 6D, excellent quality stereo optic disc photos, and absence of significant media opacities (signal strength ≥ 6). On the contrary, individuals who had undergone intraocular surgery in

the last 6 months, exhibited any retinal (including macular) or neurological disorders other than glaucoma that might influence structural measurements taken with SD-OCT, possessed a RNFLthickness of less than 30 μm, or demonstrated blink and motion artifacts were excluded. The determination of the specific cutoff value for circumpapillary RNFL thickness was guided by prior investigations into the lower limits of structural measurements using SD-OCT.[195].

While healthy eyes were defined by the absence of anterior and posterior pathology, to determine a glaucomatous eye, beyond such measurements as VF and IOP, three glaucoma specialists independently evaluated its digital optic disc photograph. They were blinded of the other clinical findings and the other imaging outcomes of the individuals. Each disc photograph was defined as normal based on the absence of superior and inferior NRR thinning, rim notch, disc hemorrhage, and RNFL defect. Eyes were excluded from the study if there were any discrepancies among the specialists. Zeiss' Cirrus HD-OCT software (version 9.0.0.281; Carl Zeiss Meditec, Dublin, CA, USA) uses the segmentation algorithm of the OCT software to identify the RNFL and measures its thickness on the circular peripapillary scan. For each eye (glaucomatous or normal) that was included in our analysis, the same algorithm was used for obtaining its high-resolution OCT RNFL thickness data at $N(= 256)$ equally-spaced ($360/N = 1.41$ degrees apart) angular-points $\{j: j = 1, 2, \ldots N\}$ in counterclockwise (TSNIT) direction around the circle (0-360 degrees).

The eyes were first stratified according to different combinations of age and disc size. Three age categories: Age1 (40-49 years), Age2 (50-59 years), and Age3 (>=60 years), and three optic disc size categories were used. These are Small (<1.6 mm^2), Average (1.6-2.6 mm^2), and Large (>2.6 mm^2) discs. The cutoffs for disc size are derived from our prior work based on clinical evaluation which was later objectively quantified by HRT measurements[196] . Thus, over all $3 \times 3 = 9$ possible combinations of age and disc size, we determined for each eye i (normal or otherwise) exactly which group (call it G) it belongs to (denoted as $i \in G$).

4.4.1 Angular Decay Calculation:

For each group G, and for each angular-point j, we take the set $R_{G,j} = \{RNFL(i_N, j): \text{normal eye } i_N \in G\}$ of RNFL values at the angular point j for every normal eye $i_N \in G$, and compute the empirical (cumulative) distribution function

$$eCDF_{G,j}(x) = |r \le x : r \in R_{G,j}| / |R_{G,j}|.$$

Each function $eCDF_{G,j}$ takes as input an RNFL thickness measurement $RNFL(i_G, j)$ of a glaucoma eye $i_G \in G$ at an angular-point j, and returns the corresponding quantile value, which lies in the range [0,1] where 0 and 1 correspond respectively to the lowest and the highest of all RNFL observations at the same angular-point j among the normal eyes that belong to G.

The Angular Quantile of RNFL data for an eye $i \in G$ at an angular-point j is computed as

$$Angular\ Quantile(i,j) = eCDF_{G,j}(RNFL(i,j)).$$

We note that the above quantile value provides a measure that is *relative* to all the RNFL observations made at a particular angular-point over the entire population of normal eyes of the corresponding group. Then the Angular Quantile is computed at each angular point j, which yields for an eye $i \in G$ a circular sequence:

$$Angular\ Quantile(i) = \{Angular\ Quantile(i,j) : j = 1, \dots N\}.$$

Thus, we define Angular Decay of an eye $i \in G$ as the following circular sequence:

$$Angular\ Decay\ (i) = \{1 - Angular\ Quantile(i,j) : j = 1, \dots N\}.$$

Hence, for a given eye, at any angular-point, Angular Decay admits a value that lies in the range [0,1]. A higher value of Angular Decay in an eye corresponds to an OCT RNFL measurement that ranks relatively lower among the RNFL observations in the normal population.

4.4.2 Directional representation of RNFL decay:

Based on the normal RNFL quantile threshold τ, we characterize the Angular Decay values of a given eye in terms of those that are above versus below τ. Each distinct sequence of contiguous Angular Decay values exceeding τ is defined as a "petal". Therefore, each petal (denoted by Π) represents an angular range that shows higher than normal decay. For a given threshold τ, we denote the number of such petals in an eye by ν_τ. By fixing τ, say, to 0.75, we can drop the suffix in the above quantity (as well as those measurements defined below for each eye) and simply refer to it as ν.

In particular, we are interested in the "widest" petal (π^*) of an eye, which is the largest contiguous sequence of Angular Decay values exceeding τ. We use an algorithm to traverse the circle counter-clockwise and at each successive angular-point $j = 1,2, \dots N$ on the circle of

eye i test whether $RNFL(i,j)$ exceeds τ. This yields the starting point (θ_s) and the end point (θ_t) of each existing petal. The counter-clockwise circular interval $[\theta_s, \theta_t]$ defines the angular range of a petal, and the widest petal, i.e., the one with the largest angular range, is identified as π^*.

Since for each angular-point j in the range of π^*, the Angular Decay value is known, we compute the *weighted circular mean* (20) of π^* where the Angular Decay value at j is used as the weight $w_j \in [0,1]$ for the angle θ_j at j. We denote this decay-weighted mean angle (of the largest contiguous region of the eye with above normal Angular Decay) by μ^*, the quadrant specific inverse of the tangent (see Equation 1.3.5 on page 13 of (12))

$$\mu^* = \tan^{-1}\left(\sum_{j\in\pi^*} w_j \sin\theta_j / \sum_{j\in\pi^*} w_j \cos\theta_j\right).$$

The weighted mean angle allows μ^* to provide a directional representation of an eye in terms of its regional concentration of relatively high RNFL loss. If no such region exists in a given eye for the chosen value of τ, then its μ^* and π^* are denoted by 'NA'.

For unsupervised identification of clusters of eyes in a given collection with distinct directional signatures, the angular data (μ^*) are fitted with a finite mixture of K univariate von Mises distributions $\sum_{k=1...K} \alpha_k \, vM(\mu_k, \kappa_k)$ where μ_k and κ_k are respectively the (angular) location and the concentration parameters of the kth component (cluster) of the above mixture whose non-negative proportions $\{\alpha_k : k = 1, ... K\}$ add up to 1. The optimal model is selected using the Bayesian Information Criterion (BIC) (21). The absence of at least 2 distinct clusters in a given collection of eyes can be checked for this data with a test of circular unimodality (12). The von Mises distribution $vM(\mu, \kappa)$ has the probability density function (pdf) $f(x|\mu, \kappa) = \exp(\kappa \cos(x - \mu))/2\pi I_0(\kappa)$, where $I_0(\kappa)$ is the modified Bessel function of the first kind of order 0 [146].

4.4.3 Indices of local and global RNFL loss:

We define 2 new indices of RNFL loss in a given eye. A measure of *local* RNFL loss in an eye is defined as the proportion of the size of the angular range of π^* to the entire circle. This local loss index is denoted by λ^*. Further, we also introduce a measure of *global* RNFL loss in an eye defined by the proportion of the total angular range of all ν petals in this eye to the entire circle. This global loss index is denoted by Λ.

$$\lambda^* = \{Number \ of \ points \ j \ in \ the \ angular \ range \ of \ \pi^*\} / N.$$

$$\Lambda = \sum_{\Pi=\Pi_1}^{\Pi_v} \{Number\ of\ points\ j\ in\ the\ angular\ range\ of\ \Pi\}\ /\ N.$$

Thus, both measures, λ^* and Λ, take values in the range [0,1], where 0 and 1 represents respectively the minimal and the maximal RNFL loss in a given eye with respect to a chosen threshold τ. We note that $0 \leq \lambda^* \leq \Lambda \leq 1$.

4.4.4 Focused RNFL diversity in the population:

For normal eyes of a given disc size, with increase in age, we tested if the population-wise RNFL thickness is conserved in any region around the circle. For a younger and older age-group pair, we performed the Kolmogorov-Smirnov test of equality of their RNFL distributions (given by the eCDF-s defined above) at each angular-point. For smoothing and visualization of local variation, a sliding window of length $K = 15$ consecutive angular-points was run over the full circle. The p-values of the K Kolmogorov-Smirnov tests within each window were combined using the Fisher's sum $X^2 = (-2) \times \sum_{k=1}^{K} \log p_k$, which follows a chi-square distribution with $2K$ degrees of freedom. Thus, a combined p-value of the above sum is computed at each window, and the focused significance of normal RNFL diversity is plotted as a circular heatmap.

4.4.5 Implementation:

The above-mentioned framework, including all its functions for analysis and visualization, is implemented using the R platform. The empirical cumulative distributions and their 2-sample Kolmogorov-Smirnov tests were computed with the ecdf() and ks.test() functions respectively. The R package 'movMF' was used for fitting a finite mixture of univariate von Mises distributions with Expectation Maximization (EM) algorithm to the RNFL decay data represented as a set of angular measurements. For visualization of the Angular Decay in a given eye, our R program plots this function on a circular scale. Against a background blue circle of radius equal to a selected normal RNFL quantile threshold τ, the Angular Decay function is shown as a circular curve which appears in blue or red when it is below or above τ respectively. Each of the contiguous red sections is a separate petal Π and the widest petal π^* with the largest angular range is shown in red (see Figure 4-4 for examples). The R package 'ctree' was used

to build binary decision tree models for glaucomatous versus normal eye classification[197]. The combined p-value was computed using the 'metap' package in R.

4.5   Results

4.5.1 Direction-specific catalog of normative RNFL thickness:

To construct a normative reference for the measure of RNFL loss in a given glaucomatous eye, we computed a population-based catalog of a total of $9 \times 256 = 2304$ direction-specific empirical distribution functions based on RNFL data from the 3985 normal eyes that belong to 9 groups (as defined in the Methods) and for $N = 256$ angular-points. The steps are illustrated with an example in Figure 4-2.
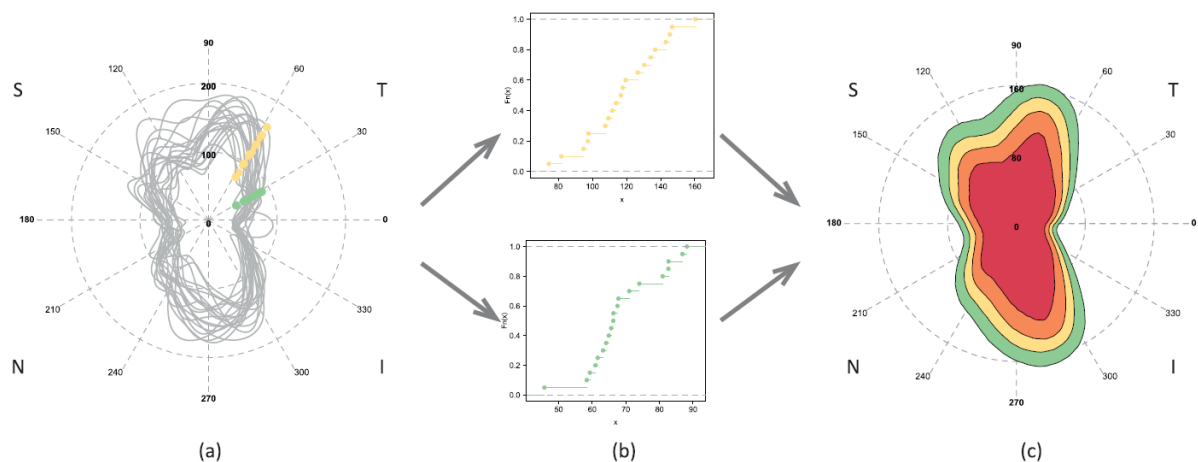
Figure 4-2 The circular OCT RNFL thickness data for normal eyes belonging to a group are overlaid and shown in plot (a). Two examples of angular-points are used to mark the OCT data as sets of green and orange points. Plot (b) shows the eCDF functions for these two OCT datasets. Plot (c) shows the ranges of normal RNFL thickness values for different Angular Quantiles. As examples, the concentric ranges for the 5th, 25th, 50th, and 75th percentiles are shown with red, orange, yellow, and green contours respectively.

