New approaches to the photopic negative response of the electroretinogram in glaucoma

---

Marc George Sarossy

MBBS (Hons) FRANZCO

ORCID: 0000-0002-6117-6662

A thesis submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

May 2022

Centre for Eye Research Australia
The Royal Victorian Eye and Ear Hospital
Department of Ophthalmology
University of Melbourne

Scientific Abstract

Background: Glaucoma is an optic neuropathy characterised by the progressive loss of retinal ganglion cells (RGCs) and is a leading cause of blindness worldwide. Prior to irreversible cell death, RGCs exhibit functional changes that if detectable, could aid in predicting the disease trajectory and help better guide the management of this condition. The photopic negative response (PhNR) of the electroretinogram (ERG) shows potential as such a functional test, but its utility has been limited due to problems with repeatability and interpretation.

Purpose: To explore novel techniques to improve the repeatability of the electroretinogram in glaucoma, describe new feature extraction techniques, and to develop models that predict glaucoma severity from the ERG.

Methods: Normal controls and individuals with glaucoma were recruited for these studies. ERGs were collected with red flashes on a blue background or with a white-on-white stimulus. Various novel denoising techniques were evaluated on the normal control group to improve measurement repeatability, and complexity measures were used as new features to better discriminate between those with and without glaucoma. Novel predictive models informed by novel features extracted with time-frequency techniques were trained and evaluated, and compared to models using various amplitude markers from the ERG.

Results: The denoising technique termed complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was shown to be effective and efficient in removing baseline drift and wander of the ERG, thus significantly improving the repeatability of the recordings. Eyes with and without glaucoma showed significant differences based on novel complexity measures of the ERG, beyond the standard PhNR measure. New models informed by time-frequency features significantly added to the predictive performance for glaucoma severity when used alongside conventional amplitude markers of the ERG.

Conclusion: This body of work developed a new method of denoising of the ERG, which also could have wider application in electrophysiology. New features based on complexity showed potential for providing important additional measures of RGC dysfunction useful for disease discrimination. Time frequency features extracted from the ERG also showed that there is substantial information in the ERG beyond what is captured by standard amplitude markers that could aid in better detecting early RGC dysfunction.

Declaration

This is to certify that this thesis comprises of my original work, except as follows:

Dr Jessica Tang assisted in the collection of the ERG data used in Chapter 2.

Dr Xavier Hadoux and Dr Zhichao Wu assisted in the collection of the ERG data used in chapter 3.

Dr Maria Triglia assisted in the collection of the ERG data used in chapter 4.

This thesis is less than 80, 000 words in length excluding figures, tables, appendices and references as approved by the Research Higher Degrees Committee.

Marc George Sarossy

May 2022

Acknowledgements

First, I would like to thank all four of my supervisors, Prof Jonathan Crowston, Prof Dinesh Kumar, Dr Anne Weymouth and Dr Zhichao Wu. My gratitude in particular goes to Zhichao who rescued this PhD and provided countless hours of support even under very difficult family circumstances.

Secondly, the team at CERA: Peter van Wijngaarden, Jacinta Mackey and Keith Martin who made the possibility of the data collection in my rooms a reality. Equally, thanks goes out to my CERA collaborators Xavier Hadoux and Jessica Tang

I am also very grateful to Olivija Tsaketas and the staff of the Essendon Eye Clinic who helped me arrange the testing of the patients in my rooms.

Finally, and most of all, to my family; my wife Maria who assisted in the electrophysiological testing and was always there for support and encouragement and my son Alex who aways greeted the participants with a friendly word and a smile.

It has been a long journey but worthwhile.

Publications and Awards

Journal Publications:

1. Frishman L, Sustar M, Kremers J, McAnany JJ, **Sarossy M**, Tzekov R, et al. ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram. Documenta Ophthalmologica. 2018:1-5.
2. **Sarossy M**, Crowston J, Kumar D, Weymouth A, Wu Z. Prediction of glaucoma severity using parameters from the electroretinogram. Scientific reports. 2021;11(1):1-9.
3. **Sarossy M**, Crowston J, Kumar D, Wu Z. Empirical mode decomposition denoising of the electroretinogram to enhance measurement of the photopic negative response. Biomedical Signal Processing and Control. 2022;71:103164.
4. **Sarossy M**, Crowston J, Kumar D, Weymouth A, Wu Z. Time–Frequency Analysis of ERG With Discrete Wavelet Transform and Matching Pursuits for Glaucoma. Translational Vision Science and Technology 2022;11(10):19-19.

Published Peer Reviewed Conference Presentations

1. **Sarossy M**, Aliahmad B, Kumar D, editors. Timing of changes in the entropy of the electroretinogram with Glaucoma. Life Sciences Conference (LSC), 2017 IEEE; 2017: IEEE.
2. **Sarossy M**, Kumar D, Wu Z, editors. Relationship between Glaucoma and Complexity Measures of the Electroretinogram. 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII); 2021: IEEE.

Conference Presentations

1. **Sarossy M**, Aliahmad B, Kumar DK. Discrete wavelet transform of the electroretinogram for glaucoma classification; choice of mother wavelet by variable ranking. Investigative Ophthalmology & Visual Science. The Association for Research in Vision and Ophthalmology 2018, Honolulu, Hawaii

Awards

1. The Hugh Noel Puckle Scholarship 2018-2021

Table of Contents

List of Figures

List of Tables

List of Enclosures

List of Abbreviations

ADC	Analog to digital converter
AIC	Aikake information criterion
ALT	Argon laser trabeculoplasty
ApEn	Approximate entropy
ASM	Angular second moment
BMO-MRA	Bruch's membrane opening minimum rim area
BMO-MRW	Bruch's membrane opening minimum rim width
CDF	Cumulative distribution function
CEEMDAN	Complete ensemble empirical mode decomposition with adaptive noise
CoR	Coefficient of repeatability
CoSaMP	Compressive sampling matching pursuit
CWT	Continuous wavelet transform
dB	Decibel(s)
DC	Direct current
DFA	Detrended fluctuation analysis
DFT	Discrete Fourier transform
DistEn	Distribution entropy
DTL	Dawson-Trick-Litzkow electrode
DWT	Discrete wavelet transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EEMD	Ensemble empirical mode decomposition
EKF	Extended Kalman filter
EMD	Empirical mode decomposition
EMG	Electromyogram
ERG	Electroretinogram
eRGC	Estimated retinal ganglion cell count
EWT	Empirical wavelet transform
FD	Fractal dimension
FET	Field effect transistor
FFT	Fast Fourier transform
FK	Fejer-Korovkin wavelet
GON	Glaucomatous optic neuropathy
GONE	Glaucomatous optic neuropathy evaluation
GWN	Gaussian white noise
ICC	Intra class correlation coefficient
ILM	Inner limiting membrane
IMF	Intrinsic mode function
InGaAs	Indium-Gallium-Arsenide
IOP	Intraocular pressure
IRC	Inner retinal complex
ISCEV	International Society for the Clinical Electrophysiology of Vision
JFET	Junction field effect transistor
KSE	Kolmogorov-Sinai entropy
LED	Light emitting diode
LogMAR	Logarithm of the minimum angle of resolution
MARS	Multivariate adaptive regression splines
MDC	Mode DC
mfERG	Multifocal electroretinogram
mmHg	Milimetres of mercury
MOSFET	Metal oxide semiconductor field effect transistor
MP	Matching pursuit
ms	Millisecond
mv	Miillivot
NFL	Nerve fibre layer
NMDA	N-methyl D aspartate
ns	Nanosecond
OCT	Optical coherence tomography
OCT-A	Optical coherence tomography angiography
OMP	Orthogonal matching pursuit
OP	Oscillitory potentials
OPL	Outer plexiform layer
ORC	Outer retinal complex
PDF	Probability density function

PERG	Pattern electroretinogram
PermEn	Permutational entropy
PhNR	Photopic negative response
PSD	Pattern standard deviation
pSTR	Positive scotopic threshold response
PVEP	Pattern visual evoked potential
PWVD	Pseudo Wigner Ville distribution
RGC	Retinal ganglion cell
RMSE	Root mean square error
RNFL	Retinal nerve fibre layer
ROMP	Regularized orthogonal matching pursuit
RPE	Retinal pigment epithelium
SampEn	Sample entropy
SAP	Standard automated perimetry
SD-OCT	Spectral domain optical coherence tomography
SITA	Swedish interactive threshold algorithm
SLIC	Simple linear iterative clustering
SLT	Selective laser trabeculoplasty
STFT	Short time Fourier transform
StOMP	Stagewise orthogonal matching pursuit
SVM	Support vector machine
TBN	Tetramethylpyrazine
TD-OCT	Time domain optical coherence tomography
UKF	Unscented Kalman filter
um	Micrometer
us	Microsecond
VCDR	Vertical cup disc ratio
VEP	Visual evoked potential
VMD	Variational mode decomposition
W	Watt
WVD	Wigner Ville distribution

Chapter 1

Literature Review

1.1 Background on Glaucoma

It is often said that the first reference to glaucoma was made by Hippocrates in his Aphorisms. The reference is in the third chapter, section 31 :

> τοῖσι δὲ πρεσβύτῃσι, δύσπνοιαι, κατάῤΡ῾οοι βηχώδεες, στραγγουρίαι, δυσουρίαι, ἄρθρων πόνοι, νεφρίτιδες, ἴλιγγοι, ἀποπληξίαι, καχεξίαι, ξυσμοὶ τοῦ σώματος ὅλου, ἀγρυπνίαι,
>
> κοιλίης καὶ ὀφθαλμῶν καὶ Ρ῾ινῶν ὑγρότητες, ἀμβλυωπίαι, γλαυκώσιες, βαρυηκοΐαι.

Translated by Adams (Hippocrates 2016) as “To old people dyspnoea, catarrhs accompanied with coughs, dysuria, pains of the joints, nephritis, vertigo, apoplexy, cachexia, pruritus of the whole body, insomnolency, defluxions of the bowels, of the eyes, and of the nose, dimness of sight, cataract (glaucoma), and dullness of hearing”

The reference is the Ancient Greek and comes from the section on diseases of old age

ἀμβλυωπία	dim-sightedness	fem nom/voc pl
ἀμβλυωπίᾳ, ἀμβλυωπία	dim-sightedness	fem dat sg (attic doric aeolic)

or “amvlyopíai”

and γλαυκώσιες

or

“glafkósies”

There is some controversy about whether the term actually denotes a colour in the Greek (Messenger 1964) or whether it best translates as ‘glaze’.

The disease was elaborated by Galen in de Usu Partium (Galen 1563) but there remains some confusion about whether he referred to what is now known to be cataract or glaucoma.

Guillemeau (Guillemeau 1585) first described possible treatments in 1585. Banister made an unacknowledged translation of the work in 1622 (Banister 1622). He did not mention glaucoma but did identify cases of raised tension in the setting of incurable cataract. Brisseau (Brisseau 1709) was probably the first to properly distinguish cataract from glaucoma stating that “Cette maladie n’attaque point du tout le cristalin” meaning “This illness does not at all attack the crystalline lens” although he goes onto conclude that glaucoma is primarily a disease of the vitreous.

Writing in Latin in 1745, the German ophthalmologist Platner (Platner 1745) is thought to be the first to associate raised pressure, detectible with firmness of the eye to touch, with glaucoma. Demours (Demours 1818) described a number of cases of acute glaucoma. Revising the notes of his father, a famous ophthalmologist in Marseilles, he notes in various cases the symptoms of acute loss of vision with pain, and the signs of congestion, pupillary distortion and the convexity of the iris : ‘’l’iris par consequent est trop convexe”, meaning “As a result the iris is too convex”. He notes that while the pupil moves, it is weak. In the work, various treatments including inducing vomiting, drawing blood, using camphor and camomile are described as tried without success. He correctly noted that the nerve and retina were paralysed. Reading the descriptions of the cases in the original French, one gets the impression that the pathophysiology is within reach of Pierre Demours but never actually grasped.

In his monumental work in 1855, Mackenzie (Mackenzie 1854) devoted a chapter to glaucoma. He still maintained that the pathology was in the vitreous – in contradistinction to cataract which by then had firmly been shown to be a disease of the lens. He further made the distinction between acute and chronic glaucoma. A reference to raised intraocular pressure is made in Mackenzie’s stage 3 of glaucoma in which he described an “abnormal hardness of the eye with immobility and unequal dilation of the pupil”. The pathophysiology remained elusive with stages 1 and 2 being “generally a disease of the crystalline alone”. He felt the cause to be inflammatory and even suggests the use of “tincture of belladonna” as a treatment. In one case he dissected an eye in which “I observed the lens tremulous” a possible case of pseudoexfoliation.

Using the ophthalmoscope invented by Helmoltz (Helmholtz 1851), Donders (Donders 1855) made careful observations of the pulsations of the optic nerve vessels, including the effect of adding digital pressure to the eye. He characterized the pulses in the arteries and veins and the synchronization with carotid and radial pulses. This was the first step towards the measurement of the intraocular pressure.

Von Graefe writing in the journal he founded in 1854 (Von Graefe 1854) carefully observed that the central retinal arterial pulse was weaker and delayed in glaucoma compared with individuals without the disease. From this he made the monumental observation that:

> Das pathologische Vorkommen des Aretienpulses bei Glaucom glaubte ich durch ein Hinderniss in der Arteria centralis erklären zu müssen, welches in ähnlicher Weise die Blutzufuhr beschränkt, wie eine Vermehrung des intraokularen Drucks

> Translated as: "I thought that the pathological occurrence of the arterial pulse in glaucoma must be explained by an obstruction in the central artery, which restricts the blood supply in a similar way to an increase in intraocular pressure"

In brief he thought that the changes in the central arterial pulse in glaucoma were restricted in a similar way to the intraocular pressure inferring for the first time that glaucoma was related to the intraocular pressure. Von Graefe went on to show that paracentesis provided a temporary relief from acute glaucoma and iridectomy a permanent cure. In 1858 Weber (Weber 1858) observed that "Careful and competent use of the ophthalmoscope will show the disc to be hollowed instead of bulging forward as has been previously maintained" – the first good description of optic disc cupping.

In an address to the thirteenth annual meeting of the British Medical Association, William Bowman FRS (Von Graefe 1854) spoke of glaucoma "the name of which, indeed is old, but the knowledge and the practice are all new". He stated categorically that "whatever the essential nature of the glaucomatous state, we as practitioners are chiefly concerned with the augmented tension of the eyeball which attends it." He went on to describe a technique of estimating the intraocular pressure with two forefingers.

Von Graefe also developed a tangent screen method for field examination. Donders' associate Haffmanns noted in 1862 the field changes in glaucoma (Haffmanns 1862) and the prognosis if the changes were advanced:

> Wenn Gesichtsfeldbeschränkung besteht, so ist die Prognose ungünstiger, besonders wenn die Beschränkung sich bis nahe an die Mittellinie erstreckt

Translated as: The prognosis is not good with field constriction especially if the field loss comes close to the centre line.

A number of tangent screen methods were developed to measure visual field but the technique received a huge boost when Hans Goldmann of Bern developed his hemispheric bowl perimeter in 1945 allowing precise control over fixation and background illumination as well as stimulus size and shape. Results of Goldmann visual field measurements are plotted as lines of equal sensitivity known as isopters. It was quite a difficult technique to assess glaucoma. One protocol used for the disease (Sommer, Quigley et al. 1984) used 4 isopters with the first at an intensity of the average threshold level of two spots 25 degrees temporal and 15 degrees above and below fixation. Second and third isopters were plotted 0.2 log units below and 0.3 log units above that level respectively. This was augmented by multiple static threshold measurements. In the study, the average time for each eye was between 45 to 60 minutes. An example of the visual field from that study is shown Figure 1-1. Other techniques for the detection of glaucomatous changes have also been described such as the Armaly technique (Rock, Drance et al. 1973) with reasonable sensitivity and specificity.

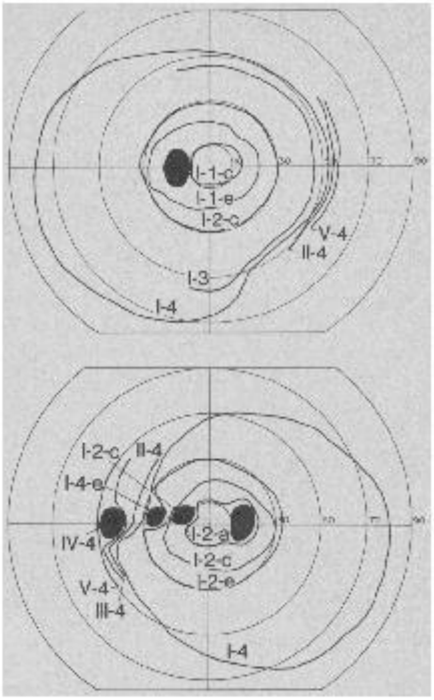


Figure 1-1: Visual field evaluation in glaucoma with Goldmann perimetry (from Sommer 1984). The upper panel shows the left eye with a normal visual field. the lower panel shows the right eye with visual field defects above and below the midline in the nasal field.

Automated static perimetry was developed in 1975 (Aulhorn 1975), (Fankhauser, Spahr et al. 1977). Static perimetry measures the threshold for perception of a stimulus at each of a large number of eccentric locations. A typical strategy collects data from 54 locations spanning the central 48 degrees of the visual field. The brightness of the dimmest stimulus seen at each location is determined to be the threshold and multiple measurements are required at each location to determine this value. The threshold at each point is measured in decibels (dB) which is a relative measure indicating the attenuation of the light intensity from the maximum. Assuming a sigmoidal probability of seeing curve, Spahr (Spahr 1975) described a method of limits strategy that forms the basis for the modern approach. In this technique, the first stimulus is presented at the average threshold for age matched controls at that location. If the stimulus is seen the brightness is reduced by 4dB and if it is not seen it is increased by 4dB. The bracketing strategy is then used to refine the threshold value down to an accuracy of 1dB (Bebie, Fankhauser et al. 1976).

Using a large dataset from patients, Olsson and colleagues (Olsson and Rootzén 1994) developed a Bayesian Markov model of thresholds which took into account parameters such as the age of the test subject and known likely direction of visual field defects. This is illustrated in Figure 1-2.

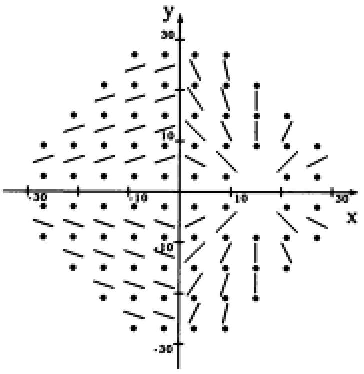


Figure 1-2: Most likely direction of visual field defects (from Olsson 1994)

This concept was used to increase the speed of visual field testing (Bengtsson, Olsson et al. 1997). Continually updating the prior probabilities of the thresholds at each point, the optimum strategy could be modified during the test to reduce the testing time required. Further development of strategies have led to even faster testing times (Glass, Schaumberger et al. 1995, Bengtsson and Heijl 1998), although sometimes speed has been at the expense of precision and reliability.

Rigorous statistical analysis of the threshold data was introduced in 1987 (Heijl, Lindgren et al. 1987). This statistical package introduced the field parameters widely used today and found within this thesis. The analysis described features of each location in terms of total deviation (the difference of each point from the age matched control) in both decibels and probability. The analysis also described global indices of the field such as the mean deviation (a weighted mean of the total deviation from age matched controls), the short-term fluctuation, the pattern deviation and the corrected pattern standard deviation. The importance of this development was in the description of global markers of disease severity. This was not something that was easy to do with the Goldmann or other perimetry techniques. The program also allowed analysis of a series of fields by examining boxplots of the raw decibel thresholds and allowing inspection of how the median and interquartile ranges changed over the series. Most of these measures are still the standard features used by glaucoma clinicians today.

The first in-vivo measurement of intraocular pressure seems to be Wahlfors (Wahlfors 1888) using an invasive method in eyes doomed by malignancies. Of course, the utilization of contact tonometry would require local anaesthesia. At the University of Wurtzburg, the noted pharmacist Joseph Rossberg supervised two junior colleagues. Anrep (von Anrep 1880) described the local anaesthetic action of the drug on animals and on himself and recommended the agent as a surgical anaesthetic. Theodor Aschenbrandt (Aschenbrandt 1883) was working in the same department and later became a German army doctor. In that role he secured a supply of cocaine and gave it to Bavarian soldiers noting the beneficial effects on fatigue. Those observations on fatigue came to the attention of Sigmund Freud, then a young neurologist. He encouraged a Viennese ophthalmologist Carl

Koller, working at the same hospital, to take the drug. Koller noticed the anaesthetic effect on his tongue. He went on to do further experiments on animals and himself and suggested it as an anaesthetic for ophthalmology.(Koller 1884). Henry Noyes of New York was at the meeting and brought the idea to the USA. Koller's article was translated into English and published in the Lancet (Koller 1884). Interestingly with his animal experiments, Koller found the topical anaesthesia adequate for the cornea but it evidently did not allow anaesthesia of the iris or intraocular surgery. Koller was proposed for the Nobel Prize on a number of occasions but was never a recipient. The European Society of Regional Anaesthesia and Pain Therapy created the Carl Koller Award in 1984 in recognition of Koller's discovery.

A large review of the state of knowledge of intraocular pressure at the time was written by Priestley Smith (Smith 1879). In it he details the principles and practice of the measurement of intraocular pressure. Following on from Monnik (Monnik 1870) he states that:

> When an impression is made in the sclera, the underlying fluid is displaced, and, being incompressible, must be found room for elsewhere. The ease with which this is accomplished depends chiefly upon the degree of pressure already exercised upon the whole internal surface of the sclera by the contents of the eye, but not upon this alone.

The ease of indentation also depended on elasticity of the sclera, the size and shape of the eye and. the fullness of the intraocular vessels. Preistley Smith developed his own tonometer which he considered superior to that of Monnick but in his own words "The estimation of 'tension' by this instrument is, it must be confessed, a tedious, and, apparently, a complicated process."

By this time, the pathophysiology of glaucoma was now well known as he states:

> Glaucoma is characterised by one constant and essential symptom — increased tension of the eyeball. Let this physical condition be added to any eye, healthy or diseased, and straightway it becomes glaucomatous, and will, under continuance of the pressure, manifest those changes which are peculiar to the disease.

Tonometry improved through various innovations of Schiötz (Schiötz 1920), Maklakoff (Maklakoff 1885) and Goldmann (Goldmann 1955), (Goldmann and Schmidt 1957) to the accurate reliable measurements we have today.

Optical Coherence Tomography (OCT) was the next major development in the characterization of glaucomatous damage. Described in 1991 by Huang (Huang, Swanson et al. 1991), the initial proof of concept was performed on the peripapillary region of the retina and the coronary artery. These experiments were in vitro. The first pilot study using the technique specifically on the retinal nerve fibre layer was published by Schuman shortly afterwards (Schuman, Hee et al. 1995) and (Schuman, Hee et al. 1995).

The first OCT implementations were adaptations of the Michaelson interferometer. These later came to be known as time domain OCT (TD-OCT) to differentiate them from the use of multiple frequency systems. The operating principle was fairly simple: a monochromatic light source is projected onto the retina as a point. At this point, the reference mirror is moved axially and the peak of the interferogram is measured. The amplitude of reflectivity at a given reference location is proportional to the height of the interferogram and the depth being measured is equal to the reference mirror location. In essence, reflectivity was measured at a single geographical point at a single depth. The 3 dimensional cube of reflectivity was built up one voxel at a time, which severely limited speed and resolution.

Simultaneous use of multiple wavelengths for OCT was first suggested in 1998 (Hellmuth and Welle 1998) and patented for use in the OCT. The theory is very complex. In brief, a smooth broadband low coherence light is used to illuminate a point on the retina. The moving mirror of the Michaelson interferometer is changed to a fixed mirror. The interferogram is now passed through a diffraction grating to resolve its spectrum. This is imaged through a very fast line camera (up to 200kHz line rate) generally based on InGaAs (Indium Gallium Arsenide) photodiodes. In this image, the position along the line (x axis) corresponds to the wavelength and the intensity is the amplitude (y axis).

This curve represents the Fourier transform of the interferogram. In most current implementations, the phase information of the light is not preserved so this is a real valued signal and is the modulus of the Fourier transform.

Mathematically, without considering refractive index, it can be shown that (De Boer, Leitgeb et al. 2017)

$$I(k) \ = S(k) \left| e^{2ikr} + \int_0^\infty a(z) e^{2ik(r+z)}\, dz \right|^2 \tag{1-1}$$

Where I(k) is the measured interferogram,S(k) is the wavelength dependent intensities from the sample , k is the wave number and a(z) is the reflectivity at depth z – the A scan at that point. 2r is the path length in the reference arm and 2(r+z) is the path in the object arm.

This can be simplified to

$$I(k) = S(k) \left[1 + 2 \int_0^\infty a(z) cos(2knz) dz + \int_0^\infty \int_0^\infty a(z) a(z') e^{-i2kn} (z - z') dz dz' \right] \tag{1-2}$$

It can be shown that a function containing a(z) can be extracted by deconvolving the square of the Inverse Fourier transform of S(k) out of the square of the inverse Fourier transform of I(k). In doing so, there are three terms left are the DC component, a(z) and its symmetrical mirror and an autocorrelation noise term from reflections within the sample. The mirror image appears because the phase information is not known or used. Thus it is necessary to have the reference position outside the area of interest so that the image and its mirror can be easily separated. The need to have the reference above or below the area of interest (rather than in the middle) does significantly reduce the maximum depth of scan. For anterior segment scans of the eye, some equipment lacks the capacity to place the reference away from the centre of interest and the scan can be mirrored as shown in Figure 1-3 from Tun. (Tun, Baskaran et al. 2017)

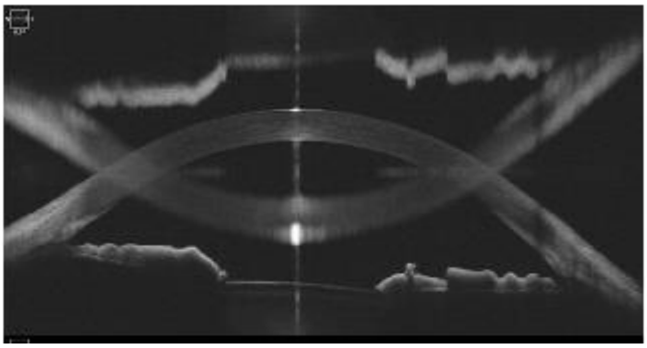


Figure 1-3: Anterior segment OCT showing duplication of image

This thesis will explore an emerging new quantitative technique for glaucoma assessment via electrophysiological measurements. The current approach – analogous to the tangent screen and finger tonometer will be developed to the equivalent of the modern perimeter, Applanation tonometer and the SDOCT.