For a given group G, the RNFL thickness data of the normal eyes in G are plotted as circular curves in Figure 4-2(a). The RNFL measurements of these eyes at two arbitrarily fixed angular-points are marked by 2 sets of points in orange (θ_1) and green (θ_2). The empirical distribution functions of the sets θ_1 and θ_2 are shown in the corresponding colors in Figure 4-2(b). We cataloged such empirical distribution functions at every angular-point for each G. These yield the quantile of normative RNFL thickness at any angular-point. For instance, in Figure 4-2(c),

for the 5th, 25th, 50th, and 75th percentiles, the increasing ranges of normative RNFL thickness are shown as successive contour curves in red, yellow, orange, and green respectively. Thus, if the RNFL circular data of any new eye i is overlaid on this contour plot, it will reveal the precise extent of any global (based on magnitude) or local (based on magnitude *and* direction) RNFL loss relative to the normative RNFL thickness of the group that i belongs to.

4.5.2 Direction-specific Angular Decay:
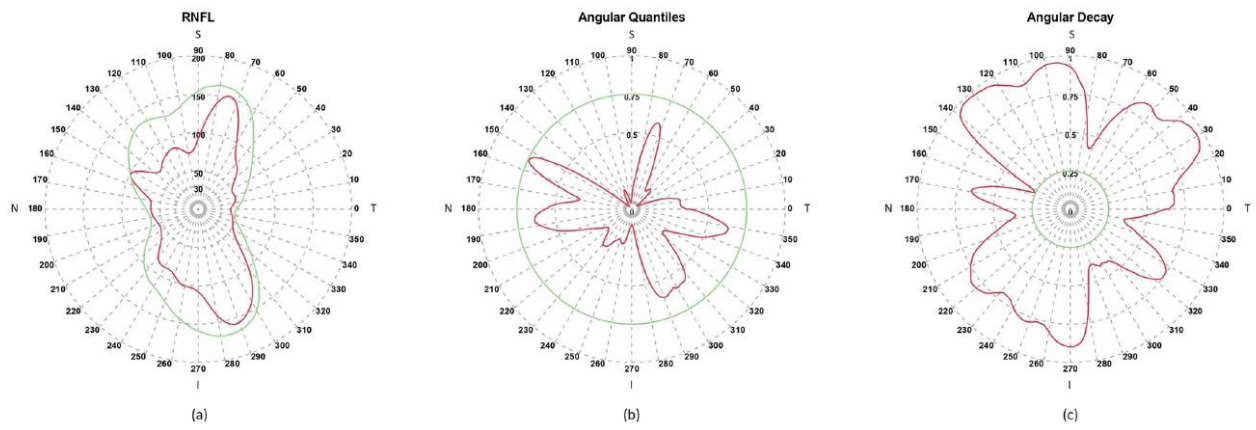
Figure 4-3 The circular OCT RNFL thickness data of a glaucomatous eye is shown as a red curve in plot (a). For reference, the 75th percentile of normal RNFL thickness is shown with a green curve. The glaucomatous and normal RNFL data are converted to their respective Angular Quantiles (of range [0,1]) and shown as the corresponding red and green curves in plot (b). In (c), the complementary plot of (b) shows the RNFL Angular Decay (of range [0,1]) as red and green curves for the glaucomatous and normal cases respectively

we calculated this as a circular function based on an eye i's high-resolution OCT RNFL thickness data at each angular-point j. Given the group G that i belongs to, and j, the corresponding distribution function from the catalog (described above) is used for calculating the precise quantile value of RNFL thickness of i at j. This is shown as the red curve in Figure 4-3(a). To focus on the last quartile (i.e., 75th-100th percentile) where the RNFL loss is likely to be common in glaucomatous eyes, a green curve shows the normative RNFL thickness at quantile 0.75 for reference. The quantile values of each RNFL curve at every angular-point yield the corresponding (green and red) Angular Quantile curves shown in Figure 4-3(b). The exactly complementary plot (computed as $Angular\ Decay = 1 - Angular\ Quantile$) produces the direction-specific values of Angular Decay of the eye i shown in Figure 4-3(c).

As the Angular Quantile function has the range [0,1], so does Angular Decay. The RNFL data and the corresponding Angular Quantile and Angular Decay functions shown as the red curves in Figure4-3(a), 4-3(b) and 4-3(c) respectively are due to an arbitrary (but real) eye that serves as an example.

4.5.3 Contiguous regions of Angular Decay:

To identify locally distinctive features of an eye's RNFL loss, we applied a circular tracing algorithm that uses a user-defined quantile threshold (τ) to demarcate the "petals" which are contiguous regions of Angular Decay values exceeding τ. Thus, we identified the petals for a given eye to compute insightful patterns of RNFL loss using 4 new indices. To illustrate, the Angular Decay functions of 6 glaucomatous eyes are shown in Figure 4-4 as examples. For each eye, the selected quantile threshold is shown as a pale blue circle of radius τ. The Decay function is shown as blue or pink curves depending on whether it takes values below or above τ. Each contiguous pink angular section represents a separate petal denoted by Π. In a given eye, for fixed τ, the total number of petals is denoted by ν; and the petal with the largest angular range is depicted as a red curve and denoted by π^*.

4.5.4 Angular representation of an eye's RNFL decay:

The systematic identification of π^* allows us to define an angular representation of an eye in terms of its largest contiguous section of above normal RNFL decay. For this purpose, we compute the weighted circular mean of π^*, which is the angle (denoted by μ^*) around which the RNFL decay within the widest petal of the eye is concentrated. The weighting used in the calculation of μ^* captures the contribution of the variations of the shape of π^*, which results in a more precise directional representation of the decay than its unweighted counterpart. The angular value of μ^* for each eye is shown as a dashed red line in the Angular Decay plots of Figure 4-4.

Sample: 1
(Group: Age 1, Cluster 1)

$\tau = 0.75$, $\Lambda = 0.45$, $\nu = 5$, $\mu^* = 113.9°$, $\lambda^* = 0.17$

Sample: 2
(Group: Age 1, Cluster 2)

$\tau = 0.75$, $\Lambda = 0.39$, $\nu = 3$, $\mu^* = 280.8°$, $\lambda^* = 0.21$

Sample: 3
(Group: Age 2, Cluster 1)

$\tau = 0.75$, $\Lambda = 0.46$, $\nu = 4$, $\mu^* = 134.6°$, $\lambda^* = 0.24$

Sample: 4
(Group: Age 2, Cluster 2)

$\tau = 0.75$, $\Lambda = 0.43$, $\nu = 3$, $\mu^* = 270.1°$, $\lambda^* = 0.23$

Sample: 5
(Group: Age 3, Cluster 1)

$\tau = 0.75$, $\Lambda = 0.09$, $\nu = 4$, $\mu^* = 157.3°$, $\lambda^* = 0.04$

Sample: 6
(Group: Age 3, Cluster 2)

$\tau = 0.75$, $\Lambda = 0.22$, $\nu = 3$, $\mu^* = 270.1°$, $\lambda^* = 0.19$
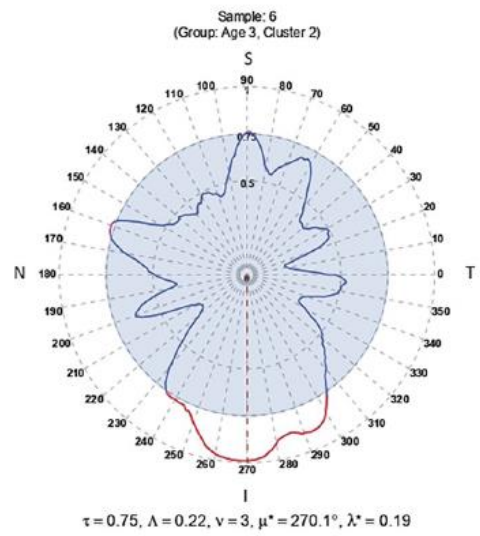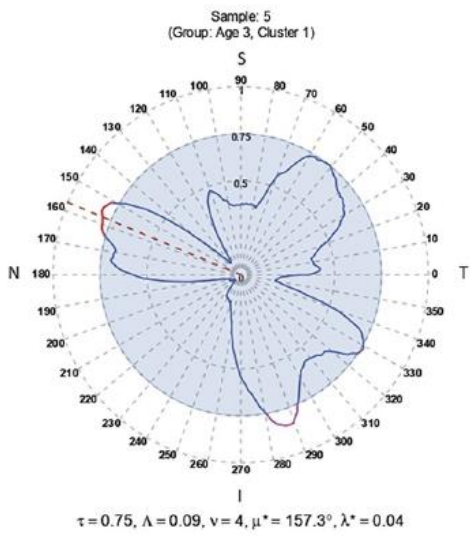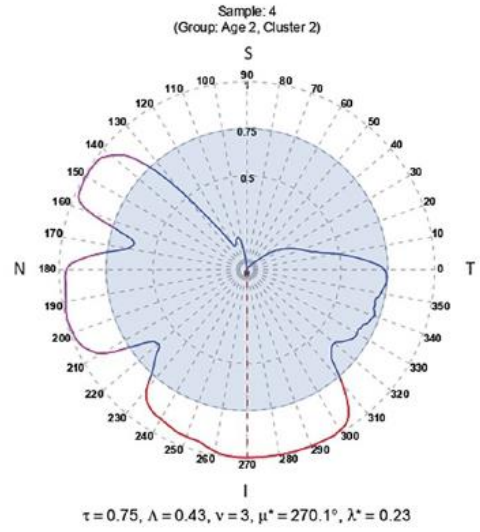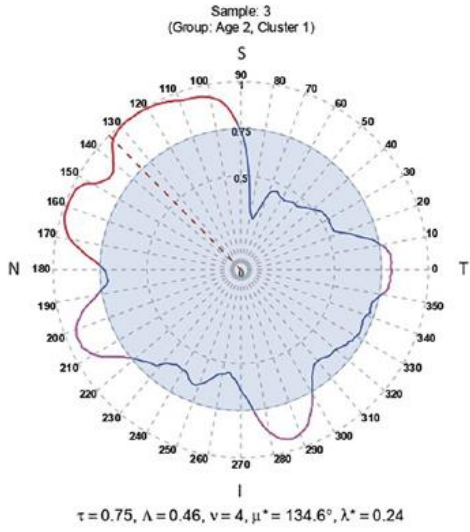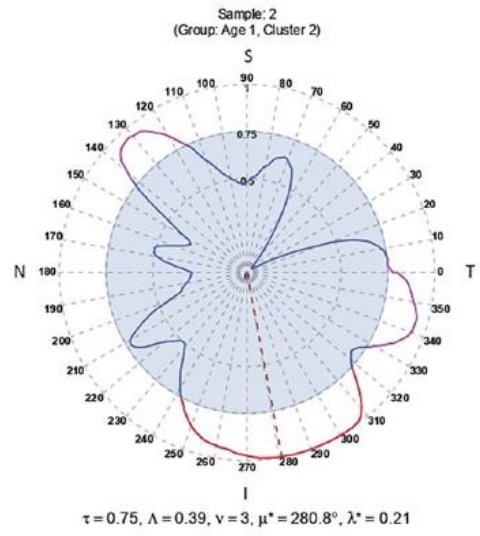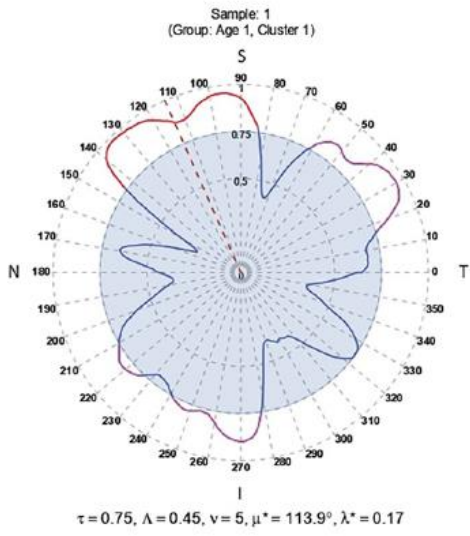
Figure 4-4 For 6 samples of glaucomatous eyes, the RNFL Angular Decay functions are plotted. Each plot contains a pale blue circle of radius equal to the threshold of decay τ set at 0.75. The petals lying above the pale blue circle are shown as pink curves. The petal with the largest angular range is marked with a red curve. The weighted circular mean angle of the red petal is shown with a dashed red line. Various indices computed for each eye are shown below the plot of each sample.

4.5.5 Local and global RNFL loss:

In addition to such directional information, we also quantified the proportion of decay that is described by π^* as it represents the widest petal of an eye. Thus, we define a measure of *local loss* of RNFL of an eye, denoted by λ^*, as the fraction of the full circle that is covered by the angular section corresponding to π^*. Further, we introduce another index for measuring the *global loss* of RNFL for an eye, denoted by Λ, as the fraction of the full circle that is covered by the union of the angular sections corresponding to all the petals in an eye. In Figure 4-4, below each eye's Angular Decay plot, the computed values of its different indices are shown.

4.5.6 Circular clustering and clinical characterization:

To demonstrate the focused structural diversity of RNFL loss, we applied finite mixture modeling, a popular technique for unsupervised learning of clusters, on the angular representations (as described above) of the glaucomatous eyes. In particular, we fitted a finite mixture of univariate von Mises distributions separately to 3 directional datasets on μ^* directional values of glaucomatous eyes. First, we partitioned the cohort of glaucomatous eyes into 3 age-groups: Age1 (40-49 years), Age2 (50-59 years), and Age3 (60+ years). Then, each group's dataset on μ^* was modeled with a mixture of $G\ (= 1, 2, \dots 5)$ von Mises distributions with the aim of checking whether there are different clusters of glaucomatous eyes in these age-groups. Here, $G = 1$ implies a unimodal distribution of data that does not multiple distinct clusters. The models were fitted using the EM algorithm and the optimal model was selected based on the value of G for which BIC is the minimum (Figure 4-5).
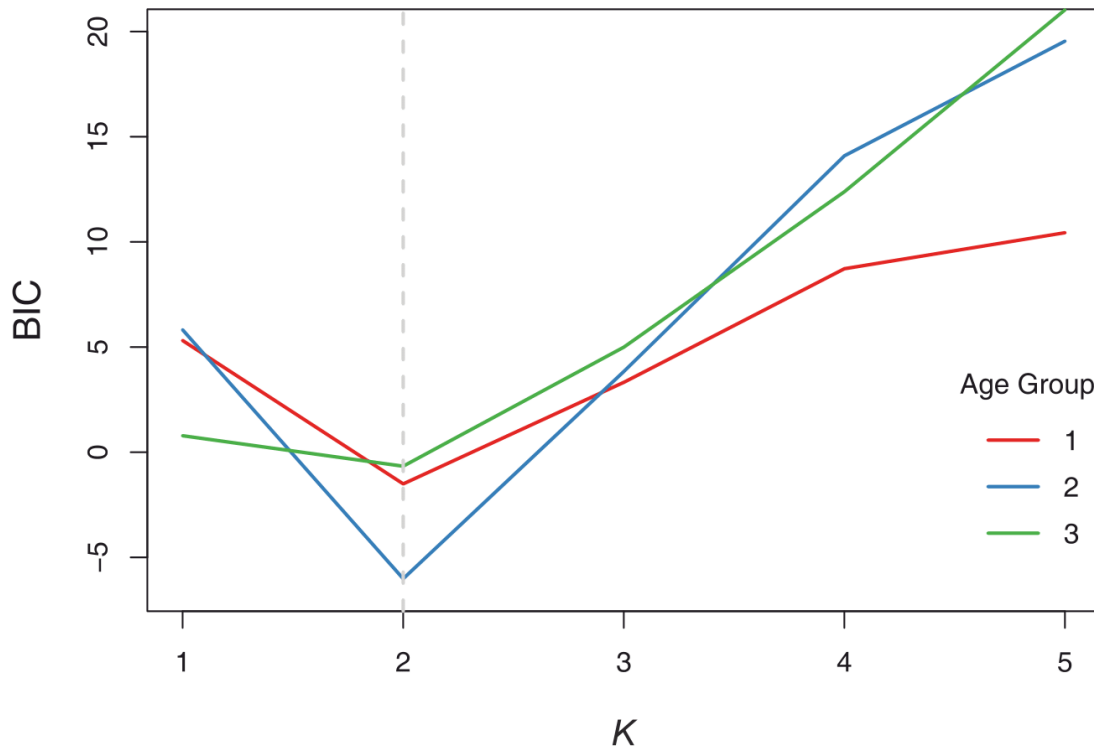
Figure 4-5 **The BIC values corresponding to different choices of the number of components, K, in the mixture model of von Mises distributions fitted to circular data from 3 age-groups.**

Interestingly, the optimally fitted von Mises mixture models identified 2 clusters (representing $G = 2$ mixture components) of glaucomatous eyes – in all 3 age-groups – with very distinct directional patterns of RNFL decay (red histograms in Figure 4-6). The estimated parameters of the selected models are shown in Table 4-2. In particular, we note that the 2 clusters align clearly across the 3 age-groups in terms of their very distinct estimates of the angular location parameter (μ) of the von Mises distribution that was fitted for identifying each component of the mixture.

Table 4-2:The estimates of the parameters of a 2-component mixture model of von Mises distributions for 3 age-groups. The Bayesian Information Criterion (BIC) value for each optimal model is shown.

Group	Component	μ (degrees)	κ	α	BIC
Age1	Cluster1	101.97	11.044	0.304	-1.506
	Cluster2	296.26	1.556	0.696	
Age2	Cluster1	120.01	0.882	0.616	-6.015
	Cluster2	286.48	13.833	0.384	
Age3	Cluster1	146.45	1.247	0.784	-0.665
	Cluster2	291.77	14.932	0.216	

Relative to the higher RNFL thickness of the normal eyes in the Superior (S) and Inferior (I) sectors (see the green distribution in Figure 4-1 (b)), the decay in the glaucomatous eyes is concentrated around these regions (the red "modes" in Fig. 7) for all 3 age-groups. The rugplot at the bottom of the red histograms in Figure 4-6 depicts the membership of the glaucomatous eyes into these 2 clusters – using pink and orange points – as assigned by the fitted mixture model. In clear contrast to the glaucomatous eyes, the distributions of angular representation of RNFL data from the normal eyes of the same age groups appear much less distinctly multimodal in the blue histograms in Figure 4-6.
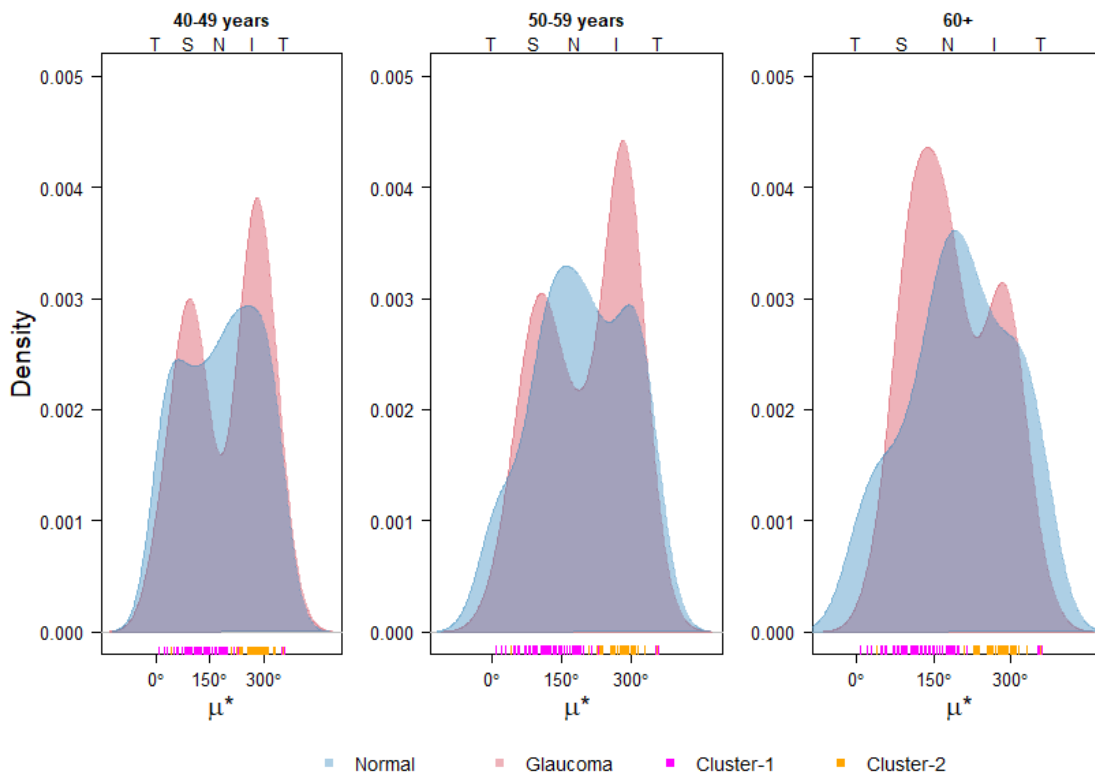
Figure 4-6 The distributions of circular data based on the angular representation of glaucomatous eyes from 3 age-groups are shown as red histograms in plots (a), (b) and (c). The corresponding normal

For further characterization of the 2 clusters of glaucomatous eyes identified in different age-groups solely by their angular representation of RNFL decay, we studied their clinical covariates. In Table 4-3, we note that many of the clinical covariates of the glaucomatous eyes that belong to these 2 clusters have distinctive mean values. The values for various parameters such as rim area, disc area, average CD ratio etc, are quite different between the normal and both the clusters of glaucomitics eye. This is a clear indication that our technique has clearly discriminated between normal and glaucoma eye.

Table 4-3: The mean values of the clinical covariates of the glaucomatous eyes belonging to the 2 identified clusters in 3 age-groups. (Abbreviation: CD = cup to disc)

Variable (unit)	Cluster1			Cluster2			Normal		
	40-49 years	50-59 years	60+	40-49 years	50-59 years	60+	40-49 years	50-59 years	60+
RIM AREA (mm^2)	1.09	1.09	1.13	1.02	1.09	1.05	1.36	1.34	1.34
DISC AREA (mm^2)	2.36	2.47	2.2	2.2	2.23	2.14	1.97	1.97	2.01
AVERAGE CD RATIO	0.72	0.72	0.68	0.7	0.68	0.7	0.52	0.53	0.54
AVERAGE THICKNESS (μm)	87.65	89.57	86.96	87.42	85.29	83.16	94.37	92.31	90.78
VERTICAL CD RATIO	0.68	0.69	0.66	0.69	0.69	0.69	0.49	0.5	0.52
CUP VOLUME (mm^3)	0.57	0.66	0.4	0.62	0.51	0.36	0.18	0.18	0.2
DISK DIAMETER (mm)	1.64	1.7	1.59	1.59	1.58	1.57	1.5	1.51	1.53

A visual comparison of such covariates across the 2 clusters in the 3 age-groups is shown in the boxplots of Figure 4-7. For consistency, we used the same pink and orange color scheme for the corresponding clusters as in Figure 4-6.
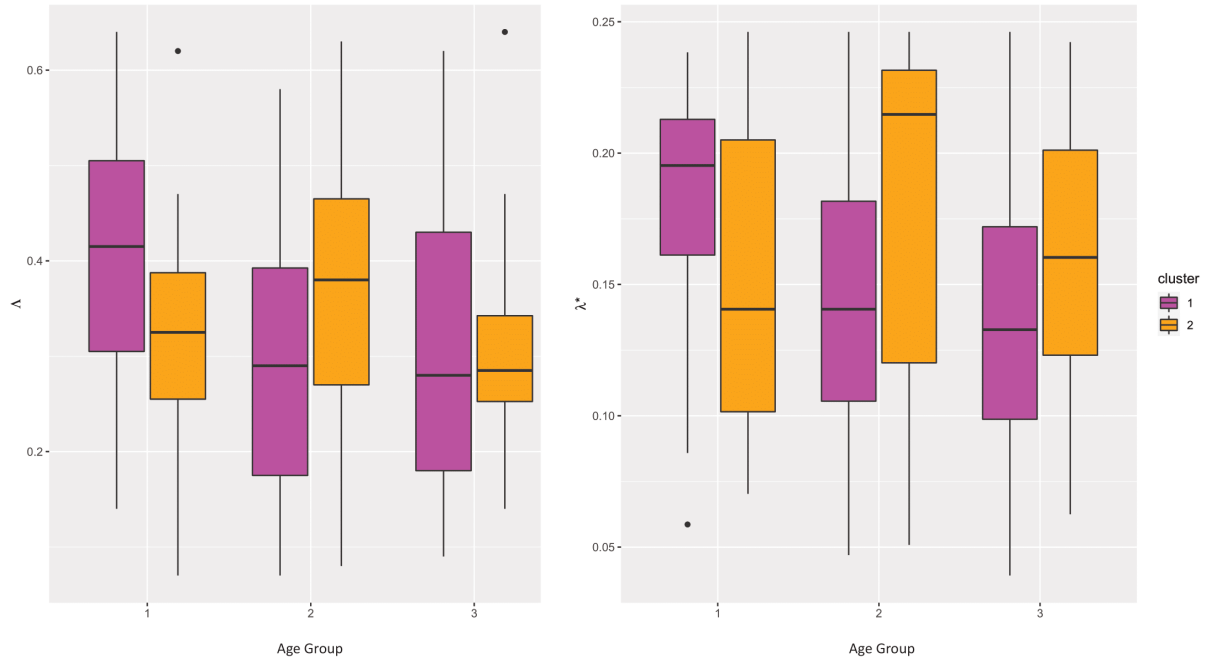
Figure 4-7 Boxplots comparing the values of the indices for (a) global and (b) local RNFL loss (y-axis) in the eyes belonging to 2 clusters in 3 age-groups (x-axis).