1.2 Definition and Features

At its core, glaucoma is a disease of the optic nerve of the eye – that is the axons of the retinal ganglion cells. Classification is based upon the cause of the disease whether it is primary or secondary, open or closed angle, congenital or acquired, adult or paediatric. Controversy does remain about the precise definition of glaucomatous optic neuropathy (Quigley 2019). For population-based studies, a rigorous definition is required. The scheme of Foster (Foster, Buhrmann et al. 2002) has been frequently used based upon the appearance of the optic nerve and changes in the visual field. As Quigley notes (Quigley 2019), a new definition is being built around changes in the OCT, correlation of OCT features with visual field changes, and finally exclusion of retinal disease to explain those changes. This discussion will, for brevity, be limited to structural changes of glaucomatous optic neuropathy (GON) and functional changes as detected by the visual field. This remains a work in progress (Iyer, Vianna et al. 2020, Iyer, Boland et al. 2020)

Ultimately, whatever the cause of raised intraocular pressure or increased susceptibility to pressure, the disease is a loss of ganglion cells through apoptosis (McKinnon 1997). Apoptosis described by Kerr in 1972 (Kerr, Wyllie et al. 1972) is a programmed, suicidal cell death with phagocytosis of the cell without inflammation. Various causes of the trigger have been described and investigated including the baric trauma on the axons, hypoxia of the axons, chronic intermittent ischaemia, glutamate excitotoxicity, defective axon transport and the presence of trophic factors (Evangelho, Mogilevskaya et al. 2019).

The first clinical challenge in glaucoma is early detection. The initial changes of open angle glaucoma are asymptomatic and much damage may have occurred before detection is possible. Various screening programs have been developed and evaluated ranging from various combinations of intraocular pressure measurement (Awoyesuku and Onua 2019), optic disc photography (Zhang, Sun et al. 2019) and OCT. New machine learning algorithms for example (Medeiros, Jammal et al. 2019) may improve the diagnostic accuracy of cheaper technologies in this regard.

The next challenge in the management of glaucoma is: given the factors which influence the susceptibility of the ganglion cells to the intraocular pressure, what is a safe intraocular pressure? That is, for an individual patient with a set of parameters that cannot be changed (age, eye geometry, co-morbidities etc), can a model be built that predicts the risk of progression as a function of intraocular pressure? Then, based upon that model, determine an acceptable intraocular pressure considering the cost – whether in terms of financial cost or side effects. This is the conscious or subconscious calculation that clinicians make when working in this area.

Before apoptosis, retinal ganglion cells undergo morphological changes – principally in their dendrites. (Liu, Duggan et al. 2011) It is likely that subtle structural changes are detectible via high resolution imaging at this stage and, further, that intervention by reducing intraocular pressure may save some of these cells. It is more likely though that this pre-apoptotic change will manifest as changes in the functional measures before the structural changes such as changes in the nerve fibre layer or disc morphology occur. This structure-function gap, in theory, will predict how much of the deficit can be recovered with optimisation of physiological conditions for the ganglion cells. The key functional tests that are currently available are the visual field test and electrophysiological testing. If electrophysiological testing can be made sufficiently reliable it may therefore inform a clinician as to the potential extent of possible recovery of function.

1.2.1 Photographic / Ophthalmoscopic features

The well-known features and risk factors apparent at the optic nerve head have been concisely described as part of the GONE project (O'Neill, Gurria et al. 2014) and include disc size, disc shape, disc tilt, peripapillary atrophy, vertical cup disc ratio, cup depth, nerve fibre layer defects (notching) and disc haemorrhages. Of these measures, the optic disc size, shape and tilt are correcting factors when considering the appearance of the nerve and do not of themselves indicate an increased susceptibility to pressure.

The vertical cup disc ratio (VCDR) is the classic clinical sign for the diagnosis of glaucoma and the measurement of its progression. Tatham (Tatham, Weinreb et al. 2013) investigated the relationship between a model of ganglion cell counts and the VCDR and found the VCDR to be quite insensitive and quite non-linear when the RGC number was plotted against the

VCDR. Part of that non-linearity can be explained by the fact that the rim area is linearly related to (1-VCDR^2).

Nerve fibre defects represent focal losses of nerve fibre bundles at the optic disc. Such defects have moderate to high specificity for glaucoma (Sugiyama, Uchida et al. 1999). Optic disc haemorrhage has a high correlation with progression of the disease (Ha, Kim et al. 2019) but is relatively insensitive. It indicates a more rapid rate of progression.

1.2.2 OCT features

The OCT has become well established in the diagnosis and management of glaucoma. The technology allows high resolution of the 3-dimensional structure of the optic nerve head as well as accurate measurements of the layers of the retina, which can detect focal or global changes in the thickness of the nerve fibre, ganglion or inner plexiform layers.

A number of types of scans are used in the modern evaluation of glaucoma:

- Cube scans centred on the macula: A sequence of high resolution linear scans are collected and then segmented to analyse the area layer by layer.
- RNFL scans: The OCT is collected as a circle centred on the optic disc and the layers are segmented.
- Optic nerve cube: A series of linear scans are collected in an area centred on the optic disc.
- OCT angiography: A new technique that uses high speed video OCT to find the vascular plexus, subtracting out the parts of the image that are not changing.

The utility of the OCT as a tool for the management of glaucoma has been improved by a number of technical improvements. As mentioned earlier, the change in basic technology from TD-OCT to SD-OCT and improvements in noise suppression and processing enabled an increase in depth resolution to the current standard of better than 11 um per pixel transversely and 3 um per pixel axially (Folgar, Yuan et al. 2014). The faster acquisition times enabled a reduction in the movement artefact. Coupled with this was the development of compensation for eye movements – so called gaze tracking.

These commonly used strategies are illustrated in the figures below

Segmentation is the process of extracting the boundaries between layers on the OCT. As the OCT is collected one line at a time, the process is usually applied to each line and represented as such. Figure 1-4 shows such a segmentation result.

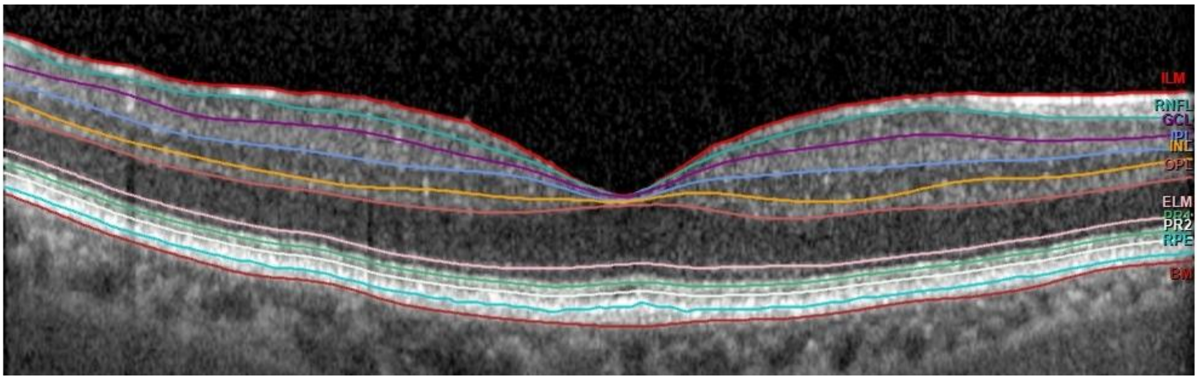


Figure 1-4: Macular OCT with segmentation . ILM= inner limiting membrane, RNFL=retinal nerve fibre layer, GCL= ganglion cell layer, IPL= inner plexiform layer, INL=inner nuclear layer, OPL=outer plexiform layer, ELM=external limiting membrane, PR1 = photoreceptor inner segments, PR2=photoreceptor outer segments, RPE=retinal pigment epithelial layer, BM = Bruch's membrane

Internally, the data for the line is represented as a 2 dimensional matrix. One axis has the linear position and the second axis the depth. The reflection intensity is encoded as the value of the cell of the matrix.

For a given position, the vector of intensities can be regarded as an 'A scan'. The A in A scan refers to amplitude modulation – essentially a plot of amplitude on the ordinate against depth on the abscissa. An example of this from Ishikawa (Ishikawa, Stein et al. 2005) is shown in Figure 1-5. In this methodology the ILM and RPE were first identified as the relevant peaks and then the largest valley as the Outer Retinal Complex (ORC – the combination of photoreceptor and outer nuclear layer). The intermediate layers were identified by the peaks and troughs. Segmentation consisted of a vector of boundary values for each transition.

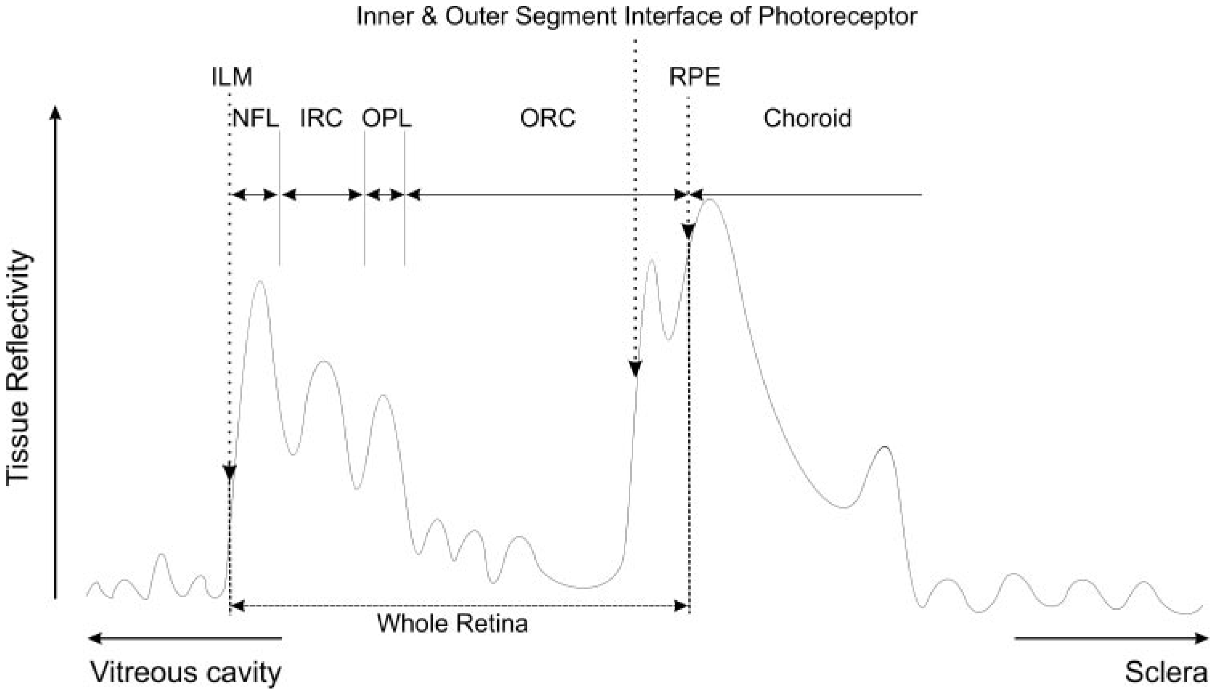


Figure 1-5: OCT A scan from Ishikawa 2005. The ordinate represents the depth of the scan. The abscissa represents the tissue reflectivity in arbitrary units. The algorithm detects the boundaries between the Inner Limiting Membrane (ILM), the Nerve Fibre Layer (NFL) the Inner Retinal Complex (IRC=Ganglion Cell layer + Inner Plexiform Layer + Inner Nuclear Layer), The Outer Plexiform Layer (OPL), the Outer Retinal Comple (ORC= Outer nuclear and photoreceptor layers), Retinal Pigment Epithelial Layer (RPE) and the Choroid

There have been many innovations since the original papers, for example Chiu (Chiu, Li et al. 2010) and Fang (Fang, Cunefare et al. 2017). The more recent algorithms use not only each A scan individually but the whole two-dimensional matrix. A variety of techniques such as Graph theory (LaRocca, Chiu et al. 2011) and deep learning (Pekala, Joshi et al. 2019) have been used. This improved segmentation has allowed extraction of reliable OCT measures. The actual segmentation algorithms within the commercial OCT machines are proprietary and details are not readily available. Algorithms using segmentation in three dimensions – i.e. considering pixels between brightness modulated scans (B-scans) have not been widely described or used. Features can be one dimensional, for example average thickness in um, two dimensional such as rim area in mm^2 or three-dimensional volumetric measures such as cup volume in mm^3. Volumetric data can be extracted from segmented or unsegmented optic nerve scans (Maetschke, Antony et al. 2019).

Well established and used OCT features of glaucomatous optic neuropathy (GON) include

- Retinal nerve fibre layer thickness: This analysis is usually performed on a circle surrounding the optic nerve. The layers are segmented and the nerve fibre layer is plotted. Parameters include the average RNFL thickness, and average thickness by clock hour, quadrant or semi-circle
- Rim parameters – using OCT characterization of the nerve head, cup disc ratio, rim thickness and rim area are calculated

- Ganglion Cell and IPL thickness: Using a scan centred on the fovea, the ganglion cell and inner plexiform layers are segregated. Parameters are extracted including averages and distributions
- Volumetric data of cup and disc volume

There are also a large number of studies proposing refinements to the measurement of OCT parameters for GON including

- More accurate characterization of the rim geometry through identification of the Bruch's Membrane Minimum Rim Opening (BMO-MRW) (Chauhan and Burgoyne 2013)
- Fourier analysis of the RNFL plot (Hsieh, Chang et al. 2020)
- OCT angiography (OCT-A) (Miguel, Silva et al. 2021)

OCT machines can be used to estimate traditional photographic optic disc parameters. This involves segmenting the image of the disc itself into disc, cup and (not disc not cup). The machine can do this either from the photograph taken for OCT targeting, or from the OCT itself. The latter method is preferable as the targeting photograph is often of low resolution. Parameters which can be extracted include Vertical cup disc ratio, horizontal cup disc ratio and rim area.

1.2.3 Visual Field Changes in Glaucoma

Standard Automated Perimetry (SAP) is generally used for the diagnosis and monitoring of glaucoma. The shape of the bowl may be spherical or aspherical (Rowe, Hepworth et al. 2019). Background illumination may be from a tungsten (Vingrys and Helfrich 1990) or a light emitting diode (LED) source (Cubbidge and Wild 2001). The stimulus itself may be delivered by a projection system or a fixed array of LEDs (Wild 1988). There has been much research into the optimum strategy for determining the threshold for each point measured by the strategy. A staircase approach is generally used which may be optimized for speed of collection by updating the strategy in real time as points are measured (Bengtsson and Heijl 1998). Results are reported as a matrix of thresholds and a number of summary statistics are collected. The most commonly used pattern of stimulus locations measures 54 locations are tested with a 6 degree spacing between points in the horizontal and vertical direction and the matrix spanning the horizontal and vertical midlines yielding points tested at 3 degrees from each midline (Rowe, Wishart et al. 2014). The data is represented as the absolute threshold – the attenuation in dB applied to yield the dimmest spot that could be seen at the location. The data is also represented as a difference between the absolute threshold and aged matched controls. This is usually also transformed into a probability plot of deviation with respect to geographical location. Drance (Drance 1972) described the classical visual field defects in glaucoma: Circumscribed paracentral defects, peripheral and central nasal steps, arcuate scotomas and sector shaped scotomas elsewhere in the visual field. These shapes correspond to the focal loss of groups of nerve fibres at the optic disc. Traditionally, recognition of these patterns and correlation with other clinical features allowed the diagnosis of glaucoma or the recognition of progression.

There has been much work done on automating the detection of progression between visual field tests (Arnalich-Montiel, Casas-Llera et al. 2009, Artes, O'Leary et al. 2014). By performing a point by point analysis of visual field threshold, the software attempts to correct for the effects of long-term fluctuations, inter-test variability, learning effects, fatigue, changes in the physiologic state of the eye and aging. (Giangiacomo, Garway-Heath et al. 2006). Another technique is the analysis of boxplots of the thresholds to detect changes in median and interquartile ranges. (Heijl, Lindgren et al. 1987).

1.3 Current interventions in glaucoma

Current interventions in glaucoma comprise topical drop therapy, laser therapy, micro-invasive surgery and conventional filtering surgery.

1.3.1 Drop treatment of glaucoma

Treatment of glaucoma with topical drops is usually the first line treatment (Ang, Fenwick et al. 2020). Agents are available from the following pharmacological classes:

- Prostaglandin analogues
- Alpha agonists
- Topical carbonic anhydrase inhibitors
- Beta blockers
- Rho kinase inhibitors

1.3.2 Laser treatment of glaucoma

Acute and subacute angle closure glaucoma is treated with laser peripheral iridotomy. The procedure, performed under local drop anaesthesia involves creating a small aperture in the peripheral iris with a brief pulse of infrared laser radiation. The procedure is effective in the long term in a large proportion of eyes (Buckley, Reeves et al. 1994, Aung, Ang et al. 2001).

Open angle glaucoma is also amenable to laser treatment. Argon Laser Trabeculoplasty (ALT) was described by Wise and Witter (Wise and Witter 1979) and comprised placing a series of small laser burns within the trabecular meshwork of the eye to improve aqueous humor outflow. The name ALT continued to be used even when solid state lasers of similar wavelengths were used.

ALT has now been superseded by Selective Laser Trabeculoplasty (SLT). The difference between the two techniques is the spot size (SLT 400um vs ALT 50um) and pulse duration (SLT 3ns vs ALT 1us) (Damji, Shah et al. 1999). The laser-tissue interaction is very different and SLT is considered superior with better retreatability and fewer side effects (Damji, Shah et al. 1999).

1.3.3 Micro invasive surgery

The latest surgical innovations are in the field of micro-invasive surgery. This involves the placement of a prosthesis either *ab externo* (Schultz, Schojai et al. 2020)or *ab interno* (Le, Bicket et al. 2019) to provide a reduction in aqueous outflow resistance and therefore a decrease in intraocular pressure.

1.3.4 Conventional glaucoma surgery

Conventional glaucoma surgery in current use comprises the construction of a bypass pathway for aqueous egress through modification of eye tissues (trabeculectomy) (Edmunds, Thompson et al. 2001) or placement of a prosthesis (Taglia, Perkins et al. 2002).

1.4 Mathematical formulation of clinical decisions in glaucoma

The practice of medicine is very complex and does not generally yield itself to explicit mathematical formulation. Nonetheless, there are a number of key clinical decisions in glaucoma which may at least be thought of in mathematical terms. In glaucoma, these must be expressed per eye. Clinicians diagnosing and managing the disease at least subconsciously must address these questions.

1.4.1 What is the probability that the eye has glaucoma?

For this question, the answer can be considered the outcome of a classifier. That is, given a vector of features (continuous and categorical) which include demographics, intraocular pressure, optic disc features, OCT parameters, visual field parameters and perhaps electrophysiological parameters

$$P_{glauc} = f(\Theta) \qquad 1\text{-}3$$

Where Θ is the vector of parameters and P_{glauc} is the probability of glaucoma. In this formulation, the classifier (such as a Support Vector Machine) yields a probability value and the acceptance of the classification is on the basis of a threshold – for example, it is accepted as a diagnosis of glaucoma if over 0.5 (Townsend, Wollstein et al. 2008). Other techniques, for example artificial neural network (Bengtsson, Bizios et al. 2005), may yield a class diagnosis such that

$$X = f(\Theta): X \in \{glaucoma, nonglaucoma\} \qquad 1\text{-}4$$

The process in equation 1-3 is essentially what an ophthalmologist working in glaucoma does in evaluating a glaucoma patient. Although there have been many machine learning implementations of both 1-3 and 1-4, practical implementation is limited to utilizing a small number of parameters (for example disc photographs) in a screening, telemedicine or developing country setting.

1.4.2 Given that the eye has glaucoma, what is the severity?

This problem is to find an answer to the equation

$$S_{glauc} = f(\psi) \tag{1-5}$$

Where S_{glauc} is a continuous bounded variable determined by the feature vector Ψ, or alternatively

$$C_{glauc} = f(\psi) \colon C_{glauc} \in \{mild, moderate, severe\} \tag{1-6}$$

In 1-6 the severity of glaucoma is given by a membership of an ordered set. This can have any number of classes but the gradations are very coarse. This sort of formulation is used for simple population studies or insurance purposes. It has only minimal value in clinical practice as it is not reasonable to only detect movements between classes rather than fine grained changes. Equation 1-5 produces a scalar measure of disease severity from the input vector.

The idea of a single continuous bounded scalar measure of glaucoma severity is attractive in principle but it has been difficult to achieve in practice. There have been a number of studies relating the findings in OCT to those of the visual field (Drasdo, Millican et al. 2007, Hood, Anderson et al. 2007)

Harwerth (Harwerth, Wheat et al. 2010) and Medeiros (Medeiros, Lisboa et al. 2012) have formulated models that use average circumpapillary retinal nerve fibre (RNFL) measures with threshold measures from each location of the visual field to predict ganglion cell counts. This is the measure used later in this thesis. .A simplified structure-function index was recently proposed by Wu(Wu and Medeiros 2021).

In current clinical practice, the aggregate measures of disease severity are crude and generally use the mean deviation of the field, the visual field index (Bengtsson and Heijl 2008), a gross measure of field constriction using a suprathreshold stimulus or an impairment measure based upon a combination of acuity and constriction along meridians of a dynamically tested visual field (Cocchiarella, Andersson et al. 2001).

1.4.3 Given that the eye has glaucoma, is the condition progressing?

One way of detecting progression would be to look for changes in the severity over time. That is from 1-5 we would look for

$$\Delta_{glauc} = S_{glauc}(t + \tau) - S_{glauc}(t) \tag{1-7}$$

Where the measurements are at time t and a time increment τ

In practice the approach in 1-7 is severely limited by the fact that the severity measures are inadequate.

However, a number of sophisticated techniques have been described to answer the related question of what is the probability that progression is taking place given two (or more) observations of clinical features.

$$P_\delta = f(\Theta_{t+\tau}, \Theta_t) \qquad \text{1-8}$$

Where P_δ represents the probability of there being progression or

$$X = f(\Theta_{t+\tau}, \Theta_t) \colon X \in \{true, false\} \qquad \text{1-9}$$

A number of successful studies have used these techniques including features sets from fields and OCT parameters and using techniques such as relevance vector machine (Bowd, Lee et al. 2012) and recurrent neural networks (Park, Kim et al. 2019).

1.4.4 Given that the eye has glaucoma and is progressing, what is the rate of progression?

Many studies have investigated the rate of progression of visual field changes or the rates of progression in OCT parameters (Suda, Akagi et al. 2018, Nguyen, Greenfield et al. 2019). It is important to note that this does not necessarily estimate the rate of progression of the underlying disease. In fact, these studies showed divergence between the visual field and OCT estimates of glaucoma. If the underlying severity could be accurately obtained as per 1-7, this measure in theory would be simple by dividing the amount of change in by the time increment. In practice, there has been no convincing demonstration of this metric.

1.4.5 Defining an individual's risk curve

It is a challenge to determine the form of the curve relating pressure in an individual eye to the risk of glaucoma progression. Even in experimental animal models, the problem is difficult to test because of the irreversible nature of glaucoma, its chronic nature and the problems in correlation between humans and other species.

Bui (Bui, Batcha et al. 2013) studied the relationship between the magnitude of pressure elevation and the amount of retinal damage in rats over a 4 week period. The group investigated various electroretinogram (ERG) parameters including the scotopic threshold response (STR), the maximum response amplitude of the photoreceptor response (Rm_{p3}), The rod bipolar scotopic response (P2), the oscillatory potentials (Ops). All of the ERG parameters demonstrated an inverted sigmoidal relationship with IOP and showed significant recovery 4 weeks after the insult. It could be deduced from that result that for an individual, the probability of any progression may have a sigmoidal relationship with IOP.

The probability of progression at a given intraocular pressure is the conditional probability of rate given pressure. This can be expressed as the well-known formula

$$P(rate|IOP) = \frac{P(rate \cap IOP)}{P(IOP)} \qquad 1\text{-}10$$

Equation 1-10 returns the normalized probability of a given rate for a certain IOP. The probability of the combination of any given rate and IOP can be considered a joint probability. The relationship, if known, can be plotted.

Figure 1-6 shows the surface/heatmap plot for a simulated joint probability distribution. This particular simulation was generated using a copula. Sklar's theorem (Sklar 1959) from financial mathematics states that under certain continuity conditions, a multivariate joint cumulative probability distribution can be decomposed into the two marginal probability distributions linked by a function known as a copula. The copula describes the nature of the dependence of the two variables and is often specified by a single parameter. In this example, the distribution shown was generated from a Normal marginal distribution for IOP and a beta distribution for the progression rate. A Gumbel copula was used with alpha of 5.

The Gumbel copula C(u,v) describes the joint cumulative probability distribution (CDF) as

$$C_\alpha(u,v) = exp\left[-\left(\left(-log(u)\right)^\alpha + \left(-log(v)\right)^\alpha\right)^{\frac{1}{\alpha}}\right] \qquad 1\text{-}11$$

Where $C_\alpha(u,v)$ is the joint CDF of marginal distributions u and v with parameter alpha. The Gumbel copula shows more dependence in the positive tail than the negative (Zhang and Singh 2007). If the joint probability is known, then the conditional probability of rate change for a given IOP or of IOP for a given rate can be determined. The conditional probability like any probability distribution will sum to 1. Figure 1-7 shows the conditional probability of a given rate of progression for different IOPs derived from the joint probability of Figure 1-6. They are, in essence the cross-sectional profile of the surface plot along the IOP line and then normalized to sum to 1.

For a joint probability of rate and IOP, two marginal probability distributions exist. These are the marginal probability of rate and the marginal probability of IOP. They represent the conditional probability of rate for all IOPs and the conditional probability of IOP for all rates. All joint probabilities have respective marginal distributions although the generation of a joint distribution from 2 marginals might not be possible with any known copula function.

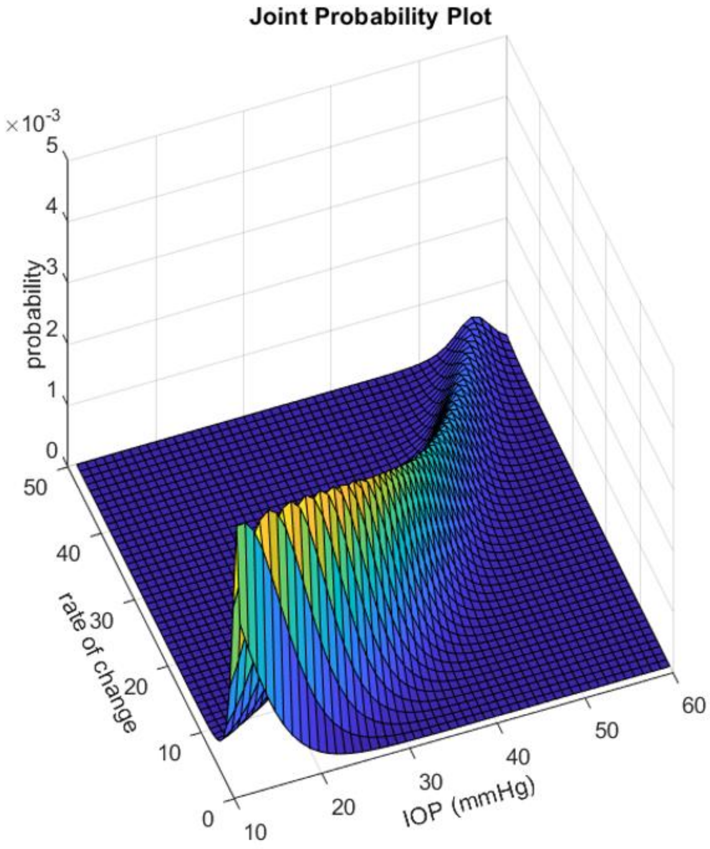


Figure 1-6: Surface plot of joint probability

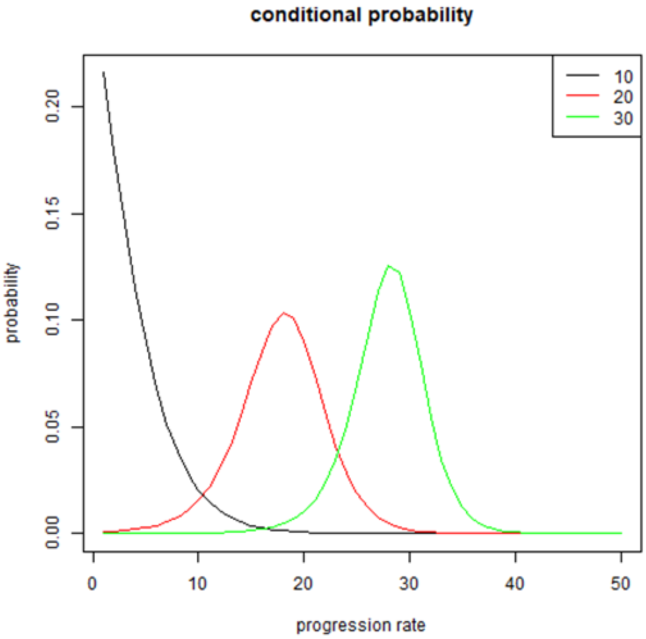


Figure 1-7: Conditional probabilities of progression rate for different IOP values

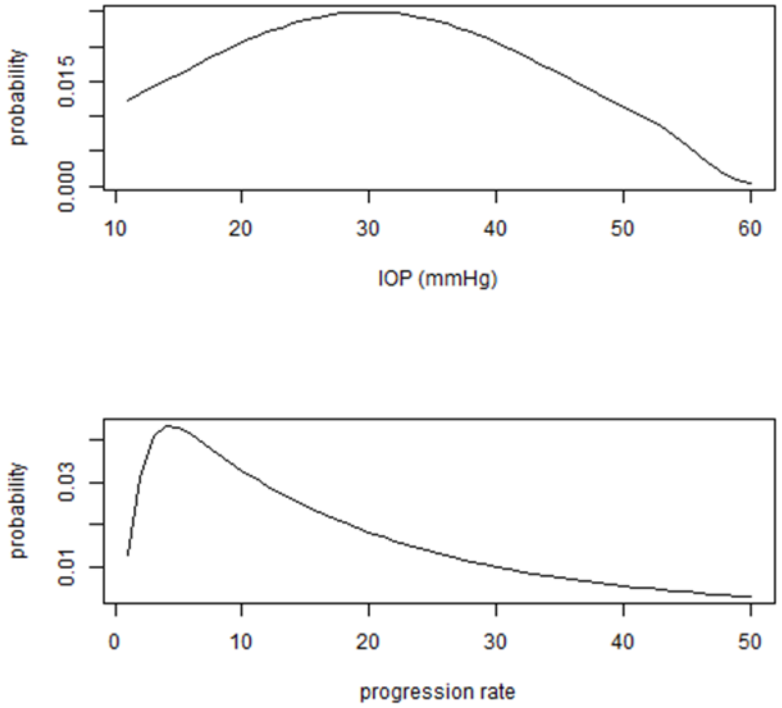


Figure 1-8: Marginal distributions of the simulated joint distribution

Figure 1-8 shows the marginal distributions from the joint distribution of Figure 1-6.

If an acceptably low rate of change can be determined – say in this case for example less than 5 of the arbitrary units used for the discussion, the conditional probability is shown in equation 1-12

$$P(IOP|[rate > 5]) \qquad 1\text{-}12$$

Graphically, this becomes the normalized marginal for IOP when the rates of 5 or less have been removed – in this case the 5 bottom rows. This is illustrated in Figure 1-9 which shows the probability distribution in the upper pane and the cumulative probability distribution in the lower pane. The line intersecting a probability of 0.05 is shown corresponding to a pressure of 20mHg

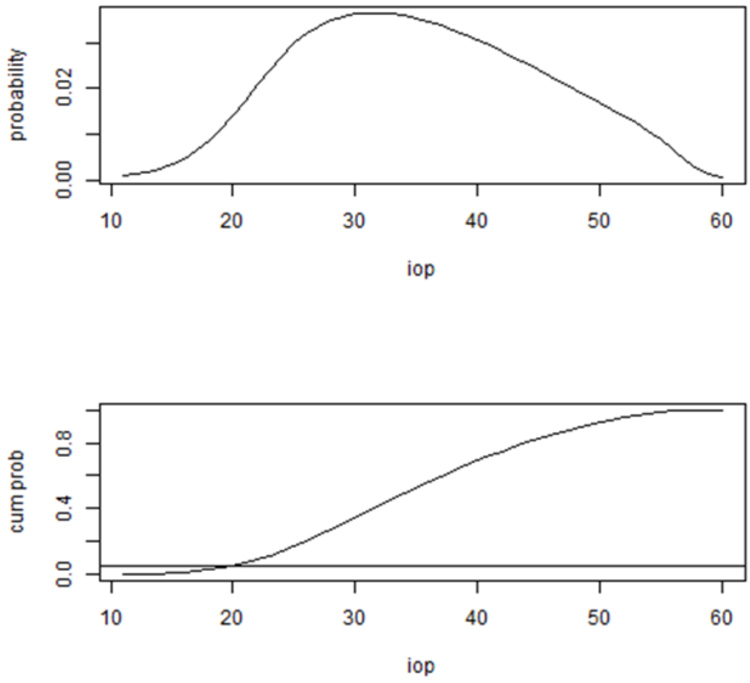


Figure 1-9: Probability of unacceptable rate of progression

This is a rigorous formulation of a target pressure. Given a joint probability distribution of rate and IOP one can determine the IOP for which the probability of an unacceptable rate of progression is less than a certain percentage.

The solution to the general problem is not yet feasible – that is, for a given vector of features about a patient with glaucoma, characterize the function yielding the risk of progression as a function of IOP and from that determine the target pressure. The conditional probability in the upper pane of Figure 1-9 is not a Normal distribution and therefore the cumulative distribution in the lower pane is not a cumulative Normal distribution. However, the cumulative Normal is probably sufficiently close for actuarial risk modelling – ie determining a cumulative normal distribution from the input feature vector, and from that predict a target pressure. A glaucoma practitioner would usually use this as a Bayesian prior distribution to update estimates of target pressure depending on what the intraocular pressure had been and whether progression had actually occurred, as these can be determined with a reasonable degree of accuracy.

This is the essence of the problem for a glaucoma practitioner treating a patient with glaucoma: Given all of the clinical information Θ, what pressure do I want to have in that eye to have a negligible risk of any progression?

With the Bayesian approach, this model would be considered the prior probability and the parameters would be updated based on observations over time. This is again the practice of the glaucoma physician: If the disease is progressing despite attainment of the target pressure, the model must be adjusted to account for that observation, and after taking into account new or unaccounted for concomitant modulating factors.

It is hoped that following on from the work in this thesis, electrophysiological parameters will form part of the vector of features that estimate the parameters of the individual risk model. To determine whether or not parameters derived from the ERG can provide additional information for the estimation of an individual risk model will require a longitudinal study which is outside the scope of this thesis but will be discussed in the conclusion.

1.5 Prevalence and impact of glaucoma

Numerous studies have investigated the prevalence of glaucoma (Tham, Li et al. 2014, Keel, Xie et al. 2018). Subgroup analysis based on race, age and type of glaucoma has been extensively performed. In 2013, the total number of people aged 30-80 in the world with glaucoma was estimated to be 64.3 million, with 60% of those in Asia. The study (Tham, Li et al. 2014) predicted that the number would increase by 74% by 2040. The disease is the second most frequent cause of legal blindness registrations in industrialized countries and the leading cause of blindness in developing countries (Delgado, Abdelrahman et al. 2019).

1.6 Background on Electroretinography

1.6.1 The electroretinogram

Du Bois-Reymond (Bois-Reymond 1849) first showed that there existed a current between the cornea and fundus of the eye in the large Tench. One of his great innovations in making this possible was the development of non-polarizable electrodes fabricated from a zinc/zinc sulphate amalgam. Holmgren (Holmgren 1870) showed in the excised frog eye that the polarity was positive at the cornea and increased in response to illumination. Dewar and McKendrick replicated Du Bois-Reymond's method in the cat, rabbit, tortoise, and goldfish using a novel reflecting electrogalavanometer. The potentials were very small and difficult to record. In his Nature paper of 1877 (Dewar 1877) Dewar reported that the potentials could be recorded in humans but the method was "too exhausting and uncertain to permit of quantitative observations being made". Gotch (Gotch 1903) recorded the first electroretinogram in the frog.

Hartline (Hartline 1925) had success in recording electroretinograms from human subjects. The first method involved the subjects lying on their backs with eyes bathed in fluid held in place by watertight goggles and Vaseline. The second, more practical method involved the subjects having the corneas anaesthetized with holocaine and the active electrode placed in the conjunctival fornix with the reference electrode in the mouth. The recording apparatus was a string galvanometer (Samojloff 1915) in which a fine platinum thread was suspended in a magnetic field and the deflection photographed. A constant 400W light source was used with a water bath as an infra-red filter and a mechanical shutter used to provide the transient light source.

Granit (Granit 1933) collaborating with Hartline recorded from the eye using directly coupled valve amplification. The circuitry was notable for reducing baseline drift to a negligible level. Granit in this paper went on to decompose the electroretinogram into three separate components. Granit and Hartline along with Wald won the Nobel prize in physiology or medicine in 1967 “for their discoveries concerning the primary physiological and chemical visual processes in the eye."

Field effect transistor (FET) amplification had been implemented by 1971 (Werblin 1971) and the use of microcomputers in the stimulation and processing of the response followed soon after. The first international standard for electroretinography was published in 1989 (Marmor, Arden et al. 1989).

The use of the Ganzfeld in psychophysics first described by Cohen (Cohen 1957), was mentioned by Gouras in 1962 (Gouras and Gunkel 1962) and first described for electroretinography by Finkelstein (Finkelstein, Gouras et al. 1968) using a sphere coated with “paint made specifically for integrating spheres by the National Bureau of Standards”. This technique forms the basis for the current standard for full field electroretinography.

1.6.2 The electroretinogram in glaucoma

Since the development of clinically feasible apparatus to measure the electroretinogram, many studies have been concluded into its application for glaucoma. One of the first such studies (Leydhecker 1950) concluded that there were no detectible changes in in the b-wave, even in blind or almost blind eyes. Although the ERG is affected at very high levels of intraocular pressure (Karlberg, Hedin et al. 1968), the effect is temporary and reverses with normalization of the pressure.

Riggs (Riggs, Johnson et al. 1964) showed that when a visual stimulus pattern reversed in contrast, a small electroretinographic signal was produced. The response was shown to be generated by the inner retina (Groneberg and Teping 1980) and therefore had potential in the evaluation of glaucoma. It was termed the pattern ERG (PERG). The stimulus can be evoked by a number of different types of patterns such as sine wave or square wave modulated linear gratings. The most popular type of patterned stimulus in common use is the checkerboard grating as it produces the largest amplitude response. The response can be measured with pattern reversal (where the black areas become white and the white ones black) or with pattern appearance (where the transition is from uniform grey to areas of black and white). The essential feature is that the overall luminance of the whole pattern does not change – therefore avoiding any flash ERG signal. Stimulus technologies are either cathode-ray tube (CRT) displays or, more recently, organic light emitting diode (OLED) displays (Cooper, Jiang et al. 2013). Conventional liquid crystal display (LCD) stimulators are unsuitable because the on-to-off transitions are not the same speed as the off-to-on transitions. The response can be elicited at a variety of spatial and temporal frequencies and at a variety of different luminances and contrasts. A positive going wave at around 50ms post pattern change is known as the P50 wave followed by a negative wave around 95ms post pattern change known as the N95 wave.

The most recent International Society for the Clinical Electrophysiology of Vision (ISCEV) standard was published in 2013 (Bach, Brigell et al. 2013) and specifies that in routine clinical practice maximum contrast and high luminance should be used with a presentation frequency of 4 reversals per second. For research, the response has been measured at reversal rates of more than 6 per second where the waveform resembles a sine wave. The P50 and N95 components of the PERG have different spatial tuning suggesting that the N95 predominately originates in the ganglion cells but the P50 may have contributions from other retinal cells (Holder 2001).

The first report of the PERG in glaucoma was by May (May, Ralston et al. 1982) who described a single patient with severe glaucoma in whom the PERG was unrecordable. Around the same time, Arden (Berninger and Arden 1988) in a series of 9 patients showed significant reduction in PERG amplitude with glaucoma and a correlation with the glaucoma severity. Kreuz (Kreuz, de Moraes et al. 2018) showed that both P50 amplitude was reduced and the latency increased in preperimetric and established glaucoma and that the parameters correlated with changes on the Optical Coherence Tomography (OCT) scans.

The response of the PERG can be analysed from the steady state response. Bach and colleagues (Bach, Unsoeld et al. 2006) conducted a long term prospective study in ocular hypertension patients using the technique. A pattern reversal stimulus was used with a presentation of 15 reversals per second. The feature extracted from the response was the amplitude of the 15Hz component of the Fourier analysis. They used two different spatial frequencies and found the 0.8 degree checks more affected than the much coarser 16 degree checks. The group concluded that the PERG can help to predict stability or progression to glaucoma at least 1 year ahead of other measures.

The multifocal electroretinogram (Sutter and Tran 1992) is a technique of extracting responses from the retina corresponding to subsections of the visual field. A large number of separate stimuli are delivered to the eyes simultaneously, each stimulus a hexagon modulated in a binary sequence affine-orthogonal to all of the other hexagons. The technique has found great utility in diseases which affect the retina in a regional manner. The technique has been evaluated in glaucoma (Sakemi, Yoshii et al. 2002) but has not been shown to be able to detect the disease.

1.7 Background on the Photopic Negative Response

1.7.1 Discovery of the response

The PhNR was first described by Viswanathan et al (Viswanathan, Frishman et al. 1999) in 1999. It has been shown to be reduced in glaucoma, both in clinical studies and in animal models (Viswanathan, Frishman et al. 2001, Machida, Gotoh et al. 2008, North, Jones et al. 2010, Nakamura, Hangai et al. 2011, Preiser, Lagreze et al. 2013, Machida, Kaneko et al. 2015). The test has significant potential in aiding the diagnosis of glaucoma and in monitoring its progression or recovery.

1.7.2 Origin of the signal

There is strong evidence that the PhNR is generated by the spiking action of ganglion cells. Li (Li, Barnes et al. 2005) showed that the response was abolished in the rat after transection of the optic nerve – which was well correlated with declines in ganglion cell counts by histology. These authors showed, like Viswanathan (Viswanathan, Frishman et al. 1999, Viswanathan, Frishman et al. 2001) that the response was also abolished by the intravitreal injection of tetrodotoxin – a blocker of the action potentials of the ganglion cells. Supportive evidence was also reported by Barnard (Barnard, Issa et al. 2011) who showed that in a mouse model of dominant optic atrophy in which mice were heterozygous to a mutated OPA1 gene (B6; C3-Opa1Q285STOP), the PhNR was significantly reduced while differences in the a-wave or b-waves were minimal when comparing the wild and OPA1 genotypes. Luo (Luo, Yu et al. 2017) used a rat model of ganglion cell loss from N methyl D Aspartate (NMDA) excitotoxity to explore the protective effect of tetramethylpyrazine nitrone (TBN). They showed that the NMDA caused loss of ganglion cells on histology and loss of visual function behaviourally as well as a reduction in the PhNR. TBN preserved visual function, PhNR and ganglion cell counts- protecting against the toxicity of the NMDA. Chrystostomou (Chrysostomou and Crowston 2013) showed good correlation between loss of PhNR and the positive scotopic threshold response (pSTR) in an acute pressure elevation study in mice. Finally , the fact that the PhNR is reduced in anterior ischaemic optic neuropathy (Rangaswamy, Frishman et al. 2004) showed further evidence that the origin of the response was in the stratum of the retina affected by that disease; the ganglion cells.

1.7.3 Detection of the timing and amplitude of the response

Even the name – photopic negative response, rather than a-wave ” indicates vagueness about the process. Difficulties in isolating the inverted peak of the response has contributed to the challenges in it finding utility. The latency of the response is even more obscure to define than its amplitude. The electrochemical process driving the PhNR clearly overlaps the b-wave (bipolar cell depolarization) and the i-wave. The i-wave itself is present in most species (Rosolen, Rigaudière et al. 2004) but importantly does not seem to be present in rats or mice. The i-wave may be generated from ganglion cells (Rosolen, Rigaudiere et al. 2003) or from the off-pathway (Gouras and Mackay 1988). The photomyoclonic response (Johnson and Massof 1982) also contaminates the PhNR response. The simplest way to measure the PhNR is the lowest voltage following the b-wave. Other authors have used a fixed point in time or the average of a set of points following a fixed point in time (Bach, Unsoeld et al. 2006).

In earlier work (Tang, Edwards et al. 2014), we compared the reliability of the identification of the PhNR using a variety of measures. The base to trough remains the most popular method. The baseline reference itself can pose a dilemma in determining the reference point to the identified peak of the response. Many electrophysiological systems utilize detrending algorithms. In recent work from our unit, Tang (Tang, Hui et al. 2018) showed that the baseline drift could be most effectively removed by fitting a third degree polynomial to the drift signal.

The amplitude of the PhNR response is only one feature of the ERG signal that may have utility in the detection and grading of glaucoma. Other features including the timing and slopes of signals may provide more information. Indeed, we have demonstrated that reliability was improved by developing a composite measure combining the amplitude of the a-wave and b-waves with the PhNR (Wu, Hadoux et al. 2016). Ostensibly, this is to normalise the overall gain of the visual system but it also incorporates other features of the signal in doing so. The addition of extra features in extracting information from the PhNR can be done with or without *a priori* assumptions about the nature of relevant features. An approach with such *a priori* assumptions in the context of glaucoma has been applied by others to the multifocal ERG signal (Barraco, Adorno et al. 2011). In this work, a vector of 13 features including timing, amplitudes and slopes of multifocal waves in glaucoma and controls was analysed and classification trained using a neural network. The machine learning strategy used in this context must be able to control multicolinearity of the feature matrix and also should ideally be interpretable. It is likely that a linear support vector machine will yield more useful results for such a dataset. The *a priori* approach does however manage the high dimensionality better than traditional time-frequency analysis. This type of analysis including wavelets and short term Fourier transform techniques has been frequently utilised in analysis of the ERG (Miguel-Jiménez, Boquete et al. 2010, Gauvin, Little et al. 2015).

The i-wave, a positive going wave is often superimposed upon the PhNR, dividing it into two waves. There has recently been interest in analysing these two waves separately as PhNR1 and PhNR2 (Ortiz, Drucker et al. 2020). In this thesis, this is taken one step further and all three components are used – the PhNR1, PhNR2 and the i-wave.

Another approach to deriving information from the PhNR is to characterize the intensity-response curve and to derive parameters from it (Tang, Hui et al. 2018). As the response is a sigmoidal curve, the response can be characterized by its maximum asymptote (V_{max}) and the stimulus intensity that gives a 50% response (I_{50})

1.7.4 Optimizing the Stimulus

The parameters for maximizing the amplitude of the PhNR have been studied, notably by Rangaswamy (Rangaswamy, Shirato et al. 2007). In that study, the authors concluded that the stimulus of a single cone type by chromatically balanced stimulus and background produced a larger amplitude PhNR than a white-on-white stimulus but the benefit reduced as the stimulus strength increased. The use of long (>200ms) flashes (Viswanathan, Frishman et al. 2001) has been described and in that study there were identifiable on and off PhNR responses but the authors have not concluded that any clinical utility exists for this sort of method. It is important to note however that the stimulus parameters that maximize the amplitude of the response do not necessarily maximize its diagnostic power – either as a classifier (presence or absence of glaucoma) or as a regressor (severity of glaucoma). This issue is investigated in this thesis.

Investigators have studied the effect of stimulus strength on the PhNR and a number have constructed fitted intensity-response curves, usually a Naka-Rushton (Naka and Rushton 1966) sigmoidal function (Tang, Hui et al. 2018), (Prencipe, Perossini et al. 2020). In most cases, the response from an individual test subject follows that model and inferences are possible from the fit parameters. It has not yet been shown that such analysis has any advantage over the analysis of a single flash intensity. In our work (Tang, Edwards et al. 2014), we showed that brighter flash intensities introduced significantly more artefact into the response.

1.7.5 Relationship to glaucoma

The PhNR is reduced in a number of conditions including glaucoma. It has also been shown reduced in optic atrophy (Gotoh, Machida et al. 2004), central retinal artery occlusion (Machida, Gotoh et al. 2004), diabetic retinopathy (Chen, Zhang et al. 2008) and retinal vein occlusion.

The PhNR signal is known to be related to glaucoma and is abolished in absolute glaucoma. It has been said to suffer from a floor effect (Wu, Hadoux et al. 2016). Another issue is that the measurements of the PhNR and retinal nerve fibre layer thickness (RNFL) are on a linear scale whereas field measurements such as mean deviation are on a logarithmic scale. There have been studies investigating correlation between PhNR and mean deviation and PhNR and RNFL thickness (Colotto, Falsini et al. 2000, North, Jones et al. 2010, Hara, Machida et al. 2020). Tang (Tang, Hui et al. 2020) showed recovery of PhNR amplitude following improvement of IOP.

One of the difficulties in finding the relationship between the PhNR and the severity of glaucoma is that there is not a good measure of overall severity of glaucoma; that is there is no good formulation of equation 1-5. In this work, I will correlate features of glaucoma with estimated ganglion cell counts to gain more intuition about the response and its relationship to the severity of the disease

Table 1-1: Summary of relevant literature on the Photopic Negative Response in Glaucoma

Study	Participants	Eyes		Stimulus/Background Colour (peak wavelength in nm)	Stimulus strength (cd.s/m^2)	Background luminance (cd/m^2)	comments
(Colotto, Falsini et al. 2000)	Glaucoma OHT Healthy	11 8 8	Skin	White/White (focal ERG)	17.5	150	Focal PhNR was reduced in glaucoma but not OHT and correlated with PERG.
(Viswanathan, Frishman et al. 2001)	Glaucoma Suspects Healthy	62 7 18	DTL	Red (630)/Blue (450)	0.05 – 1.6	5012 scot Td	PhNR reduced in glaucoma and implicit time varied with age.
(Drasdo, Aldebasi et al. 2001)	Glaucoma Healthy	18 19	DTL	Red (650)/Blue (450) Blue-green (450-535)/ Red (650)	2500 Td 1200 Td	92.5 Td 5400 Td	S-cone PhNR, L&M cone PhNR affected in glaucoma. S-cone was the most sensitive measure.
(Cursiefen, Korth et al. 2001)	Glaucoma Healthy	18 9	Henkes	White/White Orange/Blue	3.4 Td 0.43 Td	24 scot Td 316 scot Td	PhNR was not reduced in advanced glaucoma.
(Aldebasi, Drasdo et al. 2004)	Glaucoma NTG OHT Healthy	18 9 18 19	DTL	Red (650)/Blue (450) Blue-green (450-535)/ Red (650)	2500 Td 1200 Td	92.5 Td 5400 Td	S-cone PhNR, L&M cone PhNR and PERG affected in OHT and glaucoma. However, only PERG reduced in NTG and there was no correlation between PhNR and PERG.

(Machida, Gotoh et al. 2008)	Glaucoma Healthy	99 30	Contact lens	Red (644)/Blue (470)	4.8	40	PhNR amplitude and ratio were reduced in glaucoma and correlated with MD, RNFL thickness, optic disc rim area and cup/disc ratio area. a-wave correlated with MD.
(Wakili, Horn et al. 2008)	Glaucoma Healthy	37 37	DTL	Blue (440)/ Yellow (550)	0.013 – 0.052	238	S-cone PhNR amplitude and L&M-cone b-wave was significantly reduced and S-cone b-wave implicit times were increased in glaucoma. S-cone PhNR had higher sensitivity for glaucoma than b-wave.
(Sustar, Cvenkel et al. 2009)	Glaucoma Healthy	28 41	HK loop	Green (513), Amber (594), Red (635), White on Blue (470)	0.08 – 7.5	10	In controls, PhNR amplitude was greatest using red stimulus and increased with stimulus strength. Difference between glaucoma and control greater using red stimulus than white.
(Mortlock, Binns et al. 2010)	Healthy	31	DTL and skin	Red/blue	1.5	7943 Td	b-wave/PT ratio had the lowest coefficient of variation and interocular and intersession limit of agreement. Baseline PhNR measures (BT/BF) were least repeatable. Skin electrode had slightly higher variability than DTL electrode.
(North, Jones et al. 2010)	Glaucoma OHT Healthy	30 23 28	DTL	Red (650)/ Blue(450) Blue-green (450-535)/ Red (650)	0.25 24 cd/m^2	2.7 108	S-cone and L&M-cone PhNR was reduced in glaucoma and untreated OHT while treated OHT were similar to controls.

(Binns, Mortlock et al. 2011)	Healthy	18	DTL	Red (625-700)/ Blue (450)	0.05 – 3.39	7943 Tds	Naka-Rushton function a good fit for the PhNR stimulus-response data.
(Preiser, Lagreze et al. 2013)	Glaucoma preperimetric manifest Healthy	11 18 26	DTL	Red(635)/Blue (450)	0.1 – 4	10	PhNR amplitude was reduced in pre-perimetric and manifest glaucoma. PhNR and PERG had similar AUC.
(Tang, Edwards et al. 2014)	Healthy	49	DTL	Red (635)/Blue (465)	0.5, 1, 2.25, 3	10	PhNR not very repeatable measure. Better repeatability if measured from b-wave peak to PhNR trough
(Wu, Hadoux et al. 2016)	Glaucoma	43	DTL and skin	Red (635)/Blue (465)	1.5	10	Similar variability between DTL and skin electrodes which was improved by increasing the number of sweeps.
(Wu, Hadoux et al. 2016)	Healthy	20	Sensor strip	Red (621)/Blue (470)	1.5	10	PhNR recorded on hand-held ERG device (RETeval®) using sensor-strip electrodes
(Kirkiewicz, Lubiński et al. 2016)	Glaucoma Healthy	90 45	DTL	Red (640)/Blue (450)	1.6	25	PhNR amplitude and ratio were reduced in glaucoma. The sensitivity of both measures in early glaucoma was low (53/60%) compared to advanced disease (77/80%).