Further, when we compared the values of the new indices of global (Λ) and local loss (λ^*) of RNFL, we again observed clear differences in the structural changes across the same clusters in different age-groups. Intriguingly, whereas the covariates and the global loss index could distinguish between the clusters in the glaucomatous eyes of the two younger groups, for the oldest group (Age3: ≥ 60 years), the distinction is pronounced only when we compared their indices of local RNFL loss λ^*. We note that λ^* clearly admits of higher values for the eyes in cluster 2 than those in cluster 1, which further underscores the utility of directional or focused analysis of OCT data.

4.5.7 Glaucoma classification using circular data:

While the primary aim of the present study is to introduce circular representation and analysis of OCT data, we extended our application of circular data to improve the classification of glaucomatous versus normal eyes. We demonstrated this in 2 steps: first, we built classification tree models based on variables that measure the global (Λ) and local (λ^*) RNFL losses in a given eye. We show these models in Figure 4-8 (a) and (b) respectively for classification of samples in Age Group 3, which has the highest number of glaucoma samples. Interestingly, we

noted that the majority of these samples (54.5% and 68.6%) were classified by the central leaf node of either tree with relatively high impurity (Gini index 0.43 and 0.4). The overall model specificities are 0.762 and 0.761 respectively. In contrast, by using the circular clustering labels obtained earlier (by von Mises mixture) to group the eyes first, we achieved improved classification of the same Age group 3 by nodes of much less impurity (maximum Gini index 0.37) in the tree model depicted in Figure 4-8 (c), which has higher specificity of 0.804.
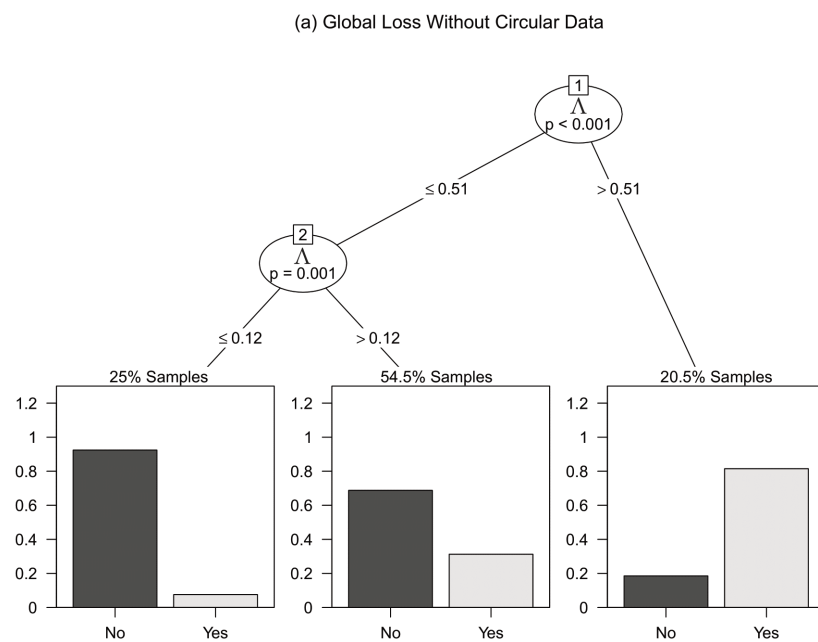
Figure 4-8 (a)

(b) Local Loss Without Circular Data

Figure 4-9 (b)

113
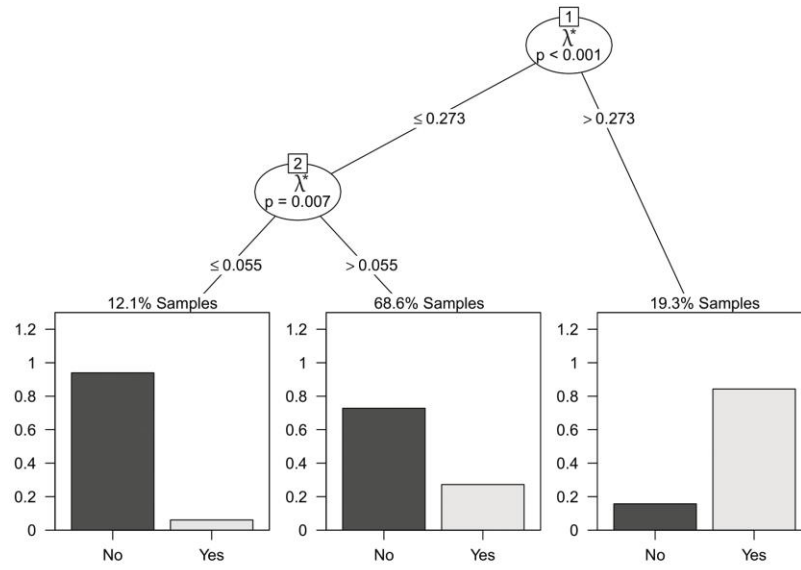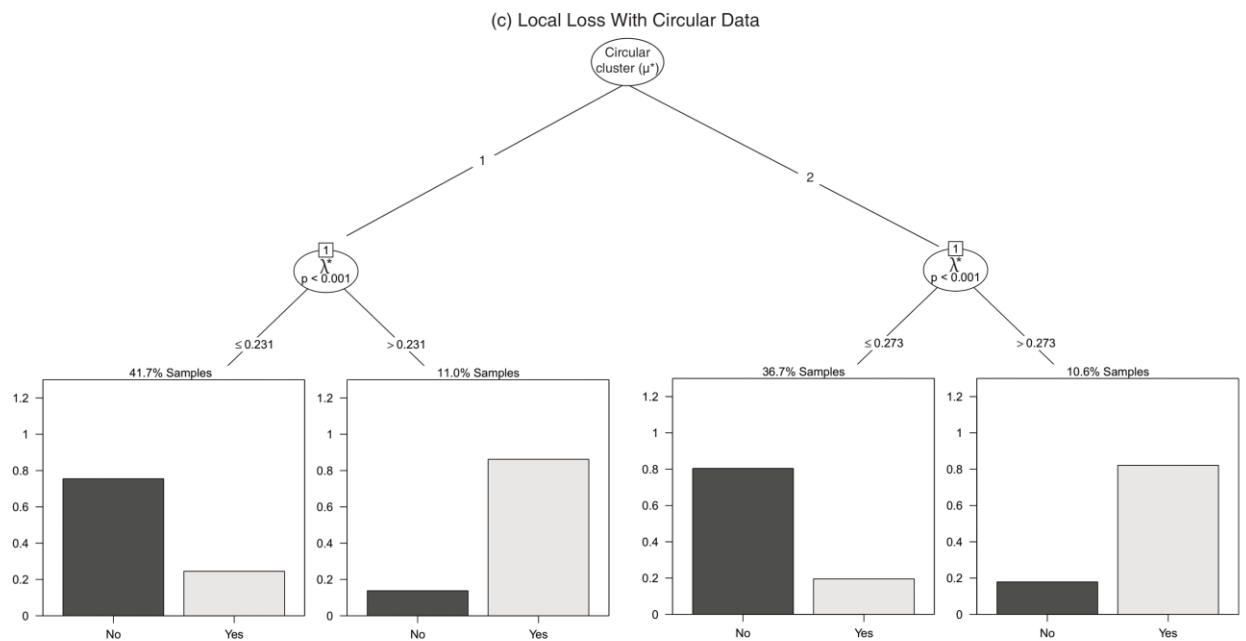
Figure 4-10 (c)

4.6   Discussion

While the investigation of localized RNFL defects in glaucoma diagnosis has been conducted in the past [198], high-resolution data from the recent platforms allow us to approach this topic more systematically. SD-OCT has the ability to produce reproducible scans of an eye that could be read out as a high-resolution sequence of measurements made at a fairly large number ($N = 256$ in the present study) of evenly-spaced angular-points defined on a circular coordinate system. Yet, both of these insightful characteristics of OCT data – its granularity (N) and circularity – generally remain unutilized in its conventional downstream statistical analyses. Indeed, clinical applications may use results that are summarized over a much smaller number of angular sections, e.g., 4 quadrants or 12 clock-hours, and then too, the RNFL measures in these sections are utilized not for their directional information but only to capture the magnitude of loss therein. Hence, any comparison of OCT RNFL data – as is often done for patient classification or comparing across time-points – is generally restricted to the linear scale whereby no angle-specific distinction could be made made among the focused losses.

In this study, we presented a new *circular statistical framework* that uses both the directionality as well as the magnitude of OCT RNFL thickness data. The framework could be straightforwardly generalized and applied to other comparable types of data such as OCT NRR, Fundus imaging, etc. In addition, we generated new high-resolution SD-OCT data. Using this input, an innovative output in the form of a directionally focused Angular Decay function was computed for each sample. Such a function provides a quantitative measure of RNFL loss – relative to the normal population – precisely at every 1.41 degree in a circular coordinate system. Notably, this allows for a well-defined data representation that takes values in the common range of [0,1] at a common sequence of angular-points around the circle with a fixed origin (0 degrees) for each sample and across all groups. Our standardized approach thus facilitates various downstream applications.

To begin with, we apply a tracing algorithm to identify the petals that are contiguous regions – defined as precisely specified angular intervals – at which the RNFL decay in a given eye exceeds a threshold that is specified with an intuitive quantile of the corresponding normal RNFL empirical distribution. To avoid a fixed choice of this threshold, our tracing algorithm uses a flexible parameter (τ). For instance, to focus on the last quartile of RNFL decay in our illustration, we used the value of $\tau = 0.75$ although we understand that this choice is arbitrary. Importantly, based on the above framework, we introduced two measures of RNFL decay – both local (λ^*) and global (Λ) – in a given eye relative to such decay in the normal eyes. In fact, we combined both the directionality (denoted by an angle) and the magnitude (by a weight) obtained from the OCT data in the form of a weighted circular mean to define a directionally focused *angular* representation (μ^*) of a given eye's RNFL decay.

The new representation opens up the potential for creative applications of the rich variety of tests and tools that exist in the fields of directional and circular statistics for insightful OCT data analysis. Such tools can provide novel insights into directionally focused segments of otherwise global phenomena such as RNFL loss in glaucoma patients. For instance, for the available left and right eye-pairs of the cohort in this study, the usual Pearson correlation of global RNFL loss (Λ) is 0.44. In contrast, the circular version of Pearson's correlation [146] computed on the angular representation (μ^*) of the same pairs of eyes is only 0.19, thus suggesting potential scope of circular statistics to expand our capacity for finer differentiation of the clinical cohorts.

Interestingly, we demonstrated this capacity by modeling the structural heterogeneity in our cohort using a finite mixture of von Mises (also known as "circular normal") distributions to the directional data generated in the present study. It led to unsupervised identification of 2 clusters with distinctive focused patterns of RNFL decay in 3 different age groups of a cohort of glaucomatous eyes from the Indian population. Moreover, the new indices of global and local RNFL loss were effective in capturing the structural differences between these glaucoma clusters. Furthermore, if the computed cluster membership of the eyes is used for grouping them first (i.e., according to the direction of focused decay), then it improved their (glaucoma versus normal) classification based on the local loss therein (as compared to conventional classification by global loss without circular clustering). These indices could be used either singly or in combination for downstream analysis of OCT data such as quantitative comparison of time-points or mining for new patterns of RNFL decay.