(Cvenkel, Sustar et al. 2017)	Glaucoma Suspects OHT Healthy	10 17 7 24	HK loop	Red (635)/Blue (470)	2.5	10	PhNR reduced in glaucoma groups compared to controls, but there were no significant differences between the glaucoma groups. The PhNR had higher area under the ROC curve than PERG.
(Banerjee, Khurana et al. 2019)	Glaucoma Healthy	25 50	BurianAllen	Blue (448) / Yellow (592) , Red (635) /Blue (450) and White/White	0.018-2	10	Red on Blue is the most sensitive stimulus for glaucoma detection
(Hara, Machida et al. 2020)	Glaucoma Healthy	84 40	BurianAllen	Red (627)/Blue(470) and White/White	0.5-3.0	10	Phnr w/w vs chromatic. Correlation of PhNR1 with MD was higher for w/w than r/b. For PhNR2 correlation was higher for r/b than w/w
(Ortiz, Drucker et al. 2020)	Glaucoma Other disease	4 181	DTL	White/White	3.0	30	PhNR 1 vs PhNR2 high correlations found
(Pham, Goldberg et al. 2020)	Glaucoma	118	DTL	Red/Blue	1.0	190	Rapid N wave is a component of the PhNR
(Hui, Tang et al. 2020)	Glaucoma	57	DTL	Red (621)/ Blue(470)	0.07-12.56	10	B3 supplementation improved retinal function by improving Vmax of PhNR sensitivity curve
(Cvenkel, Sustar et al. 2022)	Glaucoma	32	HK Loop	Red (635)/Blue (470)	2.5	10	PhNR not useful for monitoring of glaucoma; fails to pick up deterioration

OHT: Ocular hypertension, DTL: Dawson-Trick-Litzkow electrode; IOP: intraocular pressure; MD: mean deviation; NTG: normal tension glaucoma; OHT: ocular hypertension; RNFL: retinal nerve fibre layer; ROC: receiver operating characteristic; Scot: scotopic; Td: trolands. *Stimulus strength units converted to convention cd.s/m^2 where data (e.g. stimulus duration and pupil size) available.

1.8 Reliability of the signal

Electrophysiological measurements can be used in a variety of ways:

- Detection of waveform morphology to make a diagnosis. An example is the electrocardiogram (ECG) in which inversion of the T wave and changes in the relative amplitudes of Q and R waves are important in the diagnosis of ischaemia or infarction. In visual electrophysiology the so-called negative ERG is of diagnostic significance in a number of conditions such as congenital stationary night blindness (Miyake, Yagasaki et al. 1986)
- Presence or absence of a waveform. The presence of periodic sharp wave complexes in the electroencephalogram (EEG) (Steinhoff, Racker et al. 1996) is indicative of Creutzfeld Jacob disease. Their absence is said to strongly argue against that diagnosis. In the ERG, the presence of any flicker ERG signal argues against achromatopsia
- Quantitative measurements. These are measurements of amplitude or timing of waveforms from an electrophysiological signal

All quantitative amplitude measures in clinical electrophysiology are problematic. Differences in electrode positioning, changes in skin conductance, variations in the electrical properties of the body have all made the use of precise voltage measures difficult to apply. Some of the measures can vary with age and prior light exposure (Weleber 1981). Differences in the amplitude of the ERG between age matched males and females have been demonstrated (Vainio-Mattila 1951) and in women there is variation with the menstrual cycle (Brulé, Lavoie et al. 2007). In ophthalmology, there are few examples of the use of electroretinogram (ERG) amplitudes for disease monitoring. In the paediatric population, serial ERG can be useful for monitoring for vigabatrin toxicity (Westall, Logan et al. 2002). Disease activity can be monitored using the ERG for Vogt-Koyanagi-Harada syndrome (Sakata, Lavezzo et al. 2019), vitamin A deficiency through poor diet (Lee, Sarossy et al. 2015) or malabsorption (Lee, Tran et al. 2019). If a change in recorded amplitude measurements is to be considered significant, it must fall beyond the limits of agreement (Sedgwick 2013). A test with poor test-retest reliability will need a large change of amplitude to have confidence that the change reflects progression or recovery from the underlying disease.

The use of absolute voltage criteria is probably best established in the field of cardiology. For example, in the electrocardiogram (ECG), the use of a single voltage measure – the R wave in lead aVL – was shown to have poor intersession reliability (Schillaci, Battista et al. 2012). Well established derivative measures have been described for this application in cardiology with improved diagnostic accuracy and reliability (Schillaci, Battista et al. 2012) with many of these combining timing and voltage measures such as the Cornell voltage product (Molloy, Okin et al. 1992). As shall be shown later in this thesis, the electroretinogram can be transformed into the frequency domain via the discrete Fourier transform (DFT) for transient signals such as the flash ERG or the Fourier Series (FS) for periodic signals such as the 30Hz flicker response. The magnitude of frequency components derived this way is an alternative method of analysis and in some circumstances can yield more reproducible results (Hess and Baker Jr 1984)

Changes in light adaptation level via inaccuracy of timing or even day to day variability of dark adaptation rate may also contribute to inter-session variability of the ERG. A recent study of cone adaptometry showed significant intersession variability in the time constant of the rate of adaptation (Gaffney, Binns et al. 2014). One possible way of addressing this is to record the ERG at a series of flash intensities and then fit an intensity-response curve to the data. The parameters of the fitted function (such as a Naka-Rushton equation (Naka and Rushton 1966)) are then the features used in the analysis of the signal. The curve can be fitted to a-wave, b-wave or the photopic negative response and has been shown under certain conditions to be more reproducible than the amplitude measures alone (Hébert and Lachapelle 1999).

Finally, a major contributor to intersession variability in the case of the PhNR is the ability to accurately identify the (negative) peak of the waveform. Contamination by exogenous and endogenous noise impairs identification of the latency and amplitude of the response. This thesis will explore novel filter techniques to improve that identification.

Measures taken to improve the reliability of an ERG measure relevant to glaucoma will increase the confidence in the credibility of a change and will therefore make the measure more useful in the detection of progression.

Chapter 2

Adaptive Filtering of the Electroretinogram

The findings of this chapter have been published in part (see Appendix)

Sarossy M, Crowston J, Kumar D, Wu Z. Empirical mode decomposition denoising of the electroretinogram to enhance measurement of the photopic negative response. Biomedical Signal Processing and Control. 2022;71:103164.

2.1 Abstract

Background and objective: The photopic negative response (PhNR) of the electroretinogram (ERG) can be useful in the diagnosis and management of glaucoma. Its utility is limited by its large degree of measurement variability, which can be caused by low frequency oscillations and baseline wander. In this work we show a new denoising strategy using the empirical mode decomposition (EMD) which improves repeatability. This technique is compared to the variational mode decomposition (VMD) and the empirical wavelet transform (EWT).

Methods: ERGs were performed on both eyes of 56 patients using red-on-blue stimulus after 2 minutes of preadaptation, capturing 10 sweeps per recording. Two recordings were obtained to provide intrasession test-retest estimates. Denoising was performed by removing the residue after the decomposition with the EMD, the Ensemble EMD (EEMD) and the Complete ensemble EMD with adaptive noise (CEEMDAN). The VMD and EWT were also used for denoising. The PhNR was measured as the amplitude, relative to the prestimulus baseline, of the first negative going trough after the b-wave. The coefficient of repeatability (CoR) was calculated to represent 95% of the test-retest differences for the PhNR between sessions.

Results: Denoising with all of the varieties of the EMD, the VMD and the EWT improved the test-retest repeatability of the PhNR. The best results were achieved with the CEEMDAN algorithm, which for the PhNR amplitude improved the coefficient of repeatability relative to the effective dynamic range from 20.2% to 8.3%. The EMD and its variants was easier to implement than the EWT or VMD.

Conclusions: The EMD can be used to remove low frequency baseline drift and oscillation from the ERG and improve the test-retest repeatability of PhNR measurements. This technique could be deployed for a range of visual and other electrophysiological tests.

2.2 Introduction

The electroretinogram (ERG) is a time domain signal. That is, it is a time series of amplitudes, each point representing the voltage at a time point equal to the index times the reciprocal of the sample rate. Like most biosignals, the ERG is contaminated with noise. Sources of the noise can be exogenous such as powerline noise, radiofrequency interference from transmitters or switch mode power supplies, pulses from high voltage switching and electrical induction from magnetic fields. Most of the preceding types of noise are periodic or at least have a regularity. There is also noise that arises from the equipment itself including the Johnson–Nyquist noise of resistors in the electronics (Johnson 1928, Nyquist 1928). Flicker or 1/f noise (Hooge 1976) is also particularly problematic in Metal Oxide Semiconductor Field Effect Transistor (MOSFET) and Junction Field Effect Transistor (JFET) semiconductor amplifiers. These are usually deployed in biomedical preamplifiers as the input impedance is extremely high. Due to the very small amplitudes of the signals in visual electrophysiology, all of these sources of noise are highly significant. These sources of noise are highly stochastic.

The other major source of noise in the ERG is endogenous noise – arising from chemical or movement changes at the body-electrode interface, muscular (electromyogram) artefact or other bioelectric activity detected by the differential pair of electrodes. Attention to the design of the electrophysiology system is essential in minimizing the noise – in particular matching of differential semiconductor pairs, optimizing circuit paths and isolating power supply components from the input headstage (Bretschneider and De Weille 2018).

Even with the best design, filtration of the signal to reduce the noise and increase the signal to noise ratio is needed. A combination of high pass and low pass filters combined to form a band-pass filter are a part of the International Society for the Clinical Electrophysiology of Vision (ISCEV) standard for electroretinography (McCulloch, Marmor et al. 2015) which simply states that the corner frequencies should be less than 0.3 Hz and greater than 300Hz. All modern electrophysiological systems implement the majority of this filtration in software with minimal hardware filtration ahead of the Analog to Digital Converter (ADC). For such filtration to be effective, there must be good linearity of the ADC, fast settling time and sufficient digital resolution to avoid saturation. Resolutions of 24 or 32 bits are typical. The electroretinography standard published by the International Society for the Clinical Electrophysiology of Vision (ISCEV) (McCulloch, Marmor et al. 2015) states under the heading “Real-time ERG display” that the waveforms should be displayed during testing. In practice, this is always near-real time and not true real time. A small delay of fractions of a second is acceptable and this means that non-causal filters may be used, provided that the realization of the technique is sufficiently fast. More elaborate and computationally expensive filters can be used off-line in *post hoc* analysis. Filtration can be realized in the time domain or frequency domain but care must be taken to minimize changes to the shape, amplitude and timing of the responses. Zero-phase filters, although non-causal, are encouraged for this purpose. The use of notch filters for example to remove powerline noise is discouraged. If the spectrum of the noise is very similar to that of the underlying signal, filtration alone will be ineffective in increasing signal to noise ratio.

Post filtration, the method usually employed to increase signal to noise ratio is signal averaging. This technique is simply summing the time-synchronized signals at each time point and obtaining an arithmetic or trimmed mean. The technique assumes that the measured signal is an additive superposition of actual signal and noise i.e. in discrete time as:

$$y(i) = s(i) + n(i) \qquad 2\text{-}1$$

Further assumptions are (Cerutti, Bersani et al. 1987) that the signal is stationary over successive trials, the background noise is a zero-mean random signal and that it is uncorrelated with the actual signal. Under these conditions it is shown that the noise reduces as $\frac{1}{\sqrt{n}}$ in respect to the averaged signal.

Stationarity is a large assumption. It refers to the fact that neither the structure of the signal nor the structure of the noise varies over time. One way to assess this is to examine the time varying mean or mean square value or to look at the time varying frequency structure of the signal. Cohen (Cohen and Sances 1977) looked at this for the EEG – relevant to visual electrophysiology through the visual evoked potential (VEP). He found that the EEG can be considered stationary for periods of up to 14s. Porciatti (Porciatti, Sorokac et al. 2005, Porciatti and Ventura 2009) found that the pattern ERG (PERG) amplitude diminished with repeated stimulation over 2 minutes. These findings show that it can be difficult to deal with the problem of noise by simply increasing the number of sweeps.

2.3 Adaptive filter theory

Relaxing the stationarity constraint allows the consideration of filters that adapt to each sweep. One early implementation of this principle was that of Yu (Yu and Mc Gillem 1983) where he described the estimation of the optimum linear time-varying filter. Returning to equation 2-1, the objective is to estimate the signal s(i). This is the best that we can do. We cannot measure the underlying process directly. The estimator (the output of our electrophysiology system) can be described as:

$$\hat{s}(k) = \sum_{i=1}^{N} h_k\,(i)x(i) \qquad 2\text{-}2$$

Where x(i) is the signal measured over i samples and k sweeps. $h_k(x)$ is the filter for the kth sweep.

In matrix form this results in:

$$\hat{s} = Hx \qquad 2\text{-}3$$

With **H** the estimator impuse response matrix of size NxN, chosen to minimize the mean square error and the filter assumed to be linear, time varying and noncausal (new data points can affect the signal in the past). This formulation is known as the Wiener filter. A related but simpler approach has been described (Davila and Mobin 1992) by applying a weighting to each sweep based on an estimate of the signal to noise.

2.4 Latency-Dependent Averaging

Woody (Woody 1967) described a technique of dealing with evoked potentials where the latency varied. The method involved detection of the wave by cross correlation with a template formed from the first recorded sweep and then adjustment of each subsequent sweep before averaging. In that paper, the technique was applied to auditory evoked potentials where the signal was recorded directly from the cortex.

A variation of the technique was developed by Hood and Li (Hood and Li 1997) in which a stereotypical waveform was used as the template basis. In that paper, the template was stretched from a fixed start point. Wright compared this with a sliding technique for the isolation of the multifocal ERG (mfERG) (Wright, Nilsson et al. 2008) and for the pattern visual evoked potential (PVEP) (Wright, Nilsson et al. 2011) and found the sliding technique to be better in each case. Described as template averaging, the match was found over the sliding window by maximising the cross correlation with a template based on ISCEV standards. A weighted average was then constructed of the sweeps using the correlation coefficient as the weighting factor. This method is often used for mfERGs. There are no research publications exploring the technique to enhance the detection of the PhNR in glaucoma.

2.5 The Kalman Filter

Kalman (Kalman 1960) extended the idea of the Wiener filter to nonlinear and non stationary situations. The technique is used extensively in navigation and has found utility in self driving cars (Cuenca, Zhan et al. 2019). For a time series, the signal can be thought of as a series of location observations of an object, each with an uncertainty described by a

Gaussian distribution. The velocity is inferred from the difference between the elements in the vector. The velocity is a hidden variable that also has uncertainty. For each subsequent observation, the prediction for location and velocity are updated and the probability function is also updated. In brief, the filter uses a state space equation to update the prediction of the next value assuming that the process is autoregressive (Arnold, Milner et al. 1998). The filter can be used as a denoising and smoothing technique by running multiple times forward and backward (Lee, Han et al. 2010). A simple implementation is the three pass fixed interval smoother. The Kalman filter is run forward (i.e. with measurement time increasing) and again backwards (time decreasing). The forward estimate and covariance of estimation uncertainty for each point is saved. These two estimates are combined in a weighted average with the weighting based on the covariance. This smoothing technique has been used extensively for the electroencephalogram (EEG) but there are no published works utilizing it for filtering the ERG. Some interesting work has recently been published (Garcia, Lavieri et al. 2019, Garcia, Nitta et al. 2019) using the Kalman filter to predict the disease trajectory of patients with glaucoma or ocular hypertension. The authors found that machine learning algorithms based on the Kalman filter could predict the disease as measured by mean deviation (MD), pattern standard deviation (PSD) and intraocular pressure (IOP) five years into the future.

The assumptions of linearity and Gaussian probability used by the Kalman filter have been relaxed in the Extended Kalman Filter (EKF) (Lewis, Xie et al. 2017) and the Unscented Kalman Filter (UKF) (Julier and Uhlmann 1997). The EKF uses the Jacobian matrix in its update and prediction points to linearize the non-linear function of the underlying process. In the UKF, a set of sigma points is used to empirically estimate the characteristics of the posterior distribution.

2.6 The Empirical Mode Decomposition and related techniques

2.6.1 The Empirical Mode Decomposition (EMD)

The empirical mode decomposition is a mathematical technique proposed by Huang (Huang, Shen et al. 1998) to deal with the analysis of non-linear and non-stationary data. The algorithm decomposes a signal into a number of components known as intrinsic mode functions (IMFs) by iteratively finding the highest oscillating component, removing that as IMF1, finding the next highest oscillatory component as IMF2 and so on until all oscillations have been removed. The signal that is left is known as the residue. The empirical nature of the decomposition is that the algorithm makes no assumptions about the number of oscillating components, nor does the oscillation itself need to have any regularity. The IMF is defined (Blanco-Velasco, Weng et al. 2008) as a function with equal numbers of extrema and zero crossings and the envelope (all local maxima and minima) being symmetric with respect to zero. It is analogous to the harmonics of a Fourier series. The technique has been used to remove both high frequency and low frequency noise from the electrocardiogram (ECG) (Blanco-Velasco, Weng et al. 2008), EEG (Zhang, Wu et al. 2008), electromyogram (EMG) (Andrade, Nasuto et al. 2006) and the same principles have been applied to a variety of other fields, for example the forecasting of commuter flows in public transport (Wei and Chen 2012). The simplest approach to the removal of low frequency baseline noise is to remove the residue from the original signal (Lee 2016). More formal techniques have been described to enumerate the number of intrinsic mode functions (IMFs) that contain the noise and remove those (Blanco-Velasco, Weng et al. 2008).

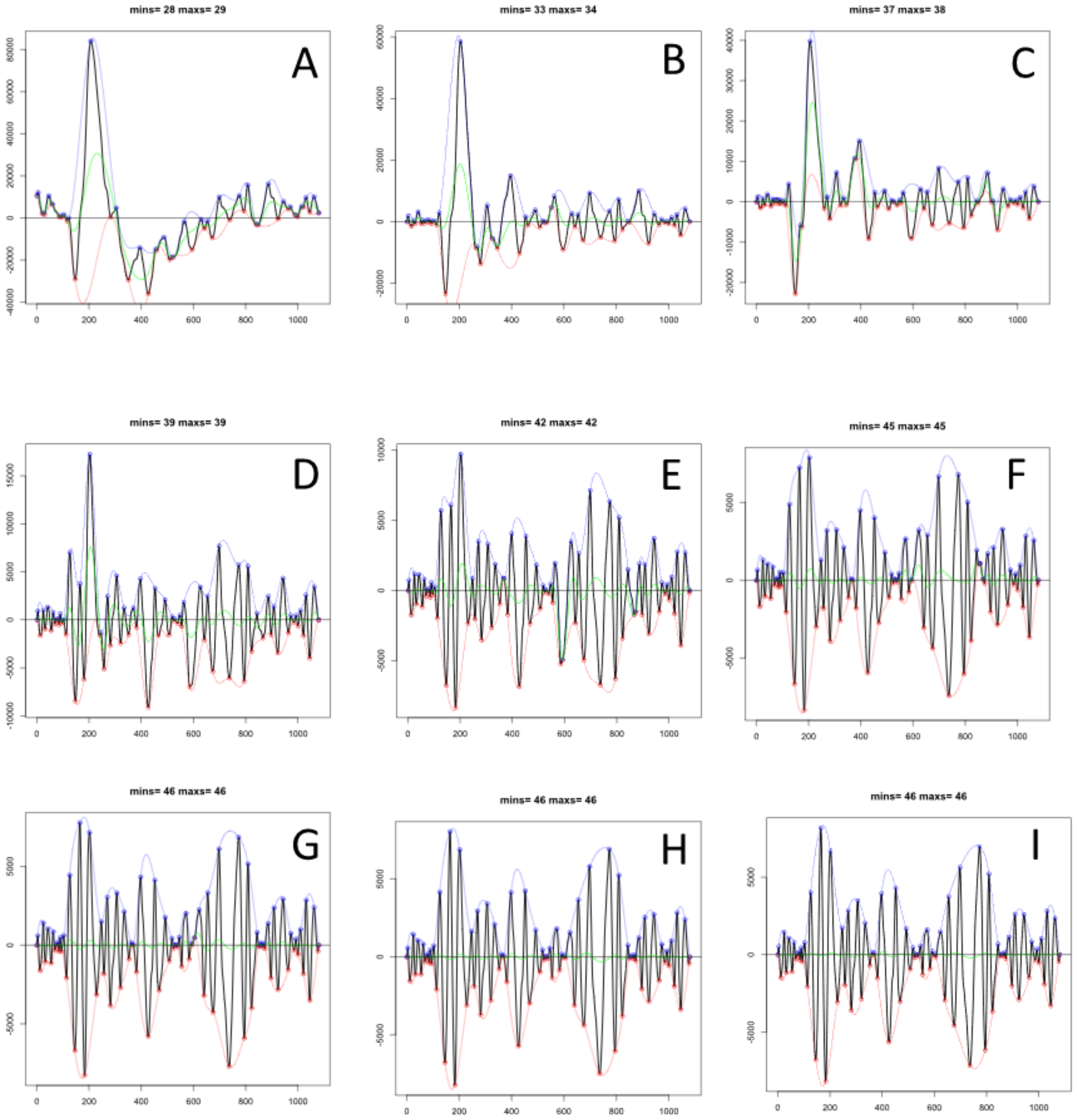


Figure 2-1: The Empirical Mode Decomposition-Sifting The panels show in order the process of extracting the first Intrinsic Mode Function (IMF) from the signal. Full discussion is in the text.

The process of sifting and extracting the first IMF is illustrated in Figure 2-1 with a real ERG from this study. In panel A the ERG is shown and the concave up and down turning points are identified. The start and end of the sequence are also considered to be both maxima and minima. A spline is used to interpolate between the local maxima (illustrated in blue) and the local minima (illustrated in red). The original algorithm (Huang, Shen et al. 1998) specified a cubic spline which is also used in the libEEMD implementation (Luukko, Helske et al. 2016). Other types of splines have been explored including local cubic splines (Chui and van der Walt 2015) and rational Hermite interpolation (Li, Xu et al. 2015). The splines interpolating the maxima and minima are called the envelopes. "Good maxima" are considered those that are positive (>0) and "good minima" those below zero. In panel A there are 5 bad maxima and 5 bad minima. The next step is to find the mean of the upper and lower envelope function – illustrated in green. This is then removed from the signal. This subroutine is called sifting. The sifted signal is then considered as the input to another sifting. In this example B is sifted to C. The process continues until a stopping criterion is reached. This is the first of either a maximum number of siftings or, more commonly, the S-number is reached with no bad minima or maxima. These are both tuneable parameters. The S-number is incremented when as the result of a sifting, the number of good maxima and minima do not change. In this example, the S-number is the default of 4. In this example, the sifting finishes at I. This signal is now considered the first IMF.

The first IMF, IMF1 is then removed from the signal and the process starts again with the signal that is left. In doing so, the highest oscillating component has been removed - in this case looking much more like a Gabor function than the cosine that is extracted by a Fourier series.

The full decomposition is illustrated in Figure 2-2. The number of IMFs extracted is a tuneable parameter and defaults to $\log_2 S$ where S is the signal length. It is apparent from this example that the decomposition is complete by IMF 6 and that IMF 7 and IMF 8 have not added anything to the decomposition. What is left over – the residue is extracted as the last element of the decomposition. In this work, the residue is removed from the original signal as a zeroing and detrending process. It will have a maximum of one turning point.

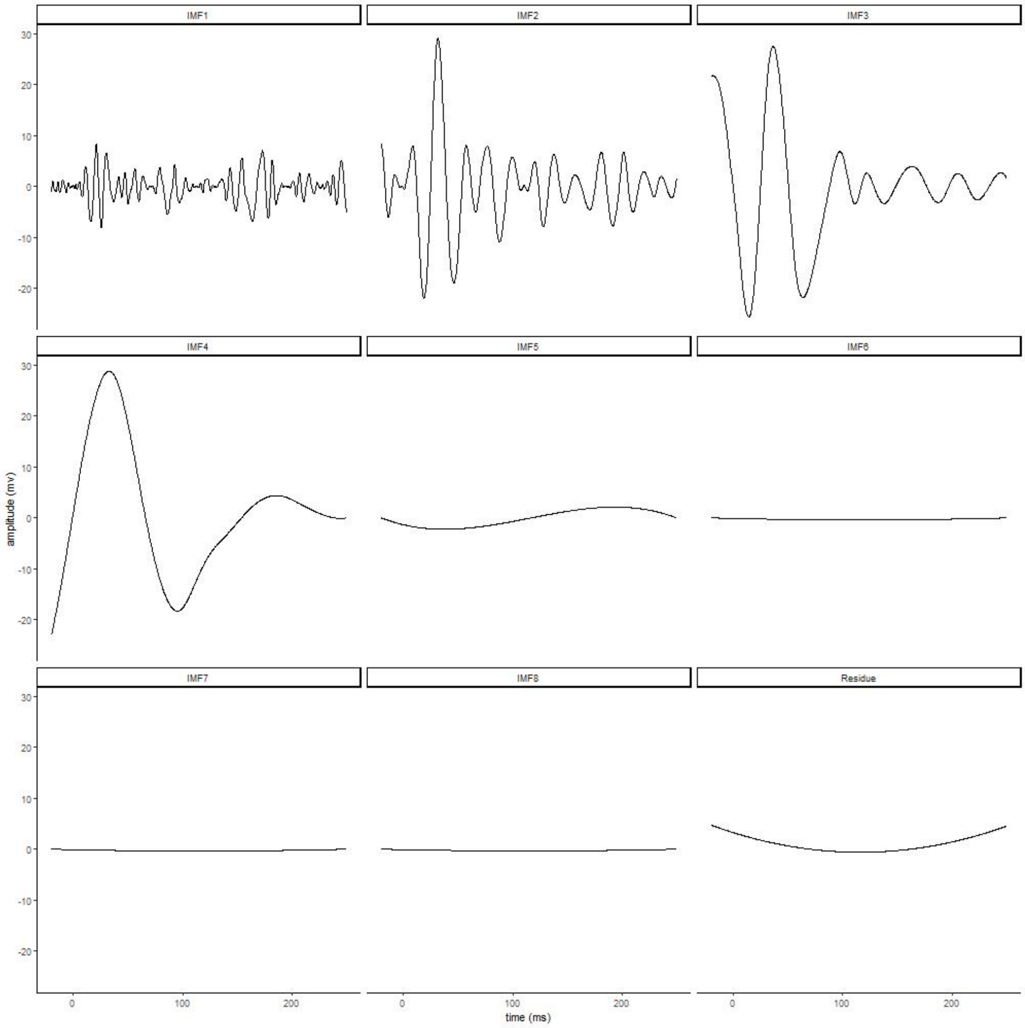


Figure 2-2: Full EMD decomposition from earlier example

2.6.2 The Ensemble Empirical Mode Decomposition (EEMD)

The performance of the EMD in the presence of intermittent frequency components can be problematic with the phenomenon of mode-mixing (Deering and Kaiser 2005) where the extracted IMF switches over from one oscillatory component to another. This reduces the inferential power of the decomposition when looking at the individual IMFs. Wu (Wu and Huang 2009) proposed a solution to the problem by composing an ensemble of signals from the original one by adding white noise of varying amplitude to the signal, performing the EMD on each signal within the ensemble and deriving the ensemble average. Chang (Chang 2010) found the EEMD superior to the EMD for decontamination of ECG signals when evaluated by mean squared error of the reconstructed signal compared to the original signal prior to noise addition. The technique had minimal effect on phase and outperformed traditional time domain filters such as the Butterworth (Butterworth 1930) infinite impulse response (IIR) filter. This variation of the EMD is illustrated in Figure 2-3.

For clarity, the figure illustrates only 2 ensemble members. This is a tuneable parameter with the default at 250. In this example, the two ensemble members are in the first two columns and the average is in the third column. Each ensemble member is created by adding Gaussian white noise to the original signal (A, B). The noise strength (measured by standard deviation of the Gaussian) is a tuneable parameter and is set to a proportion of the standard deviation of the input signal. The default is 20 percent.

The noise is shown in panels C and D and the resultant noise plus signal is shown in E and F. The EEMD is then just the average of the IMFs for each ensemble member. That is IMF1 is the ensemble average of all of the IMF1s. This is illustrated in panel I. Similarly, IMF2 for the ensemble members is shown in panels J and K and the average IMF2 in L. Finally, the residue is the average of all of the residues.

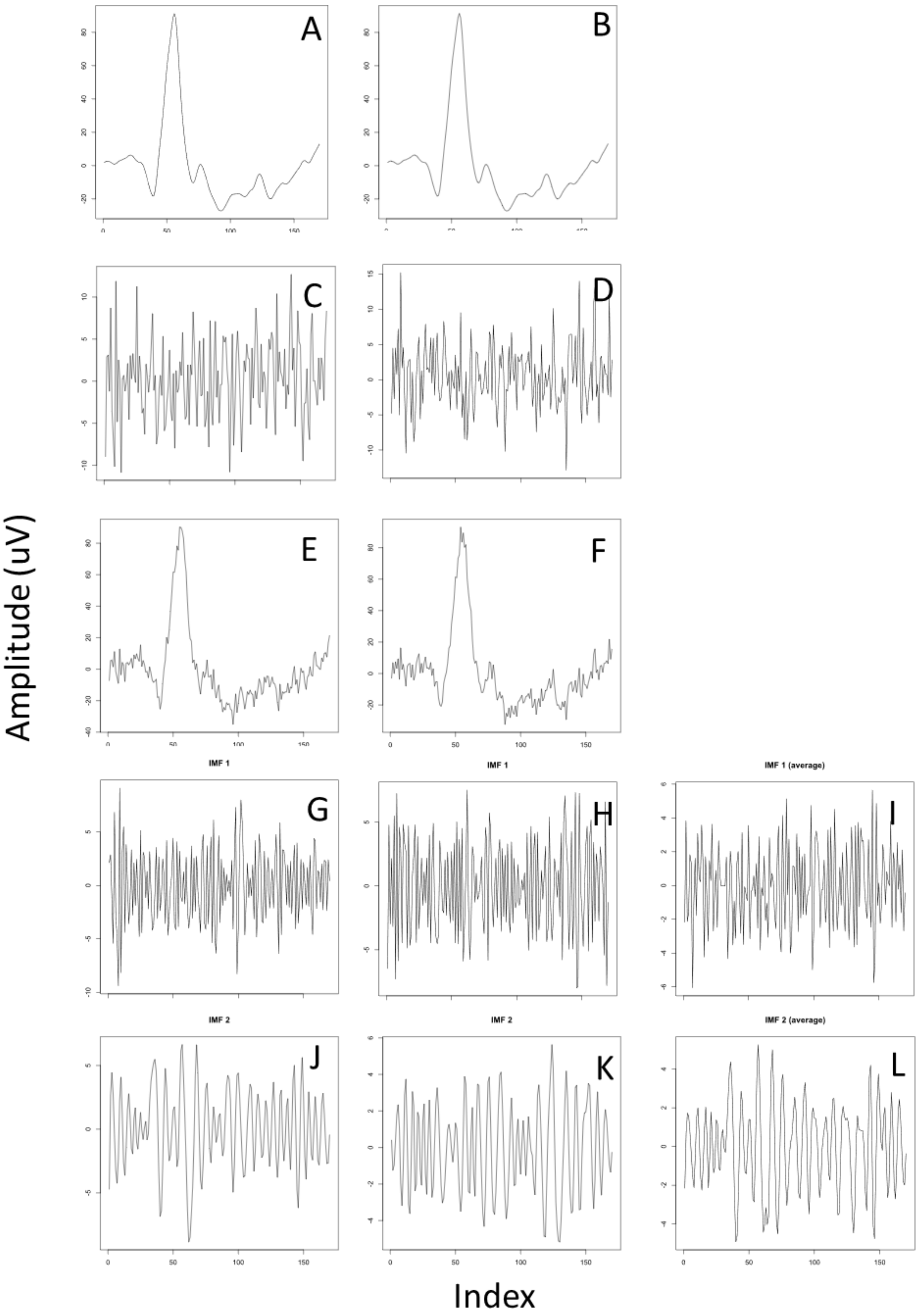


Figure 2-3: The ensemble empirical mode decomposition (EEMD). The first and second columns represent two of the ensemble members with the original signal found in panel A and B. Random Gaussian noise in panels C an D are added to yield E and F. The empirical mode decomposition is then performed on the noisy traces of each of the ensemble members and the intrinsic mode functions (IMFs) are found as the averages of the ensemble IMFs – The first two here shown at I and L..

2.6.3 The Complementary and Complete Empirical Mode Decompositions

An issue with EEMD can be that the number of IMFs is different amongst different members of the ensemble of signals, making it difficult to average. More specifically, at each stage, a residue is computed from each member of the ensemble but there is no connection between the realizations. This leads to the decomposition not being complete (Colominas, Schlotthauer et al. 2014) in that, unlike the EMD, the original signal cannot be reconstructed from the decomposition. The next development, complementary EEMD added the white noise in complementary pairs to the signal so that the white noise was eliminated from the final residue (Yeh, Shieh et al. 2010). Residual white noise did, however persist in the individual IMFs so the technique Complete Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) was evolved from the CEEMD (Torres, Colominas et al. 2011). In the CEEMDAN technique, each unique residue is determined at each stage of the ensemble decomposition before starting the next step.

The technique is illustrated in Figure 2-4. In this diagram, the original signal is shown in panels A and B. Like the EEMD, noise realizations (C and D) are added to the signal to create the noisy ensemble members (E and F). IMF1s are extracted (G and H) and averaged like the EEMD – the ensemble average IMF1 is shown in I.

The algorithm then differs significantly from the EEMD. Independently, the IMF1 is extracted and removed from the original noise. In this example, the IMF1 is extracted and then removed from C to leave residual noise H. This is similarly done for D leaving K. The derivative noise is then added to the original signal minus the ensemble IMF. In the first column this is A–I+J shown at L and for the second column B-I+K shown at M. The EMD is then performed on these signals in the same way as for panels E and F. This then continues until the stopping criterion is reached.

From this diagram, it can be seen that the sum of the ensemble averages in the third column will add up to the average of the noisy ensemble members and since the noise is random and with zero mean, it will add up to the original signal. This is the principle of “completeness” that is not present within the simpler EEMD.

Computationally, the EEMD and CEEMDAN are much more expensive than the EMD. The EEMD requires more than the ensemble size times more cycles (default 250) compared to the EMD. The CEEMDAN requires double the number of siftings compared to the EEMD as the algorithm is done on signal and noise at each iteration. Even so, the process is fast and realizable in seconds for a signal of 150 to 500 samples in length.

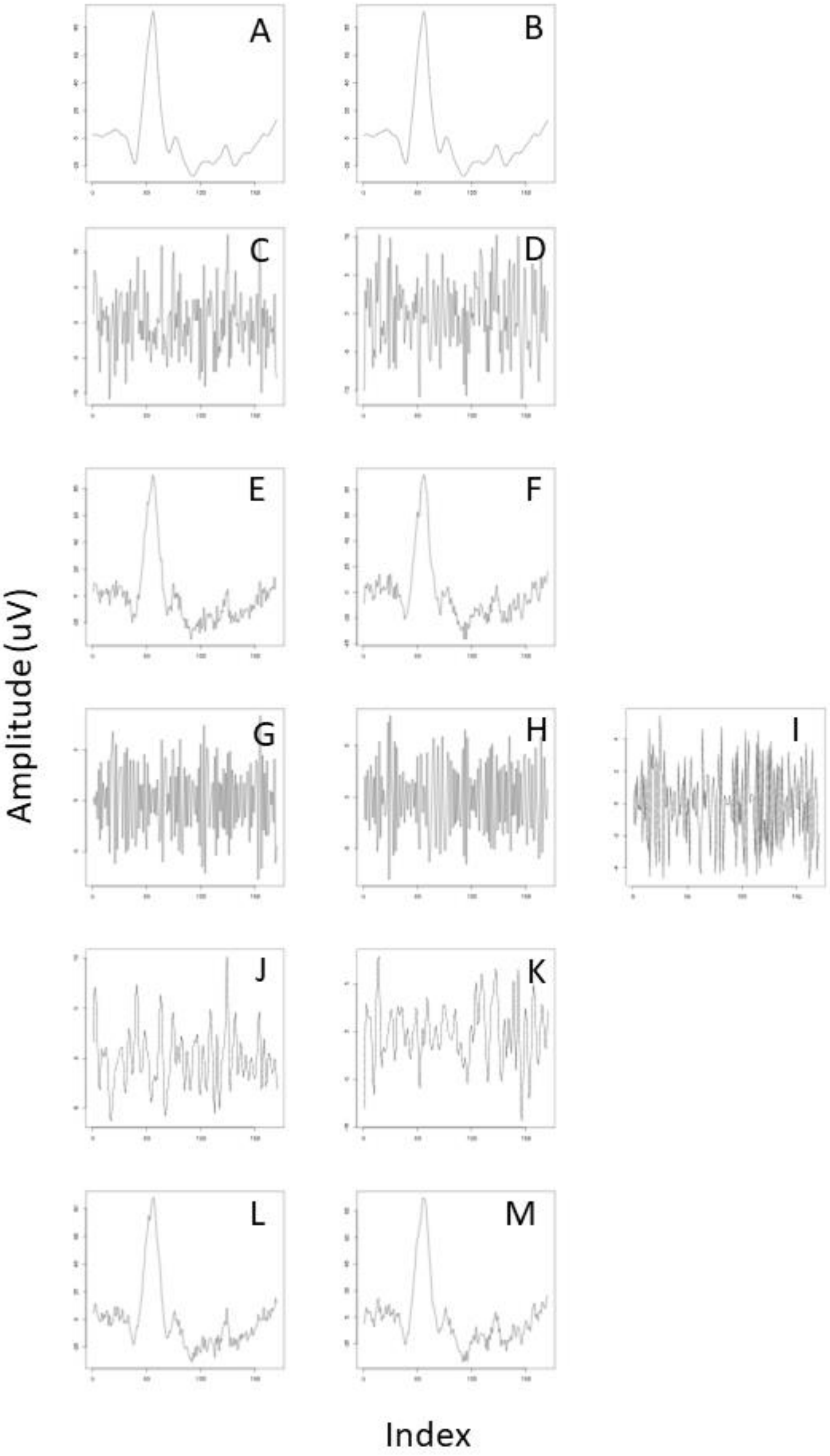


Figure 2-4: The Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN). Explanation of the process is in the text.

2.7 Variational Mode Decomposition

The Variational Mode Decomposition (VMD) was described by Dragomiretskiy and colleagues (Dragomiretskiy and Zosso 2013). It has found utility in a number of disparate areas including gas pipeline leak detection (Xiao, Li et al. 2018), wind power forecasting (Naik, Dash et al. 2018) and extraction of heart rate from photoplethysmographic signals (Sharma 2019). It sought to be a more robust decomposition than the EMD. Like the EMD, the algorithm decomposes a real valued time series into a linear sum of subsignals. The subsignals are also termed intrinsic mode functions (IMFs). These IMFs are different to those of the EMD and are chosen to have a compact bandwidth around a centre frequency. The optimization problem is formulated to minimize the bandwidth of the predetermined number of IMFs.

The constrained optimization problem (Dragomiretskiy and Zosso 2013) is written to minimize with respect to s_k and ω_k

$$\sum_{k=1}^{K} || \, \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * s_k(t) \right] e^{-j\omega_k t} || \qquad 2\text{-}4$$

Such that $\sum_k s_k(t) = r(t)$

Where k = 1..K is the index into the IMFs, $\delta(t)$ is the Dirac function, $s_k(t)$ is the kth IMF, ω_k is the centre frequency of the kth IMF and $r_k(t)$ is the original real valued signal.

In practice, the problem is solved with a Lagrangian multiplier and a quadratic penalty factor as this makes the problem unconstrained (Zhao, Wu et al. 2020). Within the Matlab and R implementations of the algorithm, the configurable parameters are the number of IMFs to recover, the tolerance for convergence and the penalty factor known as t.

Figure 2-5 shows an example of VMD denoising. Panel A shows the example ERG signal – both raw and denoised by the VMD. The CEEMDAN and EWT denoising is also shown for comparison. In this example, the default settings of alpha = 2000 and a tolerance of 1x10^{-6}. Different values were tested and minimal differences were observed. Ten components were specified for the decomposition. Panel B shows the full decomposition of the signal. The lowest frequency component is termed the “mode DC” (MDC) corresponding to the low level drift and wander. In this decomposition, the MDC is the first component determined so that the number of components extracted does not affect the MDC. Panel C shows the frequency domain representation of the original signal and panel D shows the frequency domain representation of the extracted components.

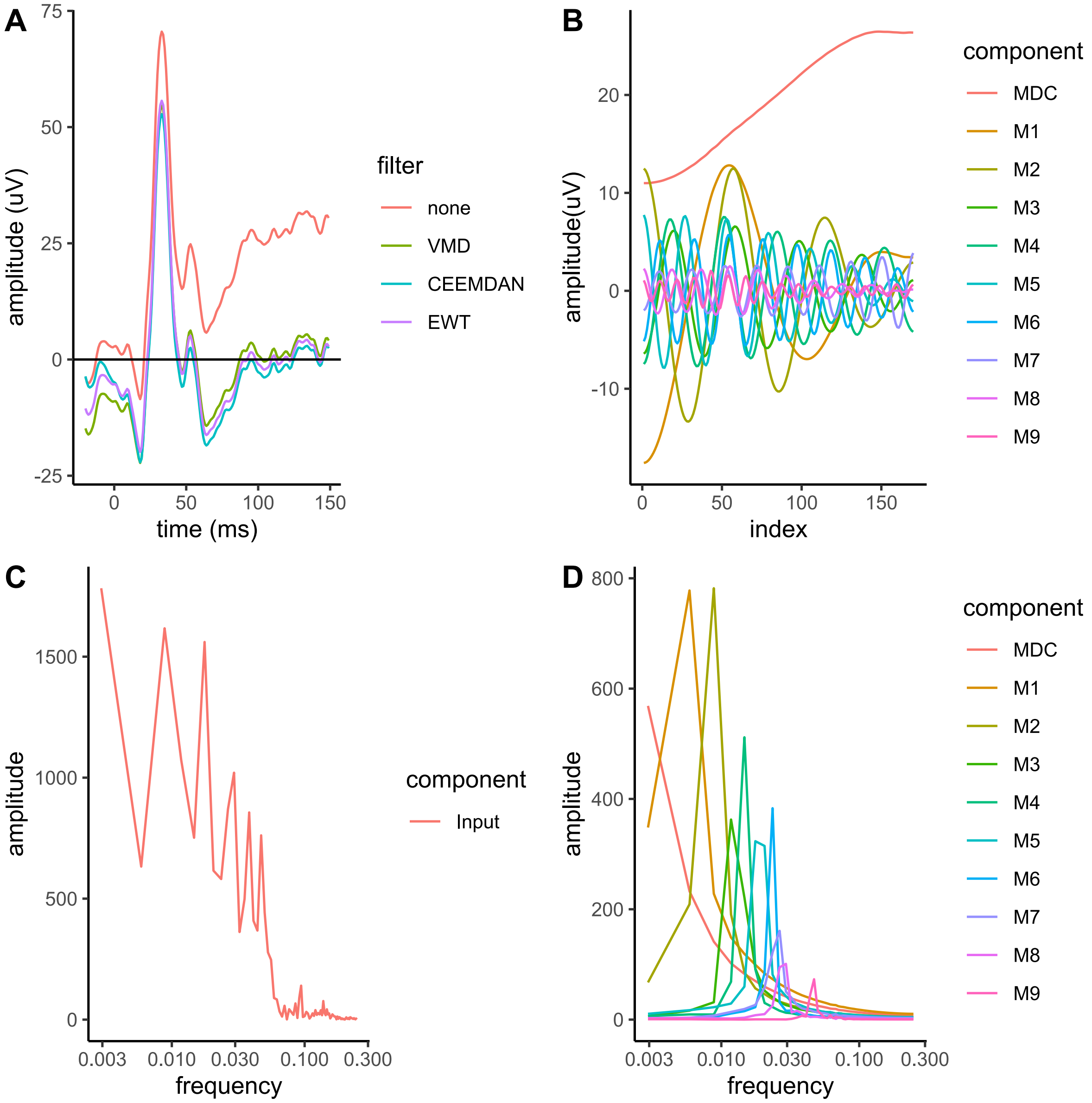


Figure 2-5: Variational mode decomposition (VMD) denoising example. Panel A shows the raw electroretinogram (ERG) trace and the denoised trace by VMD and by Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and Empirical wavelet transform (EWT). Panel B shows the time domain VMD subsignals – The MDC is the DC component removed for the denoising. Panel C is the frequency domain representation of the original signal and panel D is the frequency domain representation of the subsignals shown in B.

2.8 The Empirical Wavelet Transform

The empirical wavelet transform (EWT) was described by Gilles (Gilles 2013) as an alternative to the EMD. The idea of the algorithm is to perform a wavelet decomposition on the signal with and empirically selected bank of wavelets. The source code is available (Gilles 2021). A Python version is also available (Carvalho, Moraes et al. 2020).

The wavelets used for the EWT are derived from the Meyer Wavelet (Meyer and Roques 1993). This wavelet family is defined in the frequency domain. Figure 2-6 shows the wavelet and scaling function in both the frequency domain and the time domain.

The EWT algorithm begins by performing a Fourier transform on the input signal. Then, with N wavelets specified, the algorithm finds the boundaries in the frequency domain by finding the local minima between the N largest local maxima.

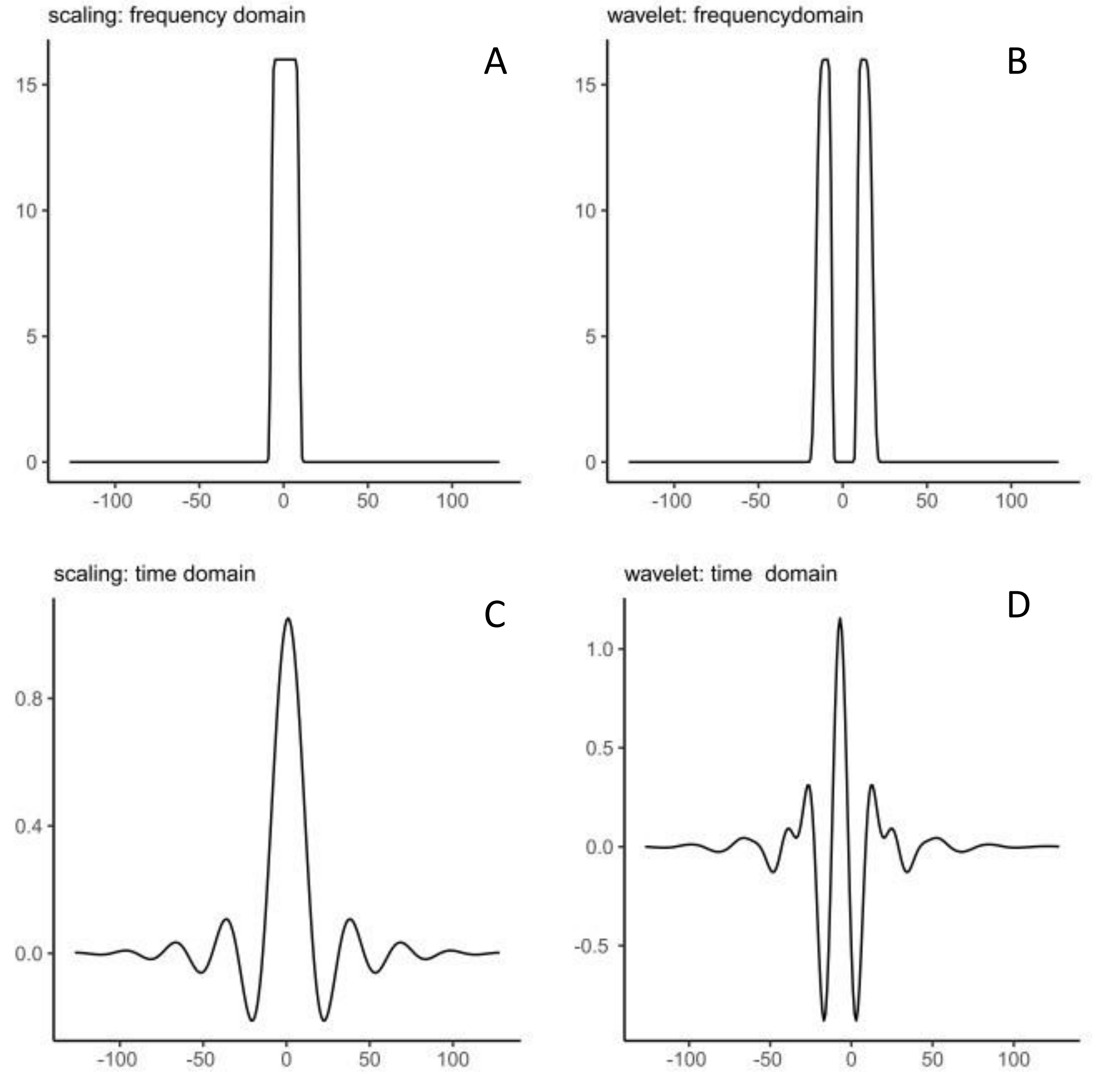


Figure 2-6: The Meyer Wavelet.

Figure 2-7 shows the initial process for the sample waveform from Figure 2-5. Panel A shows the absolute amplitude of the fast Fourier Transform (FFT) of the raw signal from the upper left pane of Figure 2-5. The boundaries have been identified and marked. Panel B shows the EWT filter bank in the frequency domain with the same normalized frequency axis. In both cases, only the positive frequencies are shown.

The next phase of the algorithm is to apply the wavelet transform. Not mentioned in the original paper but evident from the source code, the algorithm extends the signal in both directions by mirroring to reduce boundary effects. This is shown in Figure 2-8. The filter bank shown in the lower pane of Figure 2-7 is generated by the algorithm to be a matching length to the extended signal.

The frequency domain wavelets are applied to the Fourier transformed mirrored signal by simple point by point multiplication of the vectors. The decomposition is then recovered by the inverse Fourier Transform. Figure 2-9 shows the example waveform denoised with the EWT, VMD and CEEMDAN. Although the b-wave, i-wave and photopic negative response (PhNR) are very similar between the denoising techniques in this example, the baseline is quite different with only the CEEMDAN not affecting the initial baseline reading. As the PhNR is measured relative to the baseline, a difference in the magnitude of the PhNR would be expected between the different denoising techniques.

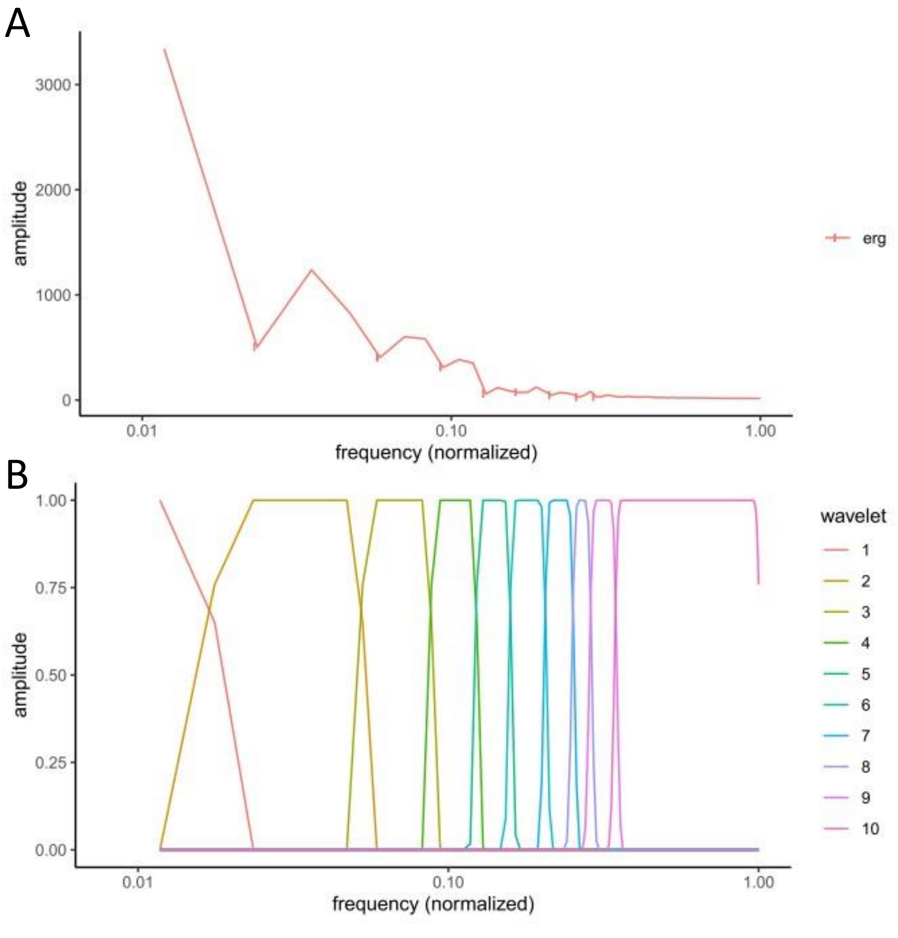


Figure 2-7: Wavelet selection for example ERG.

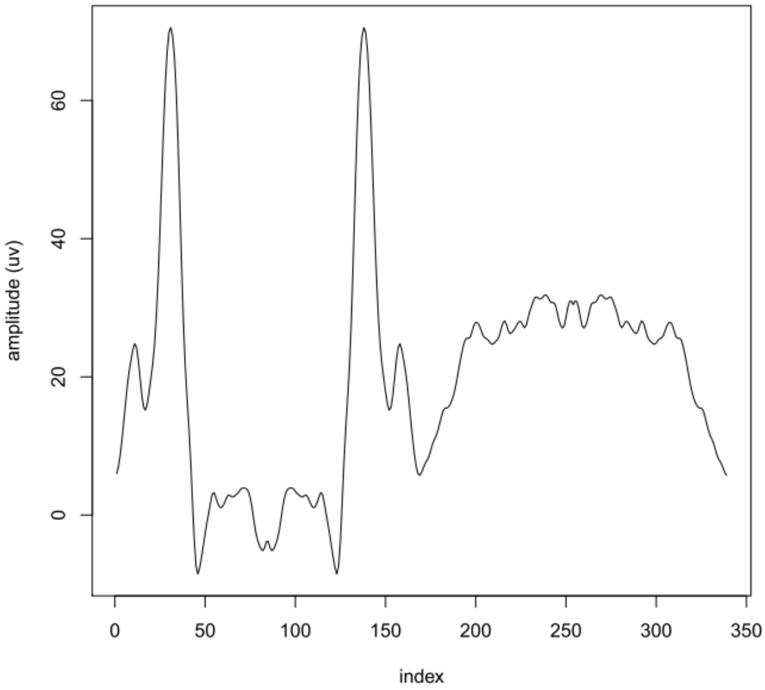


Figure 2-8: Mirrored signal. This figure shows the first step of the algorithm which is to extend the boundaries of the signal in each direction by mirroring half the signal in each direction. This has increased the length of the original signal from 170 samples to 350 samples. After the filter bank is applied, the components are trimmed back to the original size.