To accurately account for the effects of normal variation of RNFL thickness in a population, we emphasize that, ideally, normative databases of healthy eye OCT phenotypes that are large and racially representative should be developed. However, often such collections of healthy samples tend to be small or moderately sized, e.g., the normative database of Cirrus HD-OCT platform included just 284 subjects of which the representation of Indians is about 1% [170]. Clearly, this does not adequately reflect the 2020 projections about India which is to become the second in global glaucoma numbers, surpassing Europe[7]. In this direction, our rigorous cataloging of the more than 2300 empirical distribution functions of RNFL thickness, which takes into account the variations of age and optic disc sizes in the normal population, provides a quantitative reference for the focused angle-specific decay in the glaucomatous eyes. Further, we used Kolmogorov-Smirnov test to illustrate the angle-specific differences in these distributions across age groups and disc sizes. Importantly, the catalog addresses the disparity in the form of under-representation of Indian structural phenotypes in the common normative OCT databases.

We understand that there are certain limitations of the study. While our adaptation of the image processing algorithm, as implemented in the standard Cirrus HD-OCT software package, for generating the $N = 256$ point circular RNFL data is straightforward, we understand that such high-resolution data extraction is not common in OCT analysis. Moreover, the focused and

angular measure of the new directional representation of RNFL loss renders it difficult to conduct a direct comparison with any traditional indices that are based on the magnitude (but not directionality) of such decay. Finally, we identified, in 3 different age-groups, 2 recurrent clusters of glaucomatous eyes with very distinct directional signatures of RNFL decay. In future work, we will provide further characterization of the clusters with more rigorous investigation that goes beyond the scope of the present study.

While past studies have used spatial, trigonometric and Fourier analysis on OCT data, e.g., [199-202], our framework, involving directional or circular data, is methodologically quite different from those approaches. While our earlier platform CIFU [194] focuses on the aspects of shapes and curves in OCT NRR data, as in FDA, it does not explicitly address the directional characteristics of OCT data or angular measurements thereof. We have demonstrated the utility of both the granularity and the directionality of OCT data and combined these with an angle-based description of patterns that could be studied with innovative downstream analyses. Indeed, processes or platforms that generate data which are capable of being represented as angular values (e.g., wind direction, neuronal activity) require altogether different statistical approaches as compared to usual linear data [192]. We think that the new directional representation of OCT data to capture an eye's RNFL decay will pave the way for future applications of the rich methodology of directional and circular statistics[146, 192, 193] to eye data analysis.

Chapter 5

Conclusion and Future Work

5.1 Circular functional analysis of OCT data for precise identification of structural phenotypes in the eye

In this work, we contributed to the following:

(1) generated OCT NRR data in the form of 180 circular measurements of NRR thickness in a given eye,

(2) introduced CIFU, a computational pipeline for <u>CI</u>rcular <u>FU</u>nctional data modeling and analysis that is demonstrated using the OCT NRR dataset, and

(3) addressed the disparity of representation of the Asian Indian population in normative OCT databases.

In addition to its sample size and racial representation, perhaps the most noteworthy feature of the present OCT dataset is its unique high-resolution measurements of NRR thickness around a circle at every 2 degrees. As clustering with curves show, the focal variations could be more nuanced than that suggested by a general rule, e.g., ISNT, and in low-resolution data, it may be difficult to detect focal changes within the confines of pre-determined inflexible sectors. Be it circular data from OCT, or optic phenotypes in general, they seem suitable as candidate applications of circular statistics, and yet, we are unaware of any major previous studies in this regard. Further, the high-resolution also allowed the data to be closely approximated by continuous curves, and thus, specified by the corresponding functional representation. While clustering of circular point data and curve and (non-circular) functional data have been addressed by past studies, the present clustering of curves in the form of circular functions is possibly a novel application.

A prospective cohort study would be better suited to validate the predictive glaucomatous potential of the identified NRR phenotypic signature. We plan to address this in our future work. While high-resolution data could be accessed from the OCT scans, it is not commonly done by the clinical protocols. We hope that by adding user-friendly interfaces to CIFU, we may promote such data acquisition and analysis, especially since the functional representation is independent of the number of observations per sample. Indeed, there are several distinct

advantages of our approach which could be built upon further in future studies.

Indeed, straightforward extensions are feasible for similar circular data such as RNFL phenotypes and other optic neuropathies as well as related eye imaging platforms, e.g., OCTA.

5.2 Focused Analysis of RNFL Decay in Glaucomatous Eyes using Circular Statistics on High-resolution OCT data

In this work, we contributed to the following:

The present study addresses multiple important aims, including design of an analytical framework capable of defining and using a directional measure of RNFL decay based on OCT data and then applying it for focused detection of structural characteristics in glaucomatous eyes from an Indian population. In this direction, our study makes several new contributions.

(1) We used a large clinical cohort of normal Indian eyes to generate high-resolution OCT RNFL data (as noted above) for each eye.

(2) Based on the above data for normal eyes stratified by age and optic disc size, we built a catalog of empirical cumulative distribution functions of RNFL loss at each N angular point.

(3) We, developed an analytical framework based on novel applications of circular statistics to OCT data analysis. This includes specifying a directional representation of RNFL decay in a given glaucomatous eye relative to its corresponding normal population. Further, new indices were introduced to measure eye-specific global and local RNFL decay. To illustrate the utility of circular models, we fit a finite mixture of von Mises distributions to identify two clusters of directional representations of RNFL decay in glaucomatous eyes.

(4) Finally, these clusters, learned in an unsupervised manner for three different age-groups, were characterized using their clinical covariates.

While past studies have used spatial, trigonometric and Fourier analysis on OCT data, e.g., our framework, involving directional or circular data, is methodologically quite different from those approaches. While our earlier platform CIFU focuses on the aspects of shapes and curves in OCT NRR data, as in FDA, it does not explicitly address the directional characteristics of OCT data or angular measurements thereof. We have demonstrated the utility of both the granularity and the directionality of OCT data and combined these with an angle-based description of patterns that could be studied with innovative downstream analyses. Indeed,

processes or platforms that generate data capable of being represented as angular values (e.g., wind direction, neuronal activity) require different statistical approaches compared to usual linear data . We think that the new directional representation of OCT data to capture an eye's RNFL decay will pave the way for future applications of the rich methodology of directional and circular statistics to eye data analysis.

References

1.   Harasymowycz, P., et al., *Medical Management of Glaucoma in the 21st Century from a Canadian Perspective.* J Ophthalmol, 2016. **2016**: p. 6509809.
2.   Weinreb, R.N., et al., *Primary open-angle glaucoma.* Nat Rev Dis Primers, 2016. **2**: p. 16067.
3.   Tham, Y.C., et al., *Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis.* Ophthalmology, 2014. **121**(11): p. 2081-90.
4.   Mi, X.S., T.F. Yuan, and K.F. So, *The current research status of normal tension glaucoma.* Clin Interv Aging, 2014. **9**: p. 1563-71.
5.   Mroczkowska, S., et al., *Primary open-angle glaucoma vs normal-tension glaucoma: the vascular perspective.* JAMA Ophthalmol, 2013. **131**(1): p. 36-43.
6.   Wright, C., et al., *Primary angle-closure glaucoma: an update.* Acta Ophthalmol, 2016. **94**(3): p. 217-25.
7.   Quigley, H.A. and A.T. Broman, *The number of people with glaucoma worldwide in 2010 and 2020.* Br J Ophthalmol, 2006. **90**(3): p. 262-7.
8.   Foster, P.J., et al., *The definition and classification of glaucoma in prevalence surveys.* Br J Ophthalmol, 2002. **86**(2): p. 238-42.
9.   Miglior, S., et al., *Results of the European Glaucoma Prevention Study.* Ophthalmology, 2005. **112**(3): p. 366-75.
10.   Rudnicka, A.R., et al., *Variations in primary open-angle glaucoma prevalence by age, gender, and race: a Bayesian meta-analysis.* Invest Ophthalmol Vis Sci, 2006. **47**(10): p. 4254-61.
11.   Tuck, M.W. and R.P. Crick, *The age distribution of primary open angle glaucoma.* Ophthalmic Epidemiol, 1998. **5**(4): p. 173-83.
12.   Llobet, A., X. Gasull, and A. Gual, *Understanding trabecular meshwork physiology: a key to the control of intraocular pressure?* News Physiol Sci, 2003. **18**: p. 205-9.
13.   Diestelhorst, M. and G.K. Kriegistein, *Does aqueous humor secretion decrease with age?* Int Ophthalmol, 1992. **16**(4-5): p. 305-9.
14.   Gabelt, B.T. and P.L. Kaufman, *Changes in aqueous humor dynamics with age and glaucoma.* Prog Retin Eye Res, 2005. **24**(5): p. 612-37.
15.   Nguyen, C.T., et al., *Dietary omega 3 fatty acids decrease intraocular pressure with age by increasing aqueous outflow.* Invest Ophthalmol Vis Sci, 2007. **48**(2): p. 756-62.
16.   Klein, B.E., et al., *Prevalence of glaucoma. The Beaver Dam Eye Study.* Ophthalmology, 1992. **99**(10): p. 1499-504.
17.   Sommer, A., et al., *Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. The Baltimore Eye Survey.* Arch Ophthalmol, 1991. **109**(8): p. 1090-5.
18.   Zhong, Y.S., C.K. Leung, and C.P. Pang, *Glial cells and glaucomatous neuropathy.* Chin Med J (Engl), 2007. **120**(4): p. 326-35.
19.   Streit, W.J., *Microglial senescence: does the brain's immune system have an expiration date?* Trends Neurosci, 2006. **29**(9): p. 506-10.
20.   Johnson, T.V. and K.R. Martin, *Cell transplantation approaches to retinal ganglion cell neuroprotection in glaucoma.* Curr Opin Pharmacol, 2013. **13**(1): p. 78-82.
21.   Boehm, A.G., A.U. Koeller, and L.E. Pillunat, *The effect of age on optic nerve head blood flow.* Invest Ophthalmol Vis Sci, 2005. **46**(4): p. 1291-5.
22.   Racette, L., et al., *Primary open-angle glaucoma in blacks: a review.* Surv Ophthalmol, 2003. **48**(3): p. 295-313.
23.   Kosoko-Lasaki, O., et al., *Race, ethnicity and prevalence of primary open-angle glaucoma.* J Natl Med Assoc, 2006. **98**(10): p. 1626-9.

24. Tielsch, J.M., et al., *Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey.* Jama, 1991. **266**(3): p. 369-74.

25. Dielemans, I., et al., *The prevalence of primary open-angle glaucoma in a population-based study in The Netherlands. The Rotterdam Study.* Ophthalmology, 1994. **101**(11): p. 1851-5.

26. Leske, M.C., et al., *The Barbados Eye Study. Prevalence of open angle glaucoma.* Arch Ophthalmol, 1994. **112**(6): p. 821-9.

27. Mitchell, P., et al., *Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study.* Ophthalmology, 1996. **103**(10): p. 1661-9.

28. Aström, S. and C. Lindén, *Incidence and prevalence of pseudoexfoliation and open-angle glaucoma in northern Sweden: I. Baseline report.* Acta Ophthalmol Scand, 2007. **85**(8): p. 828-31.

29. de Voogd, S., et al., *Incidence of open-angle glaucoma in a general elderly population: the Rotterdam Study.* Ophthalmology, 2005. **112**(9): p. 1487-93.

30. Bengtsson, B., *The prevalence of glaucoma.* Br J Ophthalmol, 1981. **65**(1): p. 46-9.

31. Drance, S., D.R. Anderson, and M. Schulzer, *Risk factors for progression of visual field abnormalities in normal-tension glaucoma.* Am J Ophthalmol, 2001. **131**(6): p. 699-708.

32. Altintaş, O., et al., *The effects of menopause and hormone replacement therapy on quality and quantity of tear, intraocular pressure and ocular blood flow.* Ophthalmologica, 2004. **218**(2): p. 120-9.