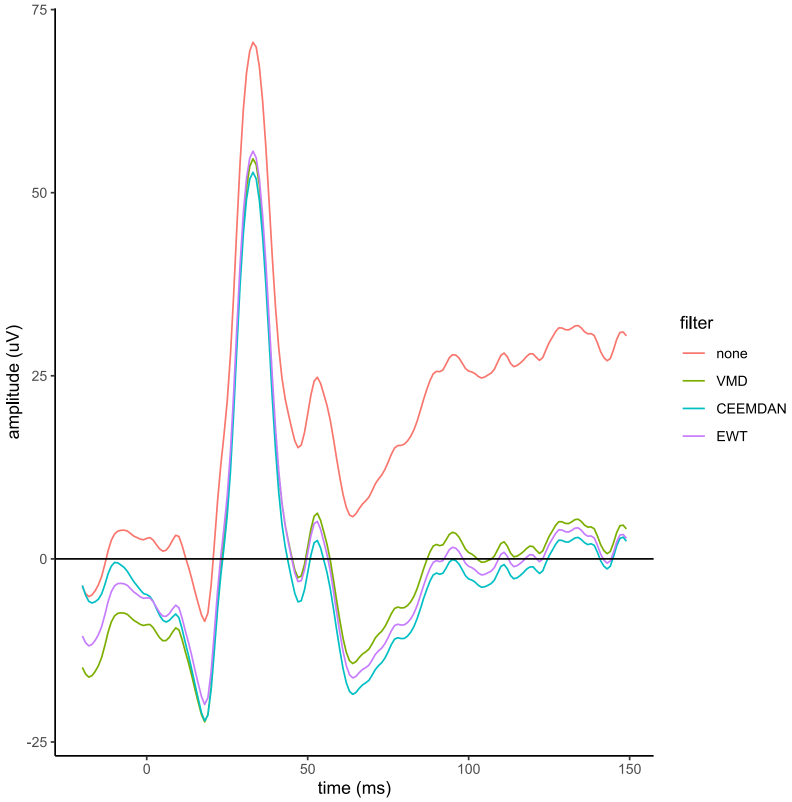


Figure 2-9: Empirical Wavelet Transform (EWT) denoising. This figure shows the original signal and the three denoising techniques. VMD= variational mode decomposition, CEEMDAN = complete ensemble empirical mode decomposition with adaptive noise, EWT= empirical wavelet transform.

2.9 Statistical Evaluation of Repeatability

Ideally, a clinical test should give the same result when the underlying condition has not changed. In the case of the PhNR, this means that if the underlying glaucoma disease has not changed, when repeated on an eye, the PhNR voltage should be the same. There are a number of ways to measure the extent to which the two measurements are similar. (Bartlett and Frost 2008). The term agreement is often used to describe the way in which two different measures of the same phenomenon align. For example, comparing retinal nerve fibre layer thickness with the mean deviation of the visual field (Cheng, Laron et al. 2007). A frequently used statistical test for agreement is the Pearson product moment correlation coefficient which is defined in the range -1 to 1. There are well documented assumptions in the use of this test (Havlicek and Peterson 1976):

- The variables are continuous
- The relationship between the variables is linear
- The joint distribution of the variables is a bivariate normal distribution
- Pairs of observations should be independent
- The observations should be free of outliers in either variable (Kim, Kim et al. 2015)

Although the effect of non-normality is small, in the presence of any significant violation of the assumptions, the Spearman rank correlation coefficient is generally used.

An approach to analysing the effect of measurement by different raters on the same subjects is the intraclass correlation coefficient (ICC) (Shrout and Fleiss 1979). This statistical test is defined on the range of 0 to 1 and has been used in the context of electroretinogram measures (Fredette, Anderson et al. 2008, Resende, Sanvicente et al. 2019). Interpretation of the level of reliability can be difficult and subjective but ICC values above 0.8 have been considered excellent reliability (Resende, Sanvicente et al. 2019). There has been criticism of the use of ICC measurements in scenarios outside of the original purpose. Significance tests of two different ICC measures are not well established and the measure itself can suffer significant bias.

In the setting of evaluating whether a test performed two times on the same subject yields the same result, the Coefficient of Repeatability (CoR) measure can be used (Vaz, Falkmer et al. 2013). This test, measured in the same units as the measurements, is related to the Limits of Agreement of the Bland Altman approach (Bland and Altman 1986). The CoR determines the maximum absolute difference between pairs of measurements that comprises 95% of the cases. These differences are usually assumed to have a folded normal distribution (Leone, Nelson et al. 1961). A folded normal distribution is the distribution of absolute measurements from a population that has a normal distribution. Where the mean is zero, the distribution is known as the half normal distribution (Daniel 1959). The quantile for 95% of the half normal distribution is equal to the quantile for 97.5% of the normal distribution; approximately 1.96. The CoR is therefore found as $1.96\sqrt{2}\,s$ where s is the standard deviation of the absolute differences of the pairs of observations. This measure is not applicable in the presence of a bias (non-zero mean) of these measures or a non-normal distribution.

2.10 Methods

Participants

The study was an approved study by the Royal Victorian Eye and Ear Human Research Ethics Committee. In this study, participants were from staff of the Centre for Eye Research Department who self-reported as having no ocular or systemic disease.

PhNR Recording

The pupils of both eyes of the participants were dilated using 1% tropicamide drops. Preadaptation was with the background room light measured at of 0.92 cd/m2. All electrophysiology recordings were performed using a commercial electrophysiology system (Espion E2/ColorDome; Diagnosys LLC, Lowell, MA, USA). Before the stimulus acquisition, preadaptation was performed to the blue background for one minute.

Red (peak wavelength, 635 nm) flashes of 4ms duration at 1 Hz of 2.25 cd s /m^2 were presented with a Ganzfeld on a blue (peak wavelength 465 nm) background of 10 cd/m^2. Dawson-Trick-Litzkow (DTL) (Dawson, Trick et al. 1979) fibre electrodes were used with a gold cup electrode at the lateral canthus as a reference and another one on the forehead as a ground.

The stimulus strength was informed by a pilot study which showed this to be a good trade-off between amplitude size and reliability. Smaller amplitude flashes produced a lower PhNR amplitude and larger amplitude flashes increased the orbicularis artefact.

Ten sweeps were collected with band pass filtering from 0.15 to 100 Hz. Sweeps were automatically rejected at amplitudes above 100uV. These filter settings were selected as the same protocol was used for photopic full field electroretinograms in the hospital electrophysiology laboratory and it was anticipated that the methods demonstrated in this thesis could be used in a *post-hoc* method.

Mathematical analysis

For the latency-dependent averaging, the template was derived from empirical data, which was used to fit a Weibull probability distribution function of the form of equation 2-5. With a k > 1, this function is initially convex rising to a turning point and is then concave with an inflexion point. This function was chosen to approximate the hypothesized underlying physiological process: This process driving the PhNR is assumed to have a fixed onset point which overlaps the offset of the bipolar response driving the b-wave. At onset, this should therefore have a negative gradient. The process will peak and then decay with a tail approximating an exponential decay. The parameters λ and k were fitted in R to the empirical data over a time epoch of 90ms using nonlinear least squares. This resulted in k=2.029 and λ=42.385. The function was negated by multiplying by negative one.

$$f(x;\ \lambda, k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} \quad x \geq 0 \qquad \text{2-5}$$

Where k is the shape parameter and λ is the scale parameter.

The probability distribution was described by Weibull (Weibull 1951) who suggested that it could be used, amongst other things, for fitting the statures of adult males in the British Isles, the fibre strength of Indian cotton and the size distribution of fly ash. In his formulation of the distribution Weibull observed that any cumulative probability function F(x) can be written as

$$F(x) = 1 - e^{\{-\varphi(x)\}} \qquad \text{2-6}$$

And then if $P = P(X \leq x) = F(z)$

$$(1 - P)^n = e^{-n\varphi(x)} \qquad \text{2-7}$$

Weibull reasoned that if the probability of a single link of a chain failing took the form of equation 2-6 then the probability of a chain of n links not failing was given was given by equation 2-7.

The function $\varphi(x)$ needed to be a positive nondecreasing function vanishing at a certain value x_u which was equal to 0 or some other value. Weibull stated that the most simple function satisfying that condition was

$\varphi(x) = \frac{(x - x_u)^m}{x_0}$

If x_u is set to zero, this then gives

$$F(x) = 1 - e^{-\frac{x^m}{x_0}} \qquad \text{2-8}$$

And as x_0 is a constant, this can be written in the form

$$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \qquad \text{2-9}$$

And then equation 2-5, the probability density function, is given as the first derivative of the cumulative density function in equation 2-9. Only the parameter k changes the shape of the function.

The template function used for this study is shown in Figure 2-10 with its first derivative shown in Figure 2-11. Using the method of Wright (Wright, Nilsson et al. 2008), template fit was evaluated by iteratively cross correlating between the template and the signal, in this case over a 20ms loop so that the first run of the iteration matched the template with the signal starting at 40ms and running to 130ms, with the next iteration running from 41 to 131ms and so on over the loop of 20. In that way a vector of correlation coefficients was created. The highest correlation coefficient was taken as the best match for each sweep and the position of the whole signal was shifted by the time corresponding to the best match and was weighted by an amount equal to the correlation coefficient. The weighted sum was then calculated and the average adjusted by the sum of all the weights. Figure 2-12 shows a sample sweep with an overlay of the response templates. For clarity only every 3rd time shift is shown. The heavy blue trace showed the best correlation corresponding to a -1ms timescale adjustment in this case.

EMD, EEMD and CEEMDAN analysis were performed with the method of (Luukko, Helske et al. 2016) via the Rlibeemd package (Helske 2017). This differs slightly from the original papers in the detection of the extrema. Denoising consisted of removing the final residual from each trace and then finding the ensemble average. The maximum number of IMFs was used to minimize the signal in the final residual. Configurable parameters include:

- The S Number: The number of iterations of no change in the crossings – for this analysis set to 4.
- The maximum number of siftings as a stopping criterion: IMF extraction will stop at this number if the S Number constraint has not been reached. In this analysis set to 50.
- The ensemble size (for the EEMD and CEEMDAN): The number of noise augmented signals – set to 250 for this analysis.
- The noise strength (EEMD, CEEMDAN): The standard deviation of the Gaussian white noise (GWN) relative to the standard deviation of the input signal – for this analysis set to 0.2.

VMD analysis was performed with the vmd package in R (Ferry 2017). The baseline wander was extracted and removed as the first mode. Ten modes were extracted with a time-step of 0 and the balancing parameter left at the default value of 2000. All other parameters were left at their default values.

EWT analysis was done in Python using the ewtpy package (Carvalho, Moraes et al. 2020). This is a Python implementation of the original Matlab code (Gilles 2013). The decomposition was set to extract 10 wavelets with boundary detection by “LocMaxMin”. LocMaxMin refers to finding the wavelet boundary in the frequency domain as the minima between the maxima. No logarithmic transformation was specified and the regularization option was set to ‘none’. The baseline wander was extracted and removed as the first element in the decomposition – corresponding to the input signal convolved by the scaling function.

Statistical analysis

Comparisons between the recording sessions – eg comparing session 1 and session 2 post latency dependent averaging were made by quantifying the coefficient of repeatability (CoR) – otherwise known as the smallest real distance. This has been shown (Beckerman, Roebroeck et al. 2001) to be superior to the intraclass correlation coefficient (Shrout and Fleiss 1979). The measure is based upon the Limits of Agreement from the Bland-Altman plot (Bland and Altman 1986) and measures the value below which the absolute differences between two measurements falls with a 95% probability (Vaz, Falkmer et al. 2013). Assuming these differences are normally distributed, the CoR is calculated by multiplying the within subject standard deviation by 1.96 √2 where 1.96 is the value corresponding to the Quantile function of the normal distribution for 0.975. The √2 term arises because of the consideration of standard deviation rather than variance and the two tailed nature of the comparison. The CoR is measured in the same units as the measurements – in the case of the PhNR it is measured in uV. It can be normalized to the effective dynamic range of the measurement if a comparison between the CoR of different parameters is needed; for example to compare the CoR of the PhNR and the CoR of the b-wave. To allow for multiple comparisons with a dependent relationship, the method of Benjamini (Benjamini and Yekutieli 2001) was employed via the mutoss package in R (Werft 2017)

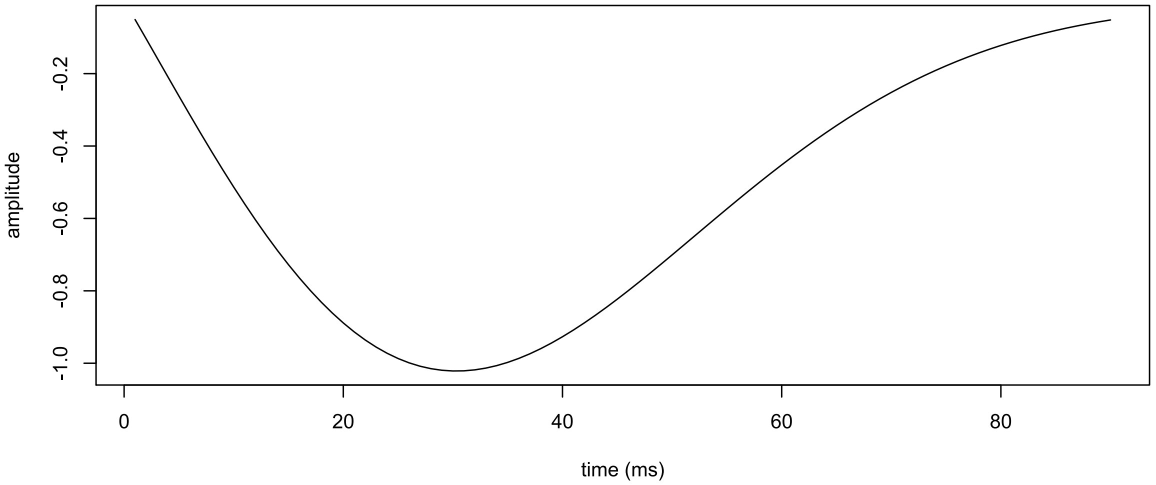


Figure 2-10: PhNR response template.

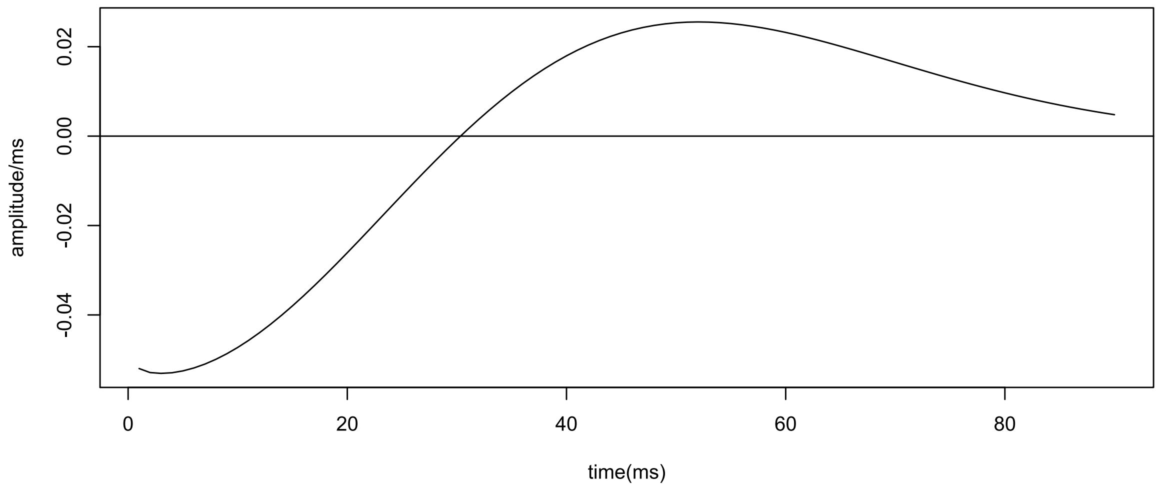


Figure 2-11: PhNR response template first derivative.

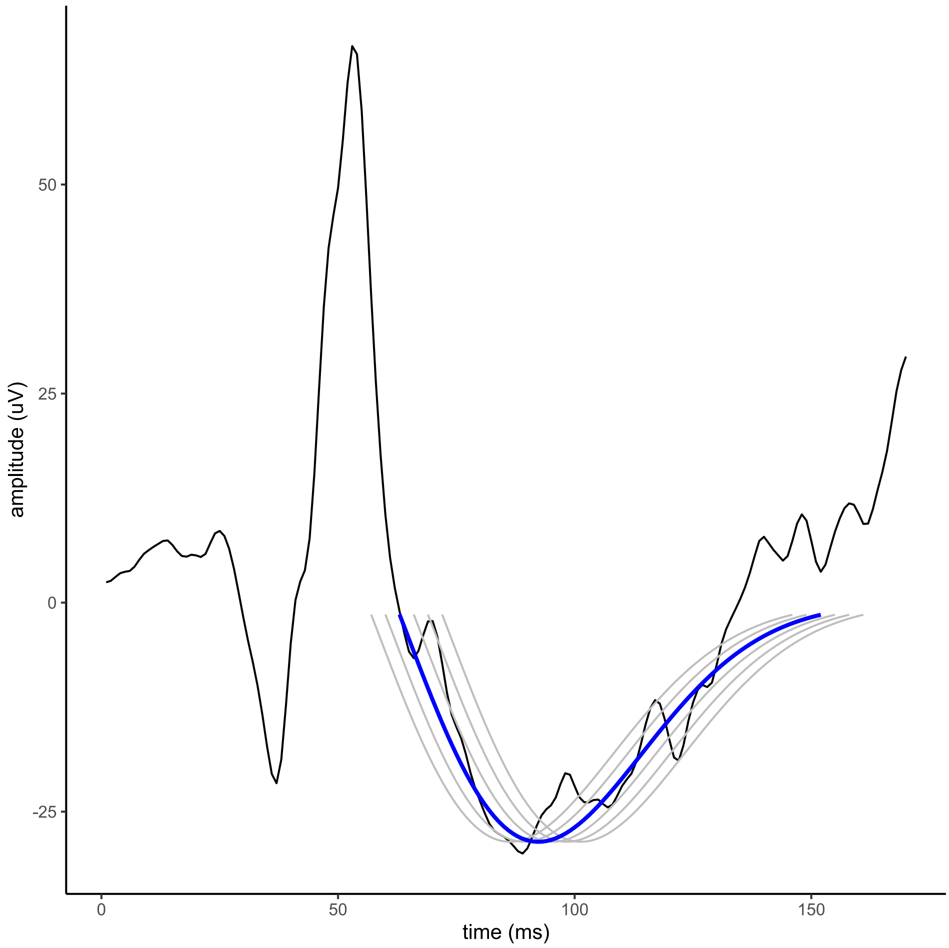


Figure 2-12: ERG sweep with PhNR templates. The best fit is shown in solid blue.

2.11 Results

Figure 2-13 shows the denoising process in action for a selected eye. The top row is the first testing session and the second row is the second session for the same eye. The thick solid trace representing the average ERG and the coloured traces the individual sweeps. The bottom row shows the test and retest averages superimposed. In this example, the CEEMDAN test and retest PhNR measurements are the closest. Figure 2-14 shows 3 further examples of the denoising in action.

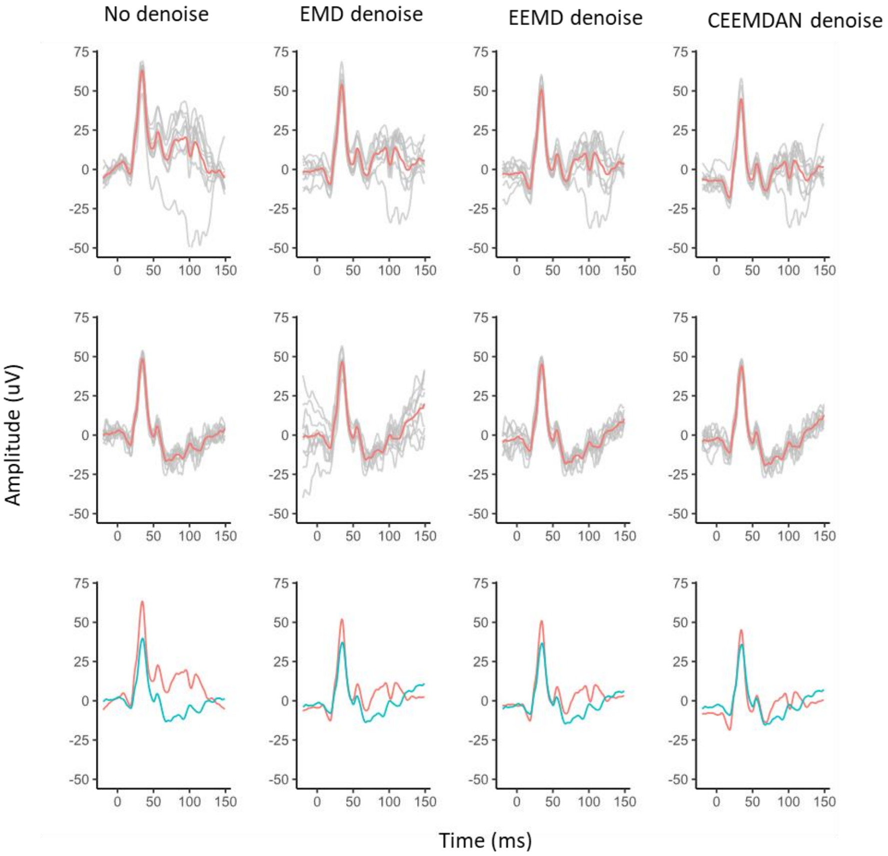


Figure 2-13: EMD denoising examples. Grey traces are individual sweeps, red solid trace is ensemble average. In the lowest row, the red and blue are the test and retest respectively.

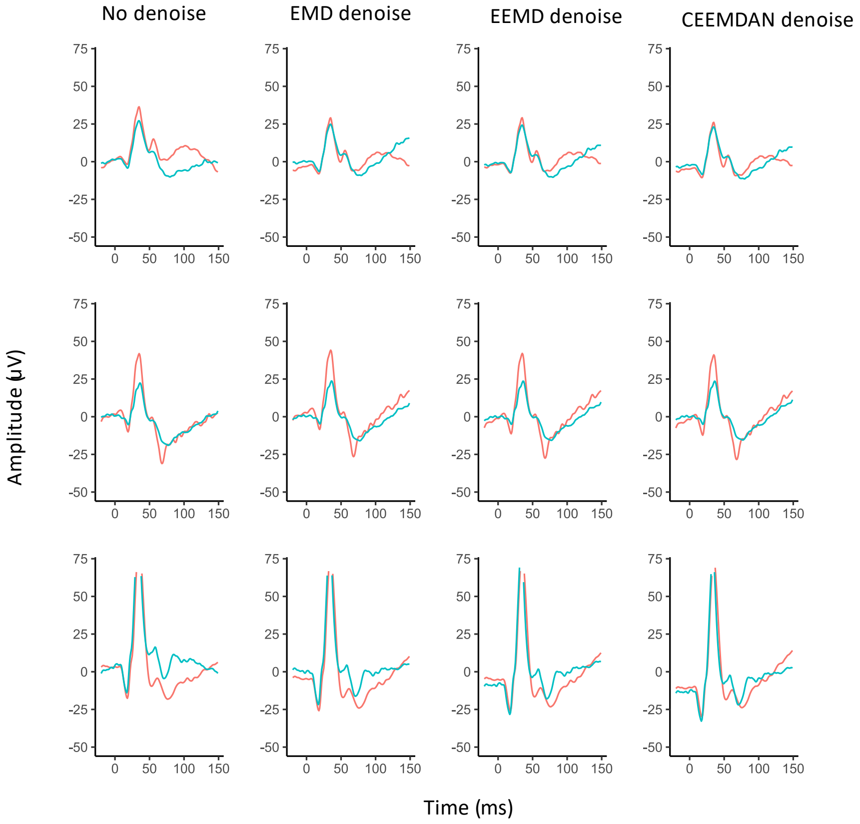


Figure 2-14: Further denoising examples. Only test and retest ensemble averages are shown.

Table 2-1: Marker calculations for all 112 eyes of 56 patients.

	a-wave (µV)	b-wave (µV)	PhNR (µV)	PhNR/B
Cohort Mean (SD)				
Raw	-10.6 (6.5)	62.4 (24.5)	-14.9 (13.9)	0.94 (0.16)
Latency Adjusted	-7.72$^{\#}$ (7.3)	54.7$^{\#}$ (30.9)	-16.3 (17.7)	1.10$^{\#}$ (0.42)
EMD	-14.5$^{\#}$ (7.9)	62.6 (24.7)	-16.7$^{\#}$ (8.8)	0.97$^{\#}$ (0.14)
EEMD	-14.9$^{\#}$ (7.4)	62.6 (24.7)	-17.3$^{\#}$ (8.6)	0.97$^{\#}$ (0.13)
CEEMDAN	-16.2$^{\#}$ (7.8)	62.5 (24.6)	-18.2$^{\#}$ (7.7)	0.97$^{\#}$ (0.13)
VMD	-16.1$^{\#}$ (7.5)	60.0 (24.8)	-14.5 (7.6)	0.99$^{\#}$ (0.18)
EWT	-18.5$^{\#}$ (7.4)	59.6 (24.9)	-12.8$^{\#}$ (5.2)	0.92 (0.09)

Notes: EMD = empirical mode decomposition; EEMD = ensemble EMD (EEMD); CEEMDAN = complete ensemble EMD with adaptive noise; VMD= variational mode decomposition; EWT= empirical wavelet transform; PhNR = photopic negative response, relative to the prestimulus baseline; PhNR/B = PhNR to b-wave amplitude divided by the a- to b-wave amplitude; $^{\#}$ = significantly different at $P < 0.05$ compared to the traces without denoising; † = significantly different at $P < 0.05$ compared to the EMD and EEMD methods. P values corrected for multiple comparisons by Benjamini-Hochberg adaptive linear method (Benjamini and Yekutieli 2001).

Table 2-2: Coefficients of repeatability for test-retest by marker and method.

	PhNR absolute (µV)	**PhNR/B absolute**	**PhNR normalised (%)**	**PhNR/B normalized (%)**
Coefficient of Repeatability				
Raw	3.0	0.04	20.2%	19.5%
Latency Adjusted	3.3	0.16$^{\#}$	20.2%	78.1%$^{\#}$
EMD	1.9$^{\#}$	0.04	11.3%$^{\#}$	13.6%$^{\#}$
EEMD	1.8$^{\#}$	0.03	10.3%$^{\#}$	13.1%$^{\#}$
CEEMDAN	1.5$^{\#}$	0.03	8.3%$^{\#}$	11.3%$^{\#}$
VMD	1.5$^{\#}$	0.03	10.3%$^{\#}$	11.3%$^{\#}$
EWT	0.94$^{\#}$	0.02	7.3%$^{\#}$	7.5%$^{\#}$

Notes: EMD = empirical mode decomposition; EEMD = ensemble EMD ; CEEMDAN = complete ensemble EMD with adaptive noise; VMD= variational mode decomposition; EWT= empirical wavelet transform; PhNR = photopic negative response, relative to the prestimulus baseline; PhNR/B = PhNR to b-wave amplitude divided by the a- to b-wave amplitude; * = relative to effective dynamic range; $^{\#}$ = significantly different at $P < 0.05$ compared to the traces without denoising. P values corrected for multiple comparisons by Benjamini-Hochberg adaptive linear method (Benjamini and Yekutieli 2001)

2.12 Discussion and Summary

Noise is a problem in clinical electrophysiology. High frequency noise is amenable to conventional filtering but low frequency non stationary noise is not. The majority of the power spectrum of the electroretinogram falls within the range of 1 to 200Hz (Gur and Zeevi 1980) and this overlaps some typical sources of noise such as 50Hz power line noise and 100Hz fluorescent light sources. Although notch filters can be used to reduce such noise, this can affect the waveform or timing of the response. , Radio frequency interference from switchmode powersupplies can be adequately removed by analog or digital filters.

Very low frequency noise such as DC trend and blink artefact are much more problematic. The influence of this type of noise is greater on a late signal like the PhNR than say the a-wave, as the PhNR is typically referenced to the baseline. Referencing the PhNR to the b-wave peak can partially offset this but adds extra unwanted variance into the signal which is now a combination of two potentially unrelated physiological processes.

The fact that the PhNR is within the latency for the photomyoclonic response and blink artefact is also problematic. Contraction of the orbicularis muscle has a complex effect on the recorded signal. First, the EMG signal will have a very different waveform and amplitude on the indifferent skin electrode to the active silver wire electrode. This will result in the failure of common mode rejection. Further, small movement of the metallic wire may cause induced voltage on the silver electrode through changes in the local tear chemistry and perhaps inductive pickup from local magnetic fields. It is highly likely that the time constant of the two electrodes (skin and DTL) are quite different – again making common mode rejection impossible.

The other key issue which affects the PhNR response is in the nature of the response itself. The response is small and slow with no well demarcated or stereotypical waveform. Although thought to be generated by the spiking of ganglion cells, it behaves quite unlike a compound action potential. The compound nerve action potential for example (Lang and Puusa 1981) elicited by a 0.2ms stimulus produces a brisk response when measured at the skin of about 2ms duration. Even the a-wave is not that brisk and propagation delays and multiple parallel and opposing pathways (inhibitory and excitatory) leads to a blurring of the response.

Nor is the response or the noise stationary. Habituation to the stimulus changes the EMG and movement artefact over the duration of the stimulus set and drift in the latency of the peak and its shape limit the extent to which simple signal averaging can be effective.

Detrending algorithms exist and are often included in the electrophysiological hardware. In this work, a simple algorithm to detrend and zero the trace was used as a comparator to more innovative techniques. A problem with this and similar algorithms is that they must ignore the actual response itself and further must assume some structure to the noise – for example linear, logarithmic, sinusoidal or polynomial. The elegance of the EMD family is that the algorithm uses the whole response, including the actual ERG and does not assume a structure to the noise.

The latency adjustment technique of Woody (Woody 1967) has a good theoretical underpinning. It does however depend on being able to match the template response in each sweep. This would work well for a-wave such as the b-wave which is a large amplitude consistent wave. It would be of little or no utility in a waveform such as the pattern ERG (PERG) where in each sweep, the noise is much larger amplitude than the signal. As such, the technique is most useful when it is not needed. In my work, this was what was found for the PhNR where the technique provided no benefit.

The Kalman filter has found great utility in navigation systems and has been used as a denoising technique. In this work, the filter worked to smooth the signal – that is it was effective in removing higher frequency noise. It had little to no effect in removing the large amplitude ultra low frequency noise that contaminates the ERG in the later parts of the response. Although there is some utility in smoothing the higher frequency components, they are easily removed with conventional analog or digital filters and this technique is probably of little use in this particular signal. Signals which are contaminated by high frequency non stationary noise such as the raw data for the multifocal ERG may benefit from this type of pre-processing.

This study showed good utility from the empirical mode decomposition as a type of adaptive filter. The filter will adapt to the noise for each trace and will remove any monotonic noise. The nature of the algorithm will also autozero the signal which is very useful for the PhNR response as it is referenced to the baseline.

In order for the EMD filter to remove the monotonic baseline noise, the number of IMFs extracted needs to be sufficiently large. This is likely to be the only important tuning parameter for the decomposition. In the example given in Figure 2-2, the decomposition can clearly be seen to be longer than is necessary, with IMF 4-6 essentially identical. Fortunately, an overly large decomposition does not affect the denoising as it is the residue that is removed as the noise. A decomposition that is too short will however not leave a monotonic residue and therefore not function properly. The implementation of the algorithm does not automatically choose the number of IMFs to be extracted and therefore it is recommended to use more than the default of $\log_2$ of the signal length. The ensemble techniques offer additional tuning parameters such as ensemble size and noise level but for the specific detrending problem the defaults can be safely used.

In this work, I have shown with real and simulated data that the ensemble EMDs offer a small advantage over the EMD. Software libraries such as that of Lukko (Luukko, Helske et al. 2016) offer implementation of the EMD and the two ensemble techniques with very fast execution times. The library can be called from Python and R which makes deployment of the technique with open source easy. Similarly, the library can be called from Matlab. For this chapter, I have rewritten the algorithm in R to explore and demonstrate its inner workings. Although slower, this technique could be implemented natively in Excel or other platforms in a similar manner to avoid the complexity of calling the C code directly from Excel if this was the preferred application.

The translational significance of this work is in the demonstration of the technique as an effective and reliable one for detrending and ultra-low frequency denoising for visual electrophysiology. Removing of this major artefact improves the reliability and ultimately the utility of measurements of the PhNR for glaucoma. Further, the technique should be of utility in other similar settings such as the Pattern ERG (PERG), the Pattern VEP (PVEP) and the On-Off ERG. I envisage that ultimately the detrend by EMD will be a selectable button on electrophysiology equipment that the operator can use on a case-by-case basis depending on the protocol and the amount of noise present at the time. That stated, extreme care would be needed to compare serial traces where one was filtered and one not. For this reason, the technique is much better implemented *post-hoc* where a decision is made to apply it to all or none of the traces. In the current implementation of the Espion software, the detrending (called baseline correction) is a protocol level option and is not reversible.

Future work will involve further optimisation of the tuneable parameters and testing of the algorithm in both glaucoma and non-glaucoma cohorts.

Chapter 3

Complexity Measures of the Electroretinogram

The findings of this chapter have been published in part (see Appendix)

Sarossy M, Aliahmad B, Kumar D, editors. Timing of changes in the entropy of the electroretinogram with Glaucoma. Life Sciences Conference (LSC), 2017 IEEE; 2017: IEEE..

Sarossy M, Kumar D, Wu Z, editors. Relationship between Glaucoma and Complexity Measures of the Electroretinogram. 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII); 2021: IEEE

3.1 Abstract

Background and objective: Biological signals are inherently complex. The photopic negative response (PhNR) of the electroretinogram (ERG) is known to originate in the retinal ganglion cells (RGCs) of the eye. Glaucoma is a disease associated with increased intraocular pressure in which the RGCs suffer dysfunction and death. This chapter explores whether measures of signal complexity can classify eyes into those with and those without glaucoma.

Methods: Electroretinograms were measured from the left eyes of 21 participants with glaucoma and 18 participants without eye disease. Skin electrodes were used with the RetEval hand held ERG device using a red flash on a blue background. Entropy and fractal measures of complexity were calculated on both individual sweeps and the ensemble average and a sliding window technique was used to assess changes of Shannon and sample entropy over the epoch. The PhNR on the ensemble averages was also determined.

Results: The PhNR was of lower amplitude in the glaucoma group (-3.6 vs -5.50 uV, p=0.012). Complexity measures calculated on individual traces and averaged showed no significant differences between the groups. Higuchi fractal dimension was higher (p=0.026) in the glaucoma group and Shannon entropy was lower (p=0.014). After the stimulus, the sample entropy of the signal returned to baseline quicker in the glaucoma group (74.7 vs 79.8 ms, p =0.04).

Conclusions: There are subtle differences in the complexity measures derived from the ERG between glaucoma and non glaucoma eyes. This may be due to loss of the number of ganglion cells firing action potentials, or to the structure of that discharge. These changes are less significant than the changes in the amplitude of the PhNR.

3.2 Introduction

Physiological systems often show great complexity. This arises from the number of contributing subunits, the multiple feedback and feed forward loops and the structures operating over multiple scales of time and space (Pukėnas, Poderys et al. 2012). A time series, like the electroretinogram (ERG) is complex and the complexity of the signal can be measured by a number of techniques. These measures could have utility for a glaucoma clinician: to inform either a classification decision – such as the presence or absence of glaucoma - or a severity assessment. Complexity measures can be broadly thought of as measures of predictability or regularity. Predictability can be considered in the temporal domain with techniques such as the fractal dimension (FD) or detrended fluctuation analysis (DFA) (Alvarez-Ramirez, Rodriguez et al. 2009). Predictability in the spatial domain can be measured by Kolmogorov-Sinai Entropy (KSE) (Lim, Khang et al. 2013) or the largest Lyapunov Exponent (Swiderski, Osowski et al. 2005). The alternative formulation of complexity is the assessment of regularity – which can be estimated with measures of entropy or the Lempel-Ziv complexity measure (Abásolo, Hornero et al. 2006). In this chapter, the use of information theory measures and fractal dimension are explored to estimate the complexity of the underlying signals and to determine if differences exist between a glaucoma and non glaucoma cohort of experimental participants.

3.3 Information Theory Analysis

The amount of information that can be carried over a noisy channel is critical in planning what can be sent over the channel and to avoid congestion. Less noisy channels can carry more information. Claude Shannon working at Bell laboratories in his seminal paper (Shannon 1948) mathematically detailed the information carrying capacity of a channel in communications. He showed that information could be sent reliably over channels that were themselves unreliable. He found, remarkably, that the form of the equation related to thermodynamic entropy.

The first law of thermodynamics states that the total energy of an isolated system is constant. The second law relates to entropy and early characterizations of the principle were by Kelvin (Kelvin 1890) and Clausius (Clausius 1856). Helmholtz, inventor of the ophthalmoscope, also made contributions to the early understanding of entropy (Helmholtz 1884). The well-known Boltzmann equation defined the entropy of an ideal gas to be related to the number of possible microstates as:

$$S = k_b ln\, W \tag{3-1}$$

Where S is the entropy, W is the number of microstates and k_b is the Boltzmann constant. In effect, the Boltzmann constant is the proportionality between the statistical entropy related to the logarithm of microstates and thermodynamic entropy expressed in Joules per degree Kelvin.

This was a very important development in thermodynamics. Boltzmann's tombstone in Vienna is inscribed with his formula. His formulation assumes an equal probability of all of the microstates and was generalized by Gibbs (Gibbs 1948) as

$$S = -k_b \sum p_i\, ln\, p_i \tag{3-2}$$

Where p_i is the probability of state I. Equation 3-2 reduces to 3-1 if all states are equal.

The optic nerve is a massively parallel structure containing around one million axons (Mikelberg, Drance et al. 1989). Each of these axons carries a quantity of information up the visual pathway and ultimately to the visual cortex. Each ganglion cell has a maintained stochastic discharge and that discharge is modulated up or down depending on whether the ganglion cell is excited or stimulated (Schottdorf and Lee 2021). Each ganglion cell can therefore be considered a noisy channel. The total information that can be carried up the nerve is related to the number of channels and their independence from each other. If the channels were perfectly independent, the total information would be the sum of the information carried in each channel or axon. The parts of the ERG that constitute the PhNR are known to be generated by the action potentials of ganglion cells (Viswanathan, Frishman et al. 1999). A change in the number of channels should lead to a decreased level of total information capacity unless offset by the dependency or the information per channel.

3.3.1 Shannon Entropy

Claude Shannon (Shannon 1948) considered the information carrying capacity of a channel. He considered the total information as a measure H in a signal Y with a finite number of possible states $\{S_1, S_2, .. S_n\}$. If the system behaves as a Markov system with ergodic properties – that is, there is a nonzero chance of leaving any state and a probability of 1 of eventually returning to every state - then consider the set of probabilities $\{p_1,p_2,..p_n\}$ corresponding to the probabilities of any of the states S. A measure H, total information, is defined:

$$H = -\sum p_i\, log\, p_i \tag{3-3}$$

One can immediately see the striking resemblance of equations 3-2 and 3-3, differing only by the Bolzmann constant. The units can be bits if the log is base 2 or 'nats' if a natural logarithm is used.

The Shannon entropy has the following properties

1. H is continuous in the p.
2. If all of the p are equal $p_i = \frac{1}{n}$, then H increases monotonically with n.
3. H = 0 if and only if all of the p_i are 0, the exception having the value unity. The system carries no information. Otherwise H is positive.

Intuitively, this result shows that the information carrying capacity of a channel is a function of the number of possible values that can be sent down the channel and of the probabilities of each of those values. If there is only one possible value, as in case 1 above, there is no information sent down the channel. If there is a very large number of possibilities there is a large amount of information carrying capacity, which is in turn dependent on the probabilities of those values. In communications theory, the units for this measure are bits (compared with Joules degrees^{-1} in thermodynamics). This is achieved by using 2 as the base for the logarithm in equation 3-3.

The concept can be illustrated with an example. Consider the English alphabet as the dictionary of information to be sent down a channel. If all of the letters were of equal probability of occurance, the entropy would be

$$-log_2 \frac{1}{26}$$

ie 4.7 bits.

Consider the probabilities of each letter occurring, then using the frequency of letters from a sample of large English Language corpora as sampled by Jones (Jones and Mewhort 2004). If the channel consists of just the letters A-Z, the entropy of the channel disregarding case is 4.18 bits (summing over the 26 probabilities). If both upper and lower case are allowed the entropy is 4.47 bits (summing over the 52 probabilities): Ah igher amount of information capacity is present in the channel. It is important to note that with Shannon entropy, the calculation is performed on the probability mass function of the channel and not on the actual channel data itself.

Shannon went on to discuss the case of the discrete channel contaminated by noise. Biological systems are inherently noisy – for example, the ganglion cell spiking rate is not regular but a Poisson process with parameter lambda controlled by the amount and pattern of light (Levick, Thibos et al. 1983). In this case, the entropy of the output channel is given by

$$H(y) = H(x, y) - H_y(x) \qquad 3\text{-}4$$

Where H(y) is the entropy of the output channel, H(x,y) is the joint entropy and $H_y(x)$ is the conditional entropy of the input. These are derived from the corresponding joint and conditional probabilities:

$$H(x,y) = -\sum p(x,y)\, log\, p(x,y) \tag{3-5}$$

And

$$H_y(x) = \sum p(x,y)\, log \frac{p(x,y)}{p(y)} \tag{3-6}$$

Shannon generalized his theory to continuous data to what is known as Differential Entropy. Equation 3-3 can easily be generalized for a continuous probability distribution to

$$H = -\int_{-\infty}^{\infty} p(x)\, log\, p(x)\, dx \tag{3-7}$$

For example, the entropy of a one-dimensional Gaussian distribution with standard deviation σ is given by

$$H(x) = log\sqrt{2\pi e}\, \sigma \tag{3-8}$$

Which is clearly dependant only on the variance and not the mean of the distribution. This scale invariance is an important property of entropy. The entropies for a range of distributions have been derived. It can be shown that for a given variance, a Gaussian random variable will have the highest entropy (Michalowicz, Nichols et al. 2013). The entropy of a continuous distribution calculated this way has quite different characteristics to that of a discrete probability distribution – not the least of which is that the entropy can take on negative values whereas for a discrete distribution, only positive values are possible. Care must be taken too with differential entropy to make the measure dimensionless before integration as the entropy calculated can vary with the units used.

To illustrate this further, consider 3 signals: a Gaussian white noise (GWN) variable, a random variable drawn from the uniform distribution and a sine wave. All signals are set to have a mean of 0 and a variance of 1.

The GWN variable will have have entropy as above of $log_e\sqrt{2\pi e}$ - approximately 1.419 in nats. The uniform distribution variable is drawn from the range of $-\sqrt{3}$ to $\sqrt{3}$ to give it uniform variance and zero mean and then the entropy is given by $log_e(a-b) = log_e(2\sqrt{3})$ – approximately 1.292. The sinusoid has entropy given by $log_e\frac{\pi A}{2} = log_e\frac{\pi}{\sqrt{2}}$ (Michalowicz, Olson et al. 2009) approximately 0.798. We shall see later in this chapter that signals contaminated with sinusoidal noise have a counter-intuitively low entropy.

For biological data, such as the electrocardiogram (ECG) or indeed the electroretinogram (ERG), the data available for analysis are not truly continuous but rather sampled continuous data. For this reason, entropy calculations are usually discrete. For sampled data, the entropy will be dependent upon the number of bins used in the sum. It is important to take this into account when comparing entropy results.

In order to use the concept of entropy to analyse a signal, the entropy must be estimated. The simplest method to do this is to estimate the probability mass function directly from the data by constructing a histogram and applying maximum likelihood techniques. This was described by Strong (Strong, Koberle et al. 1998) and is known as the plug-in estimator (Antos and Kontoyiannis 2001).

$$\widehat{H}_{mle}(p_N) = -\sum_{i=1}^{m} p_{N,i} \, log \, p_{N,i} \qquad 3\text{-}9$$

In that paper, the authors showed that the estimator converged to the true entropy at a rate proportional to the square root of the variance and inversely proportional to the square root of the sample size. This held over a number of different probability distributions.

Under some conditions, the plug-in estimator has been found to be significantly biased. Miller (Miller 1955) proposed a correction to the estimator :

$$\widehat{H}_{mm}(p_n) = \widehat{H}_{mle}(p_n) + \frac{\widehat{m} - 1}{2N} \qquad 3\text{-}10$$

Where m is the estimate of the number of bins with nonzero probability and N is the size of the alphabet.

Other techniques have been described to minimize the bias in estimating the entropy including Nemenman (Nemenman, Shafee et al. 2002). Even with bias correction, the Shannon entropy estimate is related to the number of bins (analogous to the size of the alphabet) and will monotonically rise with an increasing number of bins.

The Shannon entropy measure has used by many authors, for example Cugini (Cugini, Bernardini et al. 2001) to examine the structure of electrocardiogram signals and to detect change points within the data. It has also found utility in the analysis of electroencephalogram signals, for example by Bruhn (Bruhn, Lehmann et al. 2001) in which Shannon entropy measurements were made of electroencephalogram signals of patients under anaesthesia with desflurane for female patients undergoing gynaecologic laparotomies. A correlation was found in that study between the end tidal desflurane concentration and the entropy measure.

The information carrying capacity of retinal ganglion cells in the isolated guinea pig retina was estimated by Koch and co-authors (Koch, McLean et al. 2004). They measured spike trains via extracellular recordings, detecting spikes as a first derivative of the voltage exceeding a threshold – that is a gradient of the voltage recording being sufficiently steep. Using these spike trains, they estimated total and noise entropy by forming letters and words from the patterns of discharge (van Steveninck, Lewen et al. 1997) and used these data to directly estimate entropy via Shannon's formula. The authors estimated a coding efficiency for each axon at only around 33%. The total information carrying capacity of the optic nerve is a function of the information capacity of each axon and the correlation of the discharge patterns between axons. High degrees of decorrelation will lead to a higher overall information carrying capacity. Using a non-linear model in both salamander and macaque retinas, Pitkow (Pitkow and Meister 2012) found that the nonlinearities of the ganglion cell firing pattern nearly optimally decorrelated the spike trains between ganglion cells and this in turn optimized the overall information carrying capacity of the optic nerve.

No research thus far has been published in applying this technique to the electroretinogram or glaucoma. The plausibility of the approach however can be demonstrated as follows:

Ganglion cells of most species have a maintained discharge – that is that they fire even in complete darkness (Barlow and Levick 1969). On centre ganglion cells increase their discharge rates with stimulus onset and decrease the firing rate when the stimulus ends. Off centre ganglion cells will decrease their firing rate when the stimulus goes on and reduce the rate of discharge at the end of the stimulus. For a stimulus such as the flash electroretinogram, these responses will overlap and have complex effects on the electroretinogram after the b wave peak. The pattern of discharge is stochastic – in continuous time, the inter-spike interval at a given level of stimulus is best represented by a gamma distribution and the spikes per unit time can be represented by a Poisson distribution (Troy and Lee 1994).

A massed signal can be simulated by the addition of a number of simulated spike trains generated in this way.

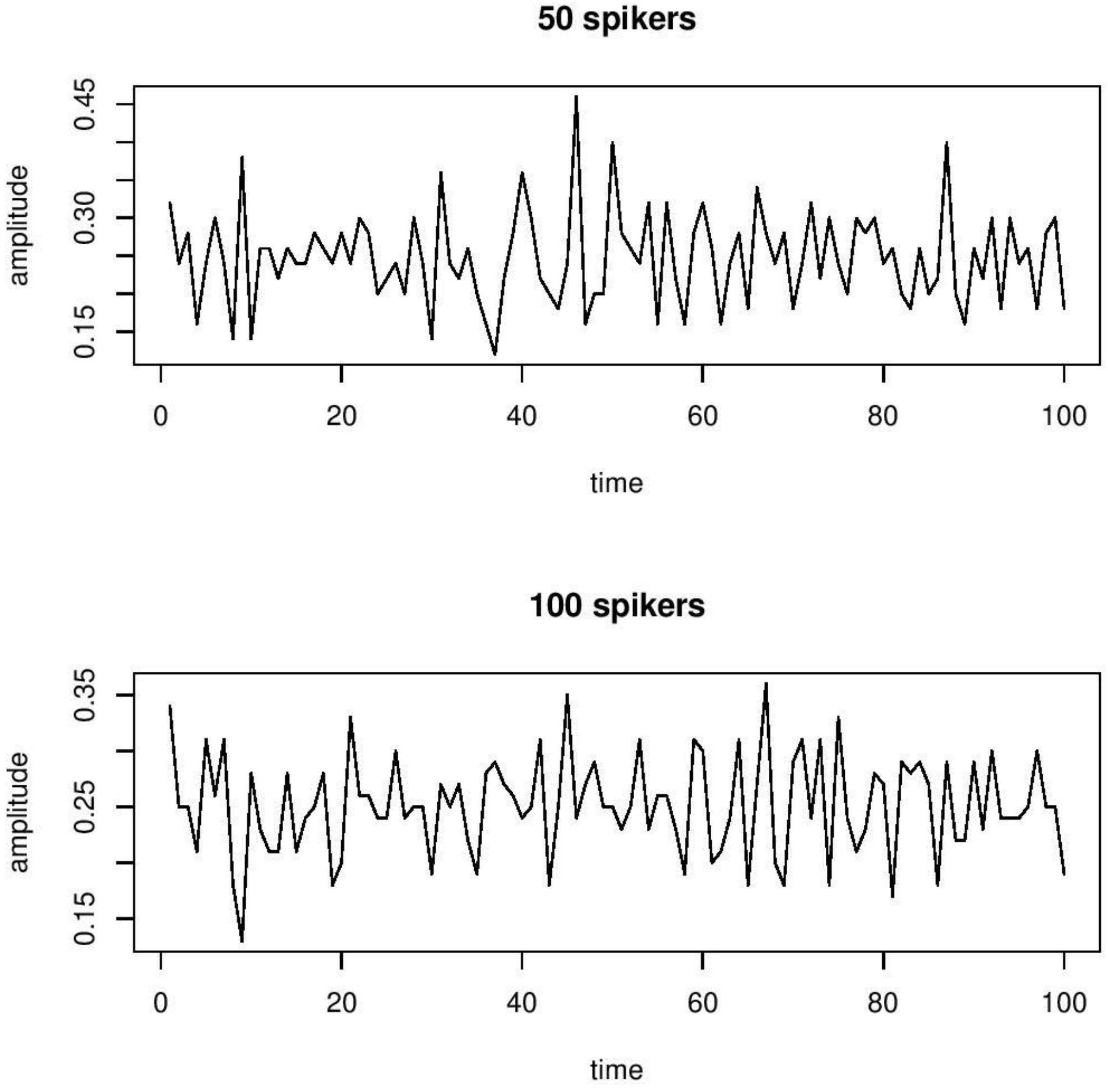


Figure 3-1: Simulated neural signal

Figure 3-1 shows two such signals – generated from 50 units and 100 units on the top and bottom respectively. This is of course a gross simplification as the units are not weighted for distance and the firing patterns are all the same. Nevertheless, if the Shannon entropy is calculated for these two synthetic signals, there is a significant difference found between them. This is illustrated in Figure 3-2 showing a higher entropy and therefore information capacity in the signal with more channels. It is hypothesized that the same difference may be apparent between glaucoma and non glaucoma test subjects.

The PhNR has been shown to have input from glial cells (Viswanathan and Frishman 1997, Frishman, Viswanathan et al. 2001) confirmed by the fact that Barium ions can be shown to eliminate the response in isolated retina by blocking Potassium channels in the glial cells (Bolz, Schuettauf et al. 2008). Amacrine cells can generate action potentials (Wu, Ivanova et al. 2011). In an experimental model of glaucoma in the rat (Kielczewski, Pease et al. 2005) numbers of amacrine cells were unaffected by glaucoma even though ganglion cell numbers declined. The amount of neurotransmitter within these cells did decline substantially in animals with glaucoma which would indicate a change in function and therefore could plausibly affect the electroretinogram and its measures of complexity.

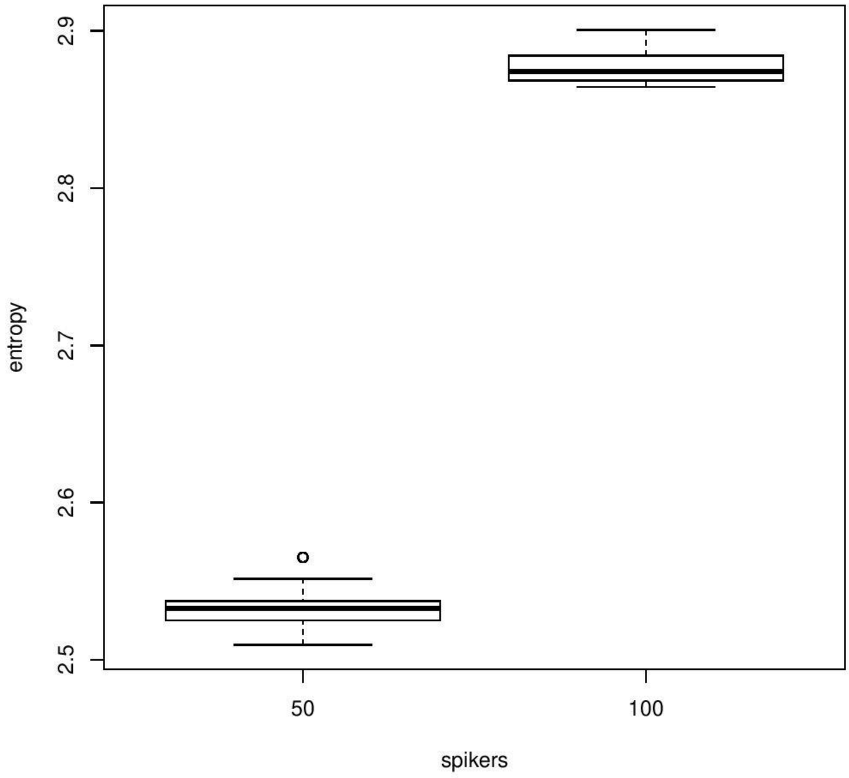


Figure 3-2: Shannon entropy of simulated signals

3.3.2 Rényi Entropy

By relaxing the constraint of additivity, Renyi (Rényi 1961) generalized the Shannon entropy into a family of entropies as

$$H = \frac{1}{1-\propto} log \sum p_f^{\propto} \qquad 3\text{-}11$$

Where α is greater than 0 and not equal to 1 but in the limit as α approaches 1 H approaches the Shannon entropy. Where α =2, the measure is known as the Renyi quadratic entropy (or collision entropy). The Renyi entropy has also been used in biosignal analysis for example Xin (Xin, Zhao et al. 2017) but not thus far for the electroretinogram.