33. Sator, M.O., et al., *Hormone replacement therapy and intraocular pressure.* Maturitas, 1997. **28**(1): p. 55-8.

34. Lee, A.J., et al., *Female reproductive factors and open angle glaucoma: the Blue Mountains Eye Study.* Br J Ophthalmol, 2003. **87**(11): p. 1324-8.

35. Ogueta, S.B., et al., *Estrogen receptor in the human eye: influence of gender and age on gene expression.* Invest Ophthalmol Vis Sci, 1999. **40**(9): p. 1906-11.

36. Hulsman, C.A., et al., *Is open-angle glaucoma associated with early menopause? The Rotterdam Study.* Am J Epidemiol, 2001. **154**(2): p. 138-44.

37. Nirmalan, P.K., et al., *Female reproductive factors and eye disease in a rural South Indian population: the Aravind Comprehensive Eye Survey.* Invest Ophthalmol Vis Sci, 2004. **45**(12): p. 4273-6.

38. Bonovas, S., V. Peponis, and K. Filioussi, *Diabetes mellitus as a risk factor for primary open-angle glaucoma: a meta-analysis.* Diabet Med, 2004. **21**(6): p. 609-14.

39. Dielemans, I., et al., *Primary open-angle glaucoma, intraocular pressure, and diabetes mellitus in the general elderly population. The Rotterdam Study.* Ophthalmology, 1996. **103**(8): p. 1271-5.

40. Klein, B.E., R. Klein, and S.C. Jensen, *Open-angle glaucoma and older-onset diabetes. The Beaver Dam Eye Study.* Ophthalmology, 1994. **101**(7): p. 1173-7.

41. Dielemans, I., et al., *Primary open-angle glaucoma, intraocular pressure, and systemic blood pressure in the general elderly population. The Rotterdam Study.* Ophthalmology, 1995. **102**(1): p. 54-60.

42. Tielsch, J.M., et al., *Diabetes, intraocular pressure, and primary open-angle glaucoma in the Baltimore Eye Survey.* Ophthalmology, 1995. **102**(1): p. 48-53.

43. Lee, A.J., et al., *Open-angle glaucoma and cardiovascular mortality: the Blue Mountains Eye Study.* Ophthalmology, 2006. **113**(7): p. 1069-76.

44. Lee, A.J., et al., *Open-angle glaucoma and systemic thyroid disease in an older population: The Blue Mountains Eye Study.* Eye (Lond), 2004. **18**(6): p. 600-8.

45. Corbett, J.J., et al., *The neurologic evaluation of patients with low-tension glaucoma.* Invest Ophthalmol Vis Sci, 1985. **26**(8): p. 1101-4.

46. Wang, J.J., P. Mitchell, and W. Smith, *Is there an association between migraine headache and open-angle glaucoma? Findings from the Blue Mountains Eye Study.* Ophthalmology, 1997. **104**(10): p. 1714-9.

47. Pache, M. and J. Flammer, *A sick eye in a sick body? Systemic findings in patients with primary open-angle glaucoma.* Surv Ophthalmol, 2006. **51**(3): p. 179-212.

48. Schmidl, D., G. Garhofer, and L. Schmetterer, *The complex interaction between ocular perfusion pressure and ocular blood flow - relevance for glaucoma.* Exp Eye Res, 2011. **93**(2): p. 141-55.

49. Klein, B.E., R. Klein, and K.L. Linton, *Intraocular pressure in an American community. The Beaver Dam Eye Study.* Invest Ophthalmol Vis Sci, 1992. **33**(7): p. 2224-8.

50. Feiner, L. and J.R. Piltz-Seymour, *Collaborative Initial Glaucoma Treatment Study: a summary of results to date.* Curr Opin Ophthalmol, 2003. **14**(2): p. 106-11.

51. Heijl, A., et al., *Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial.* Arch Ophthalmol, 2002. **120**(10): p. 1268-79.

52. Guo, L., et al., *Retinal ganglion cell apoptosis in glaucoma is related to intraocular pressure and IOP-induced effects on extracellular matrix.* Invest Ophthalmol Vis Sci, 2005. **46**(1): p. 175-82.

53. Alsbirk, P.H., *Corneal thickness. I. Age variation, sex difference and oculometric correlations.* Acta Ophthalmol (Copenh), 1978. **56**(1): p. 95-104.

54. Medeiros, F.A., et al., *Fourier analysis of scanning laser polarimetry measurements with variable corneal compensation in glaucoma.* Invest Ophthalmol Vis Sci, 2003. **44**(6): p. 2606-12.

55. Dueker, D.K., et al., *Corneal thickness measurement in the management of primary open-angle glaucoma: a report by the American Academy of Ophthalmology.* Ophthalmology, 2007. **114**(9): p. 1779-87.

56. Bechmann, M., et al., *Central corneal thickness determined with optical coherence tomography in various types of glaucoma.* Br J Ophthalmol, 2000. **84**(11): p. 1233-7.

57. Congdon, N., F. Wang, and J.M. Tielsch, *Issues in the epidemiology and population-based screening of primary angle-closure glaucoma.* Surv Ophthalmol, 1992. **36**(6): p. 411-23.

58. Foster, P.J., et al., *Determinants of intraocular pressure and its association with glaucomatous optic neuropathy in Chinese Singaporeans: the Tanjong Pagar Study.* Invest Ophthalmol Vis Sci, 2003. **44**(9): p. 3885-91.

59. Congdon, N.G., et al., *Central corneal thickness and corneal hysteresis associated with glaucoma damage.* Am J Ophthalmol, 2006. **141**(5): p. 868-75.

60. Daubs, J.G. and R.P. Crick, *Effect of refractive error on the risk of ocular hypertension and open angle glaucoma.* Trans Ophthalmol Soc U K (1962), 1981. **101**(1): p. 121-6.

61. Leske, M.C., et al., *Factors for glaucoma progression and the effect of treatment: the early manifest glaucoma trial.* Arch Ophthalmol, 2003. **121**(1): p. 48-56.

62. Chihara, E., et al., *Severe myopia as a risk factor for progressive visual field loss in primary open-angle glaucoma.* Ophthalmologica, 1997. **211**(2): p. 66-71.

63. Jonas, J.B., et al., *Predictive factors of the optic nerve head for development or progression of glaucomatous visual field loss.* Invest Ophthalmol Vis Sci, 2004. **45**(8): p. 2613-8.

64. McNaught, A.I., et al., *Accuracy and implications of a reported family history of glaucoma: experience from the Glaucoma Inheritance Study in Tasmania.* Arch Ophthalmol, 2000. **118**(7): p. 900-4.

65. Wolfs, R.C., et al., *Changing views on open-angle glaucoma: definitions and prevalences--The Rotterdam Study.* Invest Ophthalmol Vis Sci, 2000. **41**(11): p. 3309-21.

66. Wadhwa, S.D. and E.J. Higginbotham, *Ethnic differences in glaucoma: prevalence, management, and outcome.* Curr Opin Ophthalmol, 2005. **16**(2): p. 101-6.

67. Teikari, J.M. and J.P. Airaksinen, *Twin study on cup/disc ratio of the optic nerve head.* Br J Ophthalmol, 1992. **76**(4): p. 218-20.

68. Rosenthal, A.R. and E.S. Perkins, *Family studies in glaucoma.* Br J Ophthalmol, 1985. **69**(9): p. 664-7.

69.   Harper, R.A. and B.C. Reeves, *Glaucoma screening: the importance of combining test data.* Optom Vis Sci, 1999. **76**(8): p. 537-43.

70.   Parikh, R., et al., *Understanding and using sensitivity, specificity and predictive values.* Indian J Ophthalmol, 2008. **56**(1): p. 45-50.

71.   Keltner, J.L. and C.A. Johnson, *Screening for visual field abnormalities with automated perimetry.* Surv Ophthalmol, 1983. **28**(3): p. 175-83.

72.   Hitchings, R.A., D.B. Brown, and S.A. Anderton, *Glaucoma screening by means of an optic disc grid.* Br J Ophthalmol, 1983. **67**(6): p. 352-5.

73.   Tan, J.C., D.F. Garway-Heath, and R.A. Hitchings, *Variability across the optic nerve head in scanning laser tomography.* Br J Ophthalmol, 2003. **87**(5): p. 557-9.

74.   Gordon, M.O., et al., *The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma.* Arch Ophthalmol, 2002. **120**(6): p. 714-20; discussion 829-30.

75.   Rochtchina, E., P. Mitchell, and J.J. Wang, *Relationship between age and intraocular pressure: the Blue Mountains Eye Study.* Clin Exp Ophthalmol, 2002. **30**(3): p. 173-5.

76.   Whitacre, M.M. and R. Stein, *Sources of error with use of Goldmann-type tonometers.* Surv Ophthalmol, 1993. **38**(1): p. 1-30.

77.   Brandt, J.D., et al., *Central corneal thickness and measured IOP response to topical ocular hypotensive medication in the Ocular Hypertension Treatment Study.* Am J Ophthalmol, 2004. **138**(5): p. 717-22.

78.   Anderson, D.R. and V.M. Patella, *Automated Static Perimetry*. 1999: Mosby.

79.   Spry, P.G. and C.A. Johnson, *Identification of progressive glaucomatous visual field loss.* Surv Ophthalmol, 2002. **47**(2): p. 158-73.

80.   Quigley, H.A., *Neuronal death in glaucoma.* Prog Retin Eye Res, 1999. **18**(1): p. 39-57.

81.   Tuulonen, A. and P.J. Airaksinen, *Initial glaucomatous optic disk and retinal nerve fiber layer abnormalities and their progression.* Am J Ophthalmol, 1991. **111**(4): p. 485-90.

82.   Jonas, J.B. and A. Dichtl, *Evaluation of the retinal nerve fiber layer.* Surv Ophthalmol, 1996. **40**(5): p. 369-78.

83.   Medeiros, F.A., et al., *Comparison of the GDx VCC scanning laser polarimeter, HRT II confocal scanning laser ophthalmoscope, and stratus OCT optical coherence tomograph for the detection of glaucoma.* Arch Ophthalmol, 2004. **122**(6): p. 827-37.

84.   Ferreras, A., et al., *Diagnostic ability of Heidelberg Retina Tomograph 3 classifications: glaucoma probability score versus Moorfields regression analysis.* Ophthalmology, 2007. **114**(11): p. 1981-7.

85.   Garway-Heath, D.F., et al., *Vertical cup/disc ratio in relation to optic disc size: its value in the assessment of the glaucoma suspect.* Br J Ophthalmol, 1998. **82**(10): p. 1118-24.

86.   Jonas, J.B., W.M. Budde, and P. Lang, *Neuroretinal rim width ratios in morphological glaucoma diagnosis.* Br J Ophthalmol, 1998. **82**(12): p. 1366-71.

87.   Jonas, J.B., G.C. Gusek, and G.O. Naumann, *Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes.* Invest Ophthalmol Vis Sci, 1988. **29**(7): p. 1151-8.

88.   Fitzgibbon, T. and S.F. Taylor, *Retinotopy of the human retinal nerve fibre layer and optic nerve head.* J Comp Neurol, 1996. **375**(2): p. 238-51.

89.   Jonas, J.B., N.X. Nguyen, and G.O. Naumann, *The retinal nerve fiber layer in normal eyes.* Ophthalmology, 1989. **96**(5): p. 627-32.

90.   Balazsi, A.G., et al., *The effect of age on the nerve fiber population of the human optic nerve.* Am J Ophthalmol, 1984. **97**(6): p. 760-6.

91.   Dolman, C.L., A.Q. McCormick, and S.M. Drance, *Aging of the optic nerve.* Arch Ophthalmol, 1980. **98**(11): p. 2053-8.

92.   Jonas, J.B., et al., *Parapapillary chorioretinal atrophy in normal and glaucoma eyes. I. Morphometric data.* Invest Ophthalmol Vis Sci, 1989. **30**(5): p. 908-18.

93. Jonas, J.B., *Clinical implications of peripapillary atrophy in glaucoma.* Curr Opin Ophthalmol, 2005. **16**(2): p. 84-8.

94. Jonas, J.B., K.A. Königsreuther, and G.O. Naumann, *Optic disc histomorphometry in normal eyes and eyes with secondary angle-closure glaucoma. II. Parapapillary region.* Graefes Arch Clin Exp Ophthalmol, 1992. **230**(2): p. 134-9.

95. Kubota, T., J.B. Jonas, and G.O. Naumann, *Direct clinico-histological correlation of parapapillary chorioretinal atrophy.* Br J Ophthalmol, 1993. **77**(2): p. 103-6.

96. Teng, C.C., et al., *Beta-Zone parapapillary atrophy and the velocity of glaucoma progression.* Ophthalmology, 2010. **117**(5): p. 909-15.

97. Uchida, H., S. Ugurlu, and J. Caprioli, *Increasing peripapillary atrophy is associated with progressive glaucoma.* Ophthalmology, 1998. **105**(8): p. 1541-5.

98. Savatovsky, E., et al., *Longitudinal changes in peripapillary atrophy in the ocular hypertension treatment study: a case-control assessment.* Ophthalmology, 2015. **122**(1): p. 79-86.

99. Hendrickx, K.H., et al., *Cumulative incidence of patients with disc hemorrhages in glaucoma and the effect of therapy.* Ophthalmology, 1994. **101**(7): p. 1165-72.

100. Jonas, J.B. and L. Xu, *Optic disk hemorrhages in glaucoma.* Am J Ophthalmol, 1994. **118**(1): p. 1-8.

101. Kitazawa, Y., S. Shirato, and T. Yamamoto, *Optic disc hemorrhage in low-tension glaucoma.* Ophthalmology, 1986. **93**(6): p. 853-7.

102. Fercher, A.F., et al., *Optical coherence tomography - principles and applications.* Reports on Progress in Physics, 2003. **66**(2): p. 239.

103. Huang, D., et al., *Optical coherence tomography.* Science, 1991. **254**(5035): p. 1178-81.

104. Leitgeb, R., et al., *Spectral measurement of absorption by spectroscopic frequency-domain optical coherence tomography.* Opt Lett, 2000. **25**(11): p. 820-2.

105. de Boer, J.F., et al., *Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography.* Opt Lett, 2003. **28**(21): p. 2067-9.

106. Drexler, W. and J.G. Fujimoto, *Optical Coherence Tomography: Technology and Applications.* 2008: Springer Berlin Heidelberg.

107. Dong, Z.M., G. Wollstein, and J.S. Schuman, *Clinical Utility of Optical Coherence Tomography in Glaucoma.* Invest Ophthalmol Vis Sci, 2016. **57**(9): p. Oct556-67.

108. Hood, D.C. and R.H. Kardon, *A framework for comparing structural and functional measures of glaucomatous damage.* Prog Retin Eye Res, 2007. **26**(6): p. 688-710.

109. Quigley, H.A., E.M. Addicks, and W.R. Green, *Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy.* Arch Ophthalmol, 1982. **100**(1): p. 135-46.

110. Kamal, D.S., et al., *Detection of optic disc change with the Heidelberg retina tomograph before confirmed visual field change in ocular hypertensives converting to early glaucoma.* Br J Ophthalmol, 1999. **83**(3): p. 290-4.

111. Sehi, M. and D.S. Greenfield, *Assessment of retinal nerve fiber layer using optical coherence tomography and scanning laser polarimetry in progressive glaucomatous optic neuropathy.* Am J Ophthalmol, 2006. **142**(6): p. 1056-9.

112. Wollstein, G., et al., *Optical coherence tomography longitudinal evaluation of retinal nerve fiber layer thickness in glaucoma.* Arch Ophthalmol, 2005. **123**(4): p. 464-70.

113. Kuang, T.M., et al., *Estimating Lead Time Gained by Optical Coherence Tomography in Detecting Glaucoma before Development of Visual Field Defects.* Ophthalmology, 2015. **122**(10): p. 2002-9.

114. Weinreb, R.N., et al., *Diagnosis of Primary Open Angle Glaucoma: WGA consensus series - 10.* 2017: Kugler Publications.

115. Quigley, H.A., et al., *An evaluation of optic disc and nerve fiber layer examinations in monitoring progression of early glaucoma damage.* Ophthalmology, 1992. **99**(1): p. 19-28.

116.   Leung, C.K., *Diagnosing glaucoma progression with optical coherence tomography.* Curr Opin Ophthalmol, 2014. **25**(2): p. 104-11.

117.   Leung, C.K., et al., *Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: analysis of the retinal nerve fiber layer map for glaucoma detection.* Ophthalmology, 2010. **117**(9): p. 1684-91.

118.   Leung, C.K., et al., *Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: patterns of retinal nerve fiber layer progression.* Ophthalmology, 2012. **119**(9): p. 1858-66.

119.   Schuman, J.S., et al., *Reproducibility of nerve fiber layer thickness measurements using optical coherence tomography.* Ophthalmology, 1996. **103**(11): p. 1889-98.

120.   *Zeiss Cirrus HD-OCT User Manual – Models 500, 5000 Instrument and Review Software 8.1, 2015*. 2015.

121.   Almobarak, F.A., et al., *Automated segmentation of optic nerve head structures with optical coherence tomography.* Invest Ophthalmol Vis Sci, 2014. **55**(2): p. 1161-8.

122.   Reis, A.S., et al., *Optic disc margin anatomy in patients with glaucoma and normal controls with spectral domain optical coherence tomography.* Ophthalmology, 2012. **119**(4): p. 738-47.

123.   Chauhan, B.C., et al., *Bruch's Membrane Opening Minimum Rim Width and Retinal Nerve Fiber Layer Thickness in a Normal White Population: A Multicenter Study.* Ophthalmology, 2015. **122**(9): p. 1786-94.

124.   Chauhan, B.C., et al., *Enhanced detection of open-angle glaucoma with an anatomically accurate optical coherence tomography-derived neuroretinal rim parameter.* Ophthalmology, 2013. **120**(3): p. 535-543.

125.   Pollet-Villard, F., et al., *Structure-function relationships with spectral-domain optical coherence tomography retinal nerve fiber layer and optic nerve head measurements.* Invest Ophthalmol Vis Sci, 2014. **55**(5): p. 2953-62.

126.   Kasumovic, S.S., et al., *Correlation between retinal nerve fiber layer and disc parameters in glaucoma suspected eyes.* Med Arch, 2014. **68**(2): p. 113-6.

127.   Sung, K.R., J.H. Na, and Y. Lee, *Glaucoma diagnostic capabilities of optic nerve head parameters as determined by Cirrus HD optical coherence tomography.* J Glaucoma, 2012. **21**(7): p. 498-504.

128.   Louis-Paul Rivest, T.D., Aurélien Nicosia, Daniel Fortin, *A general angular regression model for the analysis of data on animal movement in ecology.* Journal of the Royal Statistical Society Series C, 2015. **65**(3).

129.   Mardia, K.V., C.C. Taylor, and G.K. Subramaniam, *Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data.* Biometrics, 2007. **63**(2): p. 505-12.

130.   Rutishauser, U., et al., *Human memory strength is predicted by theta-frequency phase-locking of single neurons.* Nature, 2010. **464**(7290): p. 903-7.

131.   Ramsay, J.O., *When the data are functions.* Psychometrika, 1982. **47**(4): p. 379-396.

132.   Ramsay, J.O. and C.J. Dalzell, *Some Tools for Functional Data Analysis.* Journal of the Royal Statistical Society. Series B (Methodological), 1991. **53**(3): p. 539-572.

133.   Grenander, U., *Stochastic processes and statistical inference.* Arkiv för Matematik, 1950. **1**(3): p. 195-277.

134.   Rao, C.R., *Some Statistical Methods for Comparison of Growth Curves.* Biometrics, 1958. **14**(1): p. 1-17.

135.   Duhamel, P. and M. Vetterli, *Fast fourier transforms: A tutorial review and a state of the art.* Signal Processing, 1990. **19**(4): p. 259-299.

136.   Ramsay, J. and B.W. Silverman, *Functional Data Analysis*. 2006: Springer New York.

137.   de Boor, C., *A Practical Guide to Splines*. 1978: Springer New York.

138.   Fugal, D.L., *Conceptual Wavelets in Digital Signal Processing: An In-depth, Practical Approach for the Non-mathematician*. 2009: Space & Signals Technical Pub.

139. World Health, O., *Global Initiative for the Elimination of Avoidable Blindness : action plan 2006-2011*. 2007, World Health Organization: Geneva.

140. Kerrigan-Baumrind, L.A., et al., *Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons.* Invest Ophthalmol Vis Sci, 2000. **41**(3): p. 741-8.

141. Garway-Heath, D.F., G. Wollstein, and R.A. Hitchings, *Aging changes of the optic nerve head in relation to open angle glaucoma.* Br J Ophthalmol, 1997. **81**(10): p. 840-5.

142. Prata, T.S., et al., *Eyes with large disc cupping and normal intraocular pressure: using optical coherence tomography to discriminate those with and without glaucoma.* Med Hypothesis Discov Innov Ophthalmol, 2014. **3**(3): p. 91-8.

143. Kwon, Y.H., et al., *Primary open-angle glaucoma.* N Engl J Med, 2009. **360**(11): p. 1113-24.

144. Mwanza, J.C., et al., *Reproducibility of peripapillary retinal nerve fiber layer thickness and optic nerve head parameters measured with cirrus HD-OCT in glaucomatous eyes.* Invest Ophthalmol Vis Sci, 2010. **51**(11): p. 5724-30.

145. Mwanza, J.C., et al., *Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes.* Ophthalmology, 2011. **118**(2): p. 241-8.e1.

146. Jammalamadaka, S.R. and A. Sengupta, *Topics In Circular Statistics-vol 5*. 2001: World Scientific Publishing Company.

147. Ieva, F., et al., *Multivariate functional clustering for the morphological analysis of electrocardiograph curves.* Journal of the Royal Statistical Society: Series C (Applied Statistics), 2013. **62**.

148. Jacques, J. and C. Preda, *Funclust: A curves clustering method using functional random variables density approximation.* Neurocomputing, 2013. **112**: p. 164-171.

149. Wang, J.-L., J.-M. Chiou, and H.-G. Müller, *Functional Data Analysis.* Annual Review of Statistics and Its Application, 2016. **3**(1): p. 257-295.

150. Matsushita, K., M.L. Puri, and T. Hayakawa, *Statistical Sciences and Data Analysis: Proceedings of the Third Pacific Area Statistical Conference*. 1993: VSP.

151. Thylefors, B. and A.D. Négrel, *The global impact of glaucoma.* Bull World Health Organ, 1994. **72**(3): p. 323-6.

152. Addepalli, U.K., et al., *LV Prasad Eye Institute Glaucoma Epidemiology and Molecular Genetic Study (LVPEI- GLEAMS). Report 1: study design and research methodology.* Ophthalmic Epidemiol, 2013. **20**(3): p. 188-95.

153. Rao, H.L., et al., *Effect of scan quality on diagnostic accuracy of spectral-domain optical coherence tomography in glaucoma.* Am J Ophthalmol, 2014. **157**(3): p. 719-27.e1.

154. Samarawickrama, C., et al., *Influence of OCT signal strength on macular, optic nerve head, and retinal nerve fiber layer parameters.* Invest Ophthalmol Vis Sci, 2010. **51**(9): p. 4471-5.

155. Wu, Z., et al., *Signal strength is an important determinant of accuracy of nerve fiber layer thickness measurement by optical coherence tomography.* J Glaucoma, 2009. **18**(3): p. 213-6.

156. Asrani, S., et al., *Artifacts in spectral-domain optical coherence tomography measurements in glaucoma.* JAMA Ophthalmol, 2014. **132**(4): p. 396-402.

157. Bouveyron, C., E. Côme, and J. Jacques, *The discriminative functional mixture model for a comparative analysis of bike sharing systems.* arXiv: Applications, 2016.

158. Bouveyron, C. and J. Jacques. *funFEM: an R package for functional data clustering*. 2015.

159. McLachlan, G.J. and T. Krishnan, *The EM Algorithm and Extensions*. 2008: Wiley.

160. Fraley, C. and A.E. Raftery, *MCLUST: Software for Model-Based Cluster Analysis.* Journal of Classification, 1999. **16**(2): p. 297-306.

161. Akaike, H., *A new look at the statistical model identification.* IEEE Transactions on Automatic Control, 1974. **19**(6): p. 716-723.

162. Biernacki, C., G. Celeux, and G. Govaert, *Assessing a mixture model for clustering with the integrated completed likelihood.* IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Computer Society, 2000. **22**(7): p. 719-725.

163. Manchester, L. and W. Blanchard, *When is a curve an outlier? An account of a tricky problem.* Canadian Journal of Statistics, 2008. **24**: p. 455-466.

164. Kaufman, L. and P. Rousseeuw, *Finding Groups in Data: An Introduction To Cluster Analysis*. 1990.

165. Sekula, M., S. Datta, and S. Datta, *optCluster: An R Package for Determining the Optimal Clustering Algorithm.* Bioinformation, 2017. **13**(3): p. 101-103.

166. Kassambara, A. and F. Mundt. *Extract and Visualize the Results of Multivariate Data Analyses [R package factoextra version 1.0.7]*. 2020.

167. Witten, D.M. and R. Tibshirani, *A framework for feature selection in clustering.* J Am Stat Assoc, 2010. **105**(490): p. 713-726.

168. Tao, Y., et al., *Profile of retinal nerve fibre layer symmetry in a multiethnic Asian population: the Singapore Epidemiology of Eye Diseases study.* Br J Ophthalmol, 2020. **104**(6): p. 836-841.

169. Bowd, C., et al., *Racial Differences in Rate of Change of Spectral-Domain Optical Coherence Tomography-Measured Minimum Rim Width and Retinal Nerve Fiber Layer Thickness.* Am J Ophthalmol, 2018. **196**: p. 154-164.

170. Knight, O.J., et al., *Effect of race, age, and axial length on optic nerve head parameters and retinal nerve fiber layer thickness measured by Cirrus HD-OCT.* Arch Ophthalmol, 2012. **130**(3): p. 312-8.

171. Rao, H.L., G.J. Babu, and G.C. Sekhar, *Comparison of the diagnostic capability of the Heidelberg Retina Tomographs 2 and 3 for glaucoma in the Indian population.* Ophthalmology, 2010. **117**(2): p. 275-81.

172. Hammel, N., et al., *Rate and Pattern of Rim Area Loss in Healthy and Progressing Glaucoma Eyes.* Ophthalmology, 2016. **123**(4): p. 760-70.

173. Siesky, B., et al., *Baseline structural characteristics of the optic nerve head and retinal nerve fiber layer are associated with progressive visual field loss in patients with open-angle glaucoma.* PLoS One, 2020. **15**(8): p. e0236819.

174. Reis, A.S., et al., *Influence of clinically invisible, but optical coherence tomography detected, optic disc margin anatomy on neuroretinal rim evaluation.* Invest Ophthalmol Vis Sci, 2012. **53**(4): p. 1852-60.

175. Garway-Heath, D.F. and R.A. Hitchings, *Quantitative evaluation of the optic nerve head in early glaucoma.* Br J Ophthalmol, 1998. **82**(4): p. 352-61.

176. Dryden, I.L. and K.V. Mardia, *Statistical Shape Analysis: With Applications in R*. 2016: Wiley.

177. Abraham, C., N. Molinari, and R. Servien, *Unsupervised clustering of multivariate circular data.* Stat Med, 2013. **32**(8): p. 1376-82.

178. Rodríguez, C.E., G. Núñez-Antonio, and G. Escarela, *A Bayesian mixture model for clustering circular data.* Computational Statistics & Data Analysis, 2020. **143**: p. 106842.

179. Gaffney, S. and P. Smyth, *Joint Probabilistic Curve Clustering and Alignment*. 2004.

180. Heard, N., C. Holmes, and D. Stephens, *A Quantitative Study of Gene Regulation Involved in the Immune Response of Anopheline Mosquitoes.* Journal of The American Statistical Association , 2006. **101**: p. 18-29.

181. Ray, S. and S. Pyne, *A computational framework to emulate the human perspective in flow cytometric data analysis.* PLoS One, 2012. **7**(5): p. e35693.

182. G.M, J. and S. C.A, *Clustering for Sparsely Sampled Functional Data.* Journal of the American Statistical Association, 2003. **98**: p. 397-408.

183. Jones, P.N. and G. McLachlan, *Fitting finite mixture models in a regression context.* Australian Journal of Statistics, 2008. **34**: p. 233-240.

184. Ray, S. and B. Mallick, *Functional Clustering by Bayesian Wavelet Methods.* Journal of the Royal Statistical Society Series B, 2006. **68**: p. 305-332.

185. Hwang, Y.H. and Y.Y. Kim, *Glaucoma diagnostic ability of quadrant and clock-hour neuroretinal rim assessment using cirrus HD optical coherence tomography.* Invest Ophthalmol Vis Sci, 2012. **53**(4): p. 2226-34.

186. Pyne, S., et al., *Automated high-dimensional flow cytometric data analysis.* Proc Natl Acad Sci U S A, 2009. **106**(21): p. 8519-24.

187. Ho, H.J., S. Pyne, and T.I. Lin, *Maximum likelihood inference for mixtures of skew Student-t-normal distributions through practical EM-type algorithms.* Statistics and Computing, 2012. **22**(1): p. 287-299.

188. Pyne, S., et al., *Joint modeling and registration of cell populations in cohorts of high-dimensional flow cytometric data.* PLoS One, 2014. **9**(7): p. e100334.

189. Qi, Y., et al., *High-speed automatic characterization of rare events in flow cytometric data.* PLoS One, 2020. **15**(2): p. e0228651.

190. Dimitriou, C. and D. Broadway, *Pathophysiology of Glaucoma*. 2013. p. 32-57.

191. Leung, C.K., et al., *Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study.* Ophthalmology, 2009. **116**(7): p. 1257-63, 1263.e1-2.

192. Cremers, J. and I. Klugkist, *One Direction? A Tutorial for Circular Data Analysis Using R With Examples in Cognitive Psychology.* Front Psychol, 2018. **9**: p. 2040.

193. Mardia, K.V. and P.E. Jupp, *Directional Statistics*. 2000: Wiley.

194. Ali, M.H., et al., *Circular functional analysis of OCT data for precise identification of structural phenotypes in the eye.* Sci Rep, 2021. **11**(1): p. 23336.

195. Bowd, C., et al., *Estimating Optical Coherence Tomography Structural Measurement Floors to Improve Detection of Progression in Advanced Glaucoma.* Am J Ophthalmol, 2017. **175**: p. 37-44.

196. Rao, H.B., et al., *Clinical measurement and categorization of optic disc in glaucoma patients.* Indian J Ophthalmol, 2009. **57**(5): p. 361-4.

197. Hothorn, T., K. Hornik, and A. Zeileis, *Unbiased Recursive Partitioning: A Conditional Inference Framework.* Journal of Computational and Graphical Statistics, 2006. **15**: p. 651-674.

198. Jonas, J.B. and D. Schiro, *Localised wedge shaped defects of the retinal nerve fibre layer in glaucoma.* Br J Ophthalmol, 1994. **78**(4): p. 285-90.

199. Zhu, W., et al., *Spatial and spatio-temporal statistical analyses of retinal images: a review of methods and applications.* BMJ Open Ophthalmol, 2020. **5**(1): p. e000479.

200. Pyne, S., et al., *Quantification of retinal nerve fiber layer changes in optical coherence tomography images reveals differential progression of glaucomas.* BLDE University Journal of Health Sciences, 2019. **4**: p. 85.

201. Essock, E.A., et al., *Fourier analysis of optical coherence tomography and scanning laser polarimetry retinal nerve fiber layer measurements in the diagnosis of glaucoma.* Arch Ophthalmol, 2003. **121**(9): p. 1238-45.

202. Hsieh, M.H., et al., *Fourier analysis of circumpapillary retinal nerve fiber layer thickness in optical coherence tomography for differentiating myopia and glaucoma.* Sci Rep, 2020. **10**(1): p. 10509.

List of Publications and Presentations

1. **Md. Hasnat Ali,** Brian Wainwright, Alexander Petersen, Ganesh B. Jonnadula, Meghana Aruru, Harsha L. Rao, M. B. Srinivas, S. Rao Jammalamadaka, Sirisha Senthil, and Saumyadipta Pyne. *"Circular functional analysis of OCT data for precise identification of structural phenotypes in the eye"* Sci Rep. 2021 Dec 2;11(1):23336. doi: 10.1038/s41598-021-02025-4. Article number: 23336 (2021)
   Publisher: Scientific Reports – Nature
   (https://www.nature.com/articles/s41598-021-02025-4)

2. **Md. Hasnat Ali**, Meghana Desai, S. Rao Jammalamadaka, Sirisha Senthil, M. B. Srinivas, Saumyadipta Pyne. *"Focused Analysis of RNFL Decay in Glaucomatous Eyes using Circular Statistics on High-resolution OCT data"* PLoS One 2023 Oct 18;18(10):e0292915. doi: 10.1371/journal.pone.0292915. eCollection 2023
   Publisher: PLoS One
   (https://doi.org/10.1371/journal.pone.0292915)

Conference

Md. Hasnat Ali, Jisy N.K, Sirisha Senthil, Saumyadipta Pyne, M. B. Srinivas. *"Automated Glaucoma Classification In High-Resolution Optic Disc Photographs Using Computer Vision Techniques and Random Forest Classifier."* ARVO 2023 Annual Meeting April 23–27 | New Orleans, La. USA.

Brief Biography

Candidate Md Hasnat Ali

Md Hasnat Ali is a Ph.D. student at Birla Institute of Technology and Science (BITS) University in Pilani's Department of Electrical and Electronics Engineering. He is currently the senior bio-statistician at L V Prasad Eye Institute. He has been an integral part of the institute's faculty since 2010, bringing his expertise and knowledge to advance research and statistical analysis in the field of ophthalmology. He has demonstrated exceptional proficiency in designing robust statistical models, analyzing complex data sets, and interpreting results precisely. He received his MBA in Public Health Informatics from Jamia Hamdard, university, New Delhi. His research interests include circular and functional data analysis, computational statistics and machine learning, predictive modeling, and Big data analysis. He has published in over 150 international publications as an author or co-author.

Supervisor Dr. M.B. Srinivas

Dr. M.B. Srinivas received his Ph.D from the Indian Institute of Science, Bangalore, India. He is currently a Professor in the Department of Electrical and Electronics Engineering at Birla Institute of Technology and Science (BITS) Pilani. His research interests include In-Memory Computing, Memory Technologies, VLSI arithmetic and Deep Learning for Fundus Image Processing. He is an active member of IEEE and has served as Chairman of IEEE Hyderabad Section during 2007 and 2008 and founding Chairman of CAS/EDS Joint Chapter during 2012 and 2013. He is a recipient of Microsoft Research 'Digital Inclusion' award in 2006 and Stanford Medicine 'MedTech Innovation' award in 2016.

Co-supervisor Dr. Saumyadipta Pyne

Dr. Pyne's areas of research interest are computational statistics, machine learning and health data science. He received his Ph.D. in Computer Science from the State University of New York (SUNY) at Stony Brook and postdoctoral training at the Broad Institute of MIT and Harvard University. He held different positions at premier institutions such as University of California Santa Barbara, University of Pittsburgh, Harvard Medical School, Indian Statistical Institute, Indian Institute of Public Health, and National Institute of Medical Statistics, New Delhi. He has served as the PC Mahalanobis Chair and Head of Bioinformatics, Full Professor, Research Scientist, Scientific Director, Ramalingaswami Fellow, Senior Research Fellow and Adjunct Professor. Presently, Dr. Pyne is a Senior Research Fellow and National Service DATA Scholar at the United States National Institutes of Health (NIH) and Adjunct Full Professor at the Department of Statistics and Applied Probability at University of California Santa Barbara.

Co-supervisor Dr. Sirisha Senthil

Dr. Sirisha Senthil is a highly accomplished Ophthalmologist with over 22 years of experience in the field of Glaucoma. She holds an MS in Ophthalmology from Aravind Eye Hospital, Madurai, and an FRCS in Ophthalmology from Edinburgh. She is currently the Head of Glaucoma service at L V Prasad Eye Institute, where she has been a faculty member since 2007. Dr. Senthil specializes in managing refractory adult and pediatric glaucomas, with a particular interest in glaucoma drainage implants and secondary glaucomas and genetics of inherited childhood glaucoma. She has published over 169 scientific papers in peer-reviewed journals, authored several book chapters and received multiple awards for her work, including Dr.Vengal Rao Medal in 2012 and the American Academy of Ophthalmology Achievement award in 2018.