3.3.3 Approximate Entropy

Recall that the Shannon and related entropy measures do not account for the sequence of data within the signal. Further, it is generally thought that a large dataset is required for accurate estimation of the entropy measure. Signals, such as biosignals, consist of an ordered sequence of observations corresponding to amplitude measures at a regular sequence of time points - known as a time series. Kolmogorov (Kolmogorov 1985) proposed a new method of considering entropy as the rate of which new information is acquired in a signal as the length of the subsignals increases. Grassberger (Grassberger and Procaccia 1983) showed how this could be simplified and applied to a time series and Pincus (Pincus, Gladstone et al. 1991) proposed a variation: Approximate Entropy (ApEn) as a new statistic using these concepts to quantify regularity in data. The first application suggested was for heart rate data. ApEn has been used on a variety of biosignals including ECG (Holzinger, Stocker et al. 2012), EEG (Srinivasan, Eswaran et al. 2007) and EMG (Chen, Wang et al. 2006). Nair (Nair and Joseph 2014) used the technique to analyse the electroretinogram in patients with a variety of retinal dystrophies compared with controls. There was a significant difference between the groups but was less than the authors found with the Hurst exponent. The technique has not been otherwise described for the electroretinogram and not for glaucoma. Richman (Richman and Moorman 2000) suggested that ApEn, by matching each subsequence within the time series with itself, introduces significant bias and makes it heavily dependent on the record length; being uniformly lower than expected for short records and lacking consistency. The authors developed a new related measure known as sample entropy described in the next section. As ApEn is generally considered to be flawed as a measure, it will not be considered further in this thesis.

3.3.4 Sample Entropy

Sample entropy – SampEn (Richman and Moorman 2000) is the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point excluding self matches. The authors with a simulated data set found the measure much less biased than ApEn and required less computational time to estimate. They also claim it agrees with theory more closely than ApEn. SampEn has found application in the analysis of ECG time series data (Bornas, Llabrés et al. 2006), EEG (Abásolo, Hornero et al. 2006) and EMG (Zhang and Zhou 2012) and has largely supplanted ApEn. Nair (Nair and Joseph 2014) used SampEn in combination with other features to differentiate retinal dystrophies from normal controls. It has not otherwise been used in electroretinography.

3.3.5 Permutation Entropy

A related approach to SampEn and ApEn is permutation entropy (Bandt and Pompe 2002). In this algorithm, the entropy is calculated on a Shannon or Renyi basis on the probability of permutations of the subseries. For example, in the series

$$x = (4,7,9,10,6,11,3)$$

For a permutation order of 3, the subseries are constructed:

$$\{(4,7,9),(7,9,10),(9,10,6),(10,6,11),(6,11,3)\}$$

Then each of the subseries are compared to the next one for magnitude order. So (4,7,9) yields 012, whereas (10,6,11) yields 102. The probabilities for all of the permutations in the series i.e. (012,021,102,120,201,210) are computed across the signal and the entropy calculated as per Shannon's (or Renyi's) formulation

$$H = -\sum p(\pi)\, log\, p(\pi) \tag{3-12}$$

Where the sum runs over all n! permutations π of order n.

As with many other measures of entropy, the technique has been used for EEG, ECG and EMG but there are no published works using it for electroretinography

3.3.6 Distribution Entropy

A recent modification of SampEn has been described, known as distribution entropy (DistEn) (Li, Liu et al. 2015). To illustrate this method, consider a small extract of an ERG signal plotted as Figure 3-3. The amplitude values have been multiplied by 1000 and rounded for clarity

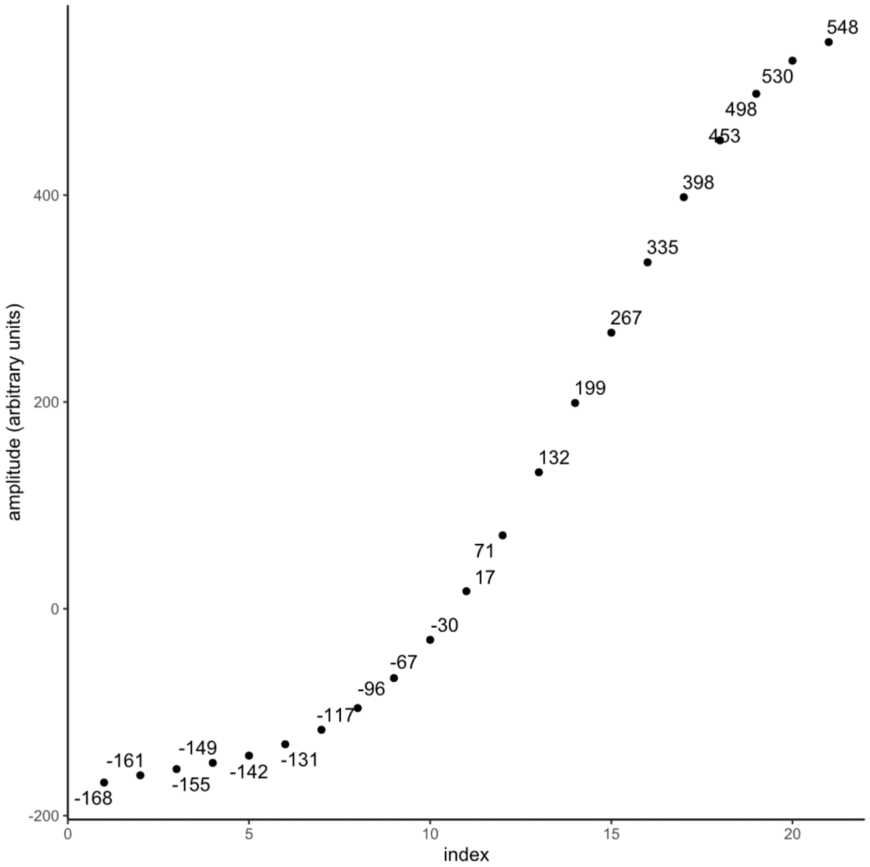


Figure 3-3: Part of the ERG trace converted to integers

A series of vectors of length equal to the embedding dimension are created Figure 3-4. In this case the embedding dimension is 2 but it can be of any integer length greater than 1. From this list of integers, a matrix is made up consisting of all of the Chebyshev distances (the greatest difference between any of the co-ordinates) between the vectors. The distances of the vectors with themselves (on the lower left to top right diagonal) are removed. This is shown in Figure 3-5.

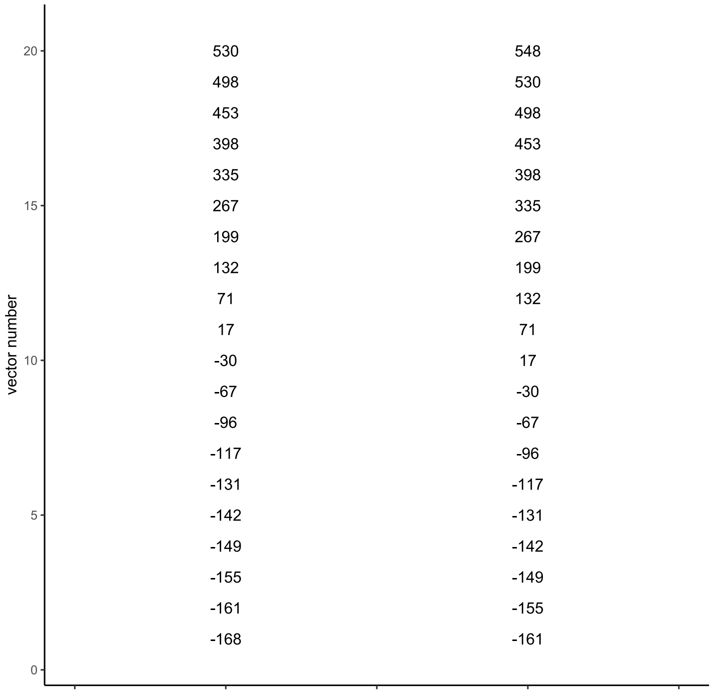
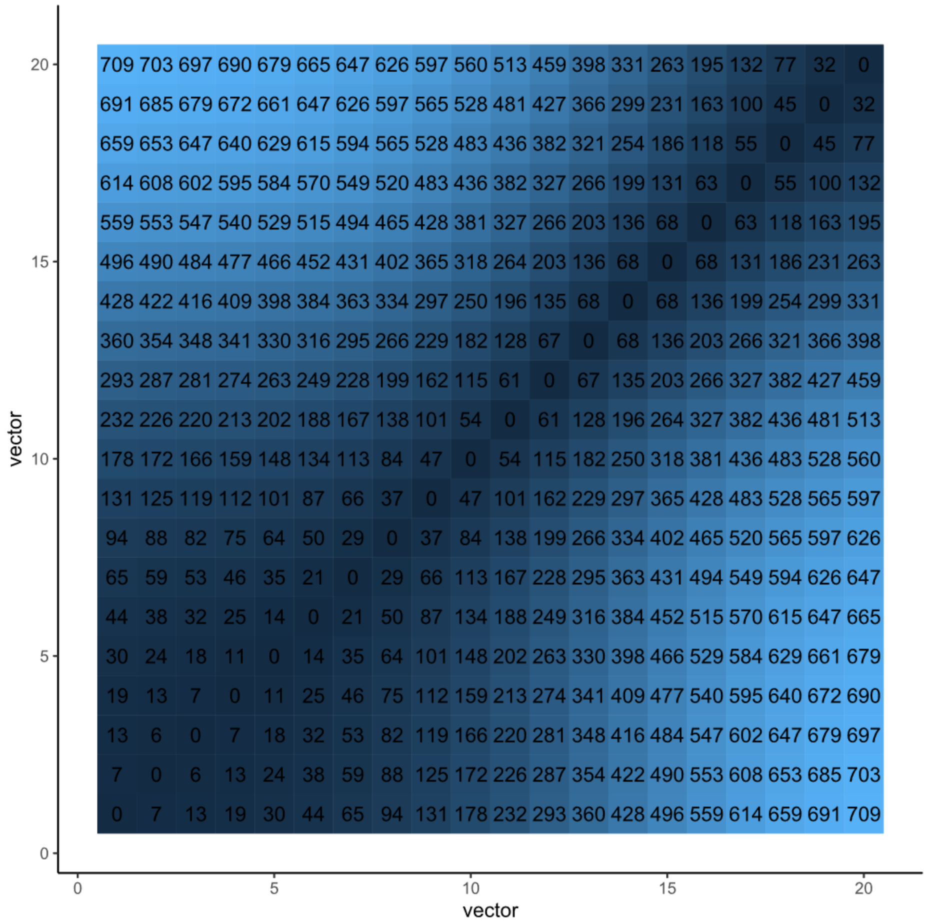


Figure 3-4: Vectors created for distribution entropy



709	703	697	690	679	665	647	626	597	560	513	459	398	331	263	195	132	77	32	0
691	685	679	672	661	647	626	597	565	528	481	427	366	299	231	163	100	45	0	32
659	653	647	640	629	615	594	565	528	483	436	382	321	254	186	118	55	0	45	77
614	608	602	595	584	570	549	520	483	436	382	327	266	199	131	63	0	55	100	132
559	553	547	540	529	515	494	465	428	381	327	266	203	136	68	0	63	118	163	195
496	490	484	477	466	452	431	402	365	318	264	203	136	68	0	68	131	186	231	263
428	422	416	409	398	384	363	334	297	250	196	135	68	0	68	136	199	254	299	331
360	354	348	341	330	316	295	266	229	182	128	67	0	68	136	203	266	321	366	398
293	287	281	274	263	249	228	199	162	115	61	0	67	135	203	266	327	382	427	459
232	226	220	213	202	188	167	138	101	54	0	61	128	196	264	327	382	436	481	513
178	172	166	159	148	134	113	84	47	0	54	115	182	250	318	381	436	483	528	560
131	125	119	112	101	87	66	37	0	47	101	162	229	297	365	428	483	528	565	597
94	88	82	75	64	50	29	0	37	84	138	199	266	334	402	465	520	565	597	626
65	59	53	46	35	21	0	29	66	113	167	228	295	363	431	494	549	594	626	647
44	38	32	25	14	0	21	50	87	134	188	249	316	384	452	515	570	615	647	665
30	24	18	11	0	14	35	64	101	148	202	263	330	398	466	529	584	629	661	679
19	13	7	0	11	25	46	75	112	159	213	274	341	409	477	540	595	640	672	690
13	6	0	7	18	32	53	82	119	166	220	281	348	416	484	547	602	647	679	697
7	0	6	13	24	38	59	88	125	172	226	287	354	422	490	553	608	653	685	703
0	7	13	19	30	44	65	94	131	178	232	293	360	428	496	559	614	659	691	709

Figure 3-5: Chebyshev distance matrix

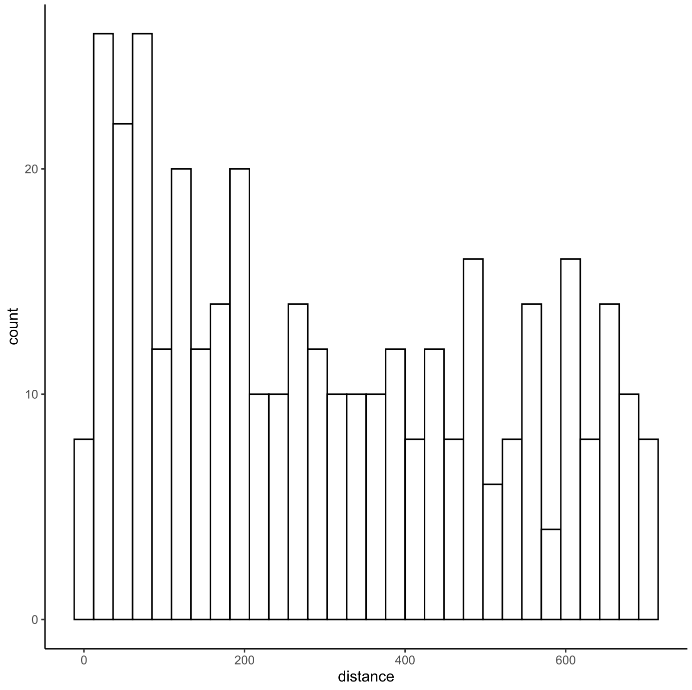


Figure 3-6: Histogram of distance matrix

Finally a histogram is constructed from the distance matrix as shown in Figure 3-6. The Shannon entropy is then calculated from this histogram. The authors of this technique claim that it is less sensitive to the number of bins than SampEn is to the value of the radius.

3.3.7 Fuzzy Entropy

For ApEn and SampEn, the subsequence matches must be classified by a Heaviside function which classifies the subsequence as a match or non match – a binary classifier. Chen (Chen, Wang et al. 2007) proposed a new entropy measure based around the notion of fuzzy sets (Zadeh 1965) where a class can have a continuum of grades of membership. The concept requires a measure of membership degree with a fuzzy function returning a value in the range [0,1]. Chen utilized an exponential as the fuzzy measure of the similarity of two vectors. The technique has been used and compared with other entropy measures in time domain EEG and EMG signals (Monge, Gómez et al. 2015)

3.3.8 Windowed Entropy

All of the previous measures of entropy are of the entire signal. It is possible to perform any of these on a window of the signal, then sliding the window along the length of the signal. Like time-frequency analysis, this time-entropy analysis is a tradeoff between resolution in the time domain compared with resolution of entropy changes. The technique has been used, for example, with EEG signal data (Bezerianos, Tong et al. 2003) to detect time point changes in the EEG of rats during cardiac ischaemia. It is a signal analysis technique that has some interesting properties and is dependent upon the variance and the distribution of the observations. As noted earlier, the measure is invariant to scale as normalization occurs as part of the calculation.

Consider for example, a change in the status of recording channel from being grounded to having straight line drift. When grounded, the signal can take no value other than 0 and will have an entropy of 0. Straight line drift would yield a uniform distribution and this, in the discrete case, will give an entropy of logN where N is the number of bins for the calculation. Not only is this scale invariant but it is also independent of the gradient of the drift. In the limit, a straight line is a uniform distribution which will have an entropy

$$H = log(a) - log(b) \qquad 3\text{-}13$$

Where the support is in the interval (a,b). Thus at a specific point in time, the entropy of the signal has changed from 0 to the value given in 3-13. By using a sliding window approach across the signal, it may be possible to detect the point in time where the change occurred and hence infer the point in time that the underlying system changed.

A Normal distribution will have a higher entropy for the same variance. In fact, for a given variance, in the limit, the normal distribution has the highest entropy. The entropy of the window will be an estimate of the entropy at the time represented by the middle of the window. Estimation accuracy is affected by the number of bins used in the formulation of the histogram. Temporal accuracy is affected by the size of the window which is acting like a moving average.

Changes in the entropy represent changes in the structure of the underlying signal as the process driving the recording changes. For example, a change from baseline drift to Gaussian white noise will show a change in the entropy. Pinpointing the temporal location of such a change may give some intuition about changes in the underlying driving process as for example in the EEG example given before.

Figure 3-7 shows a simulation of such a change in process. In the upper panel is a simulated waveform. The first half of the signal is a random variable generated from an exponential distribution and the second half of the signal is a random variable generated from a normal distribution. The mean and variance are the same for both components. The transition point can be seen at time 750 where some of the signal becomes negative.

The entropy calculated by the sliding window method is shown in the lower panel. A sliding window of 300 samples was used with 30 bins. The second half of the signal has a higher entropy – recalling that the normal distribution has a higher entropy than the exponential. The change is smooth because of the averaging over the window length – there will be 299 windows with elements generated from both processes.

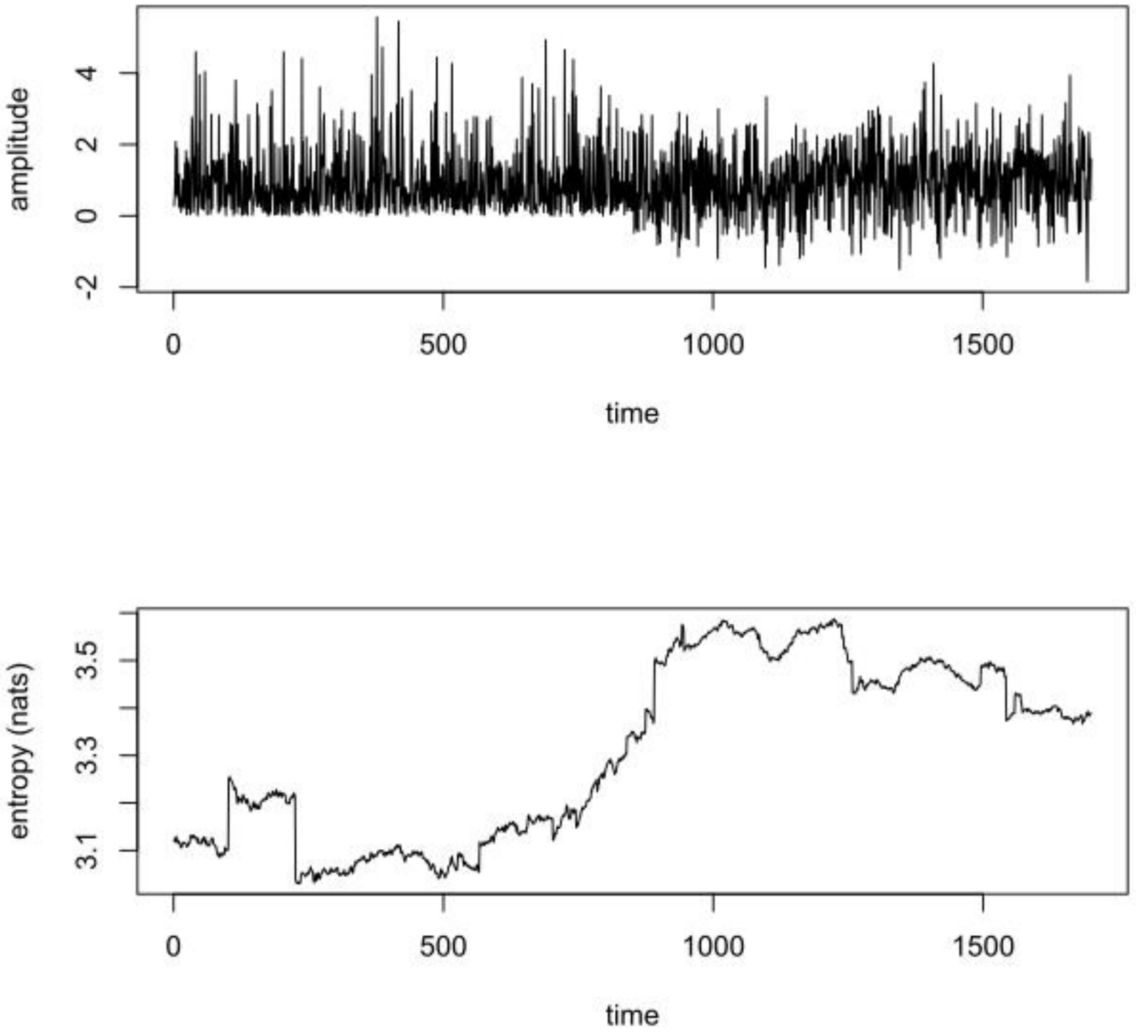


Figure 3-7: Entropy process change simulation

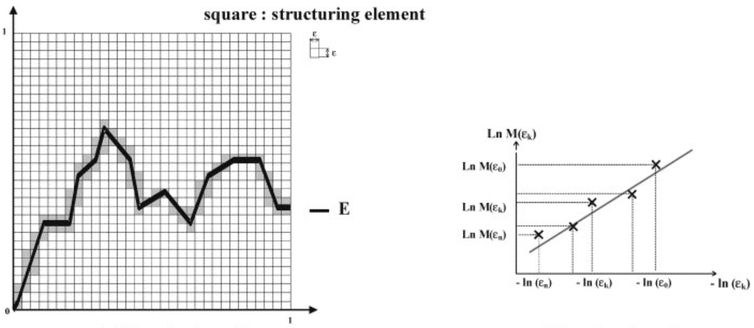
3.4 Fractals

In 1890, building on the work of Cantor (Cantor 1883), Peano (Peano 1890) described a curve that passed through all points in a 2 dimensional space – the Peano curve. This ability of a one dimensional trajectory to fill space was analogous to the ergodic hypothesis of Bolzman (Von Plato 1991) which is usually understood as the assumption that the trajectory of an isolated mechanical system runs through all states compatible with the total energy of the system. The question arose to the dimensionality of the curve – it can be thought of as more than one dimension but less than 2. Hausdorff (Hausdorff 1918) formalized the theory of fractional dimensions in 1919. Mandelbrodt (Mandelbrot 1967) showed that the coast of Britain contained properties of increasing complexity with scale and self-similarity. He calculated that the contour had a Hausdorff dimension of 1.25. He coined the term fractal in 1975 and introduced the concept to the wider public with his book The Fractal Geometry of Nature (Mandelbrot 1983).

Entropy is one method of characterizing the complexity of a system. Fractal analysis is another way to measure complexity – more specifically the rate at which complexity changes with scale. Although taking different approaches, there are some similarities between the two analyses. The two approaches have been compared for the EEG in sleep (Ma, Shi et al. 2018) and Esgiar (Esgiar, Naguib et al. 2002) found a high correlation between fractal dimension (FD) and entropy in the analysis of images of histopathological sections of colon cancers. An object is a fractal if it shows self similarity and scale invariance. A classic example is the analysis of the coast of Britain (Mandelbrot 1967). As one increases the magnification, more detail becomes apparent and the length of the measurement increases. Contrast this with the length of a line, parabola, ellipse or some other regular geometric shape. In those cases, the dimension of the line is 1. In the case of the coast of Britain, the dimension is considered to be more than 1 because as the scale is increased, the length also increases and is in the limit infinite. The rate at which the length changes with scale is the FD and it is equal to the gradient of the line of the log log plot of length vs scale. In the case of the coast of Britain, Mandelbrodt estimated the FD to be 1.25. In contrast, the much smoother curve of 1.13 was found for the Australian coast. Most objects in nature and signals exhibit fractal like behaviour over a limited range of scales and then to only a limited extent. In addition, quantification limits such as – in the case of biological signals – amplitude and time resolution will affect the ability to resolve the object at various scales.

3.4.1 Estimation of Fractal Dimension

Boxcounting dimension. The boxcounting dimension is evaluated by covering the signal with squares of side length r and evaluating how many squares are intersected by the signal as the linear size is increased. Figure 3-8 taken from Zeng (Zeng, Koehl et al. 2001) shows the concept for an arbitrary waveform. Panel (a) shows the waveform with an overlaid grid of squares. The squares intersected by the waveform are shown in grey. This process is repeated at different square sizes and a log-log plot is made with the number of squares intersected plotted against the number of total squares. This is shown in panel (b). The fractal dimension is equal to the gradient of this curve.



(a) Quantization of the space (b) Log-log plot

Figure 3-8: Estimation of fractal dimension by box counting

Katz fractal dimension. Katz was inspired by Madelbrodt and considered an example of the fractal dimension of a river. He considered the fractal dimension (FD) to be related to the length of the river versus the distance between mouth and source. In Madelbrodt's example, the river measurements are in 2 dimensions – for example north and south. Katz (Katz 1988) estimates the FD by taking the log of the ratios of the length of a curve between successive points and the diameter estimated as the distance between the first point of the sequence and the point in that sequence providing the farthest distance. Essentially, he creates a 2 dimensional plane with axes time and amplitude. The algorithm is sensitive to the units used and places too much importance on the first element in the series. There are likely serious flaws in the formula and it has been claimed that it does not estimate the fractal dimension at all.(Castiglioni 2010).

Higuchi Fractal dimension. This method (Higuchi 1990) is often used for timeseries. The length of the signal is considered for each of the members of the ensemble as follows: With every element included , with every second element, with every third element and so on represented as the series corresponding to k=1,2,3 etc. Then if there is a power law relationship between the length and k, the timeseries is fractal with the dimension equal to the exponent. (Cervantes-De la Torre, González-Trejo et al. 2013).

Variogram and Madogram. The variance of a variable z (or signal in this case) is the second central moment of that variable. It can be expressed as:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2 \qquad 3\text{-}14$$

Where z_i is the ith of n value of z and s^2 is the empirical variance

The variance can also be calculated from the difference between all of the pairs in the set i.e.:

$$s^2 = \frac{1}{2n(n-1)}\sum_{all i \neq j}\left(z_j - z_i\right)^2 \qquad 3\text{-}15$$

In this case, the n(n-1) comes from the number of distinct (i,j) pairs in the sample. A variogram is a measure of the variance between different lags of a variable that can be either one or more dimensions. A time series is of one dimension but it can be in 2 dimensions in geography over the (x,y) plane or can be extended to 3 or more dimensions. The madogram is a related measure which uses the absolute rather than the squared difference. In general for a time series:

$$y_p(t) = \frac{1}{2}E|X_u - X_{t+u}|^p \tag{3-16}$$

Where p is the order of the generalized variogram and E is the expectation (mean) function. The variogram is given by order p=2 and the Madogram by order p=1. The sum over all lags of the variogram is simply the variance, so the variogram itself is the variance split up by lags.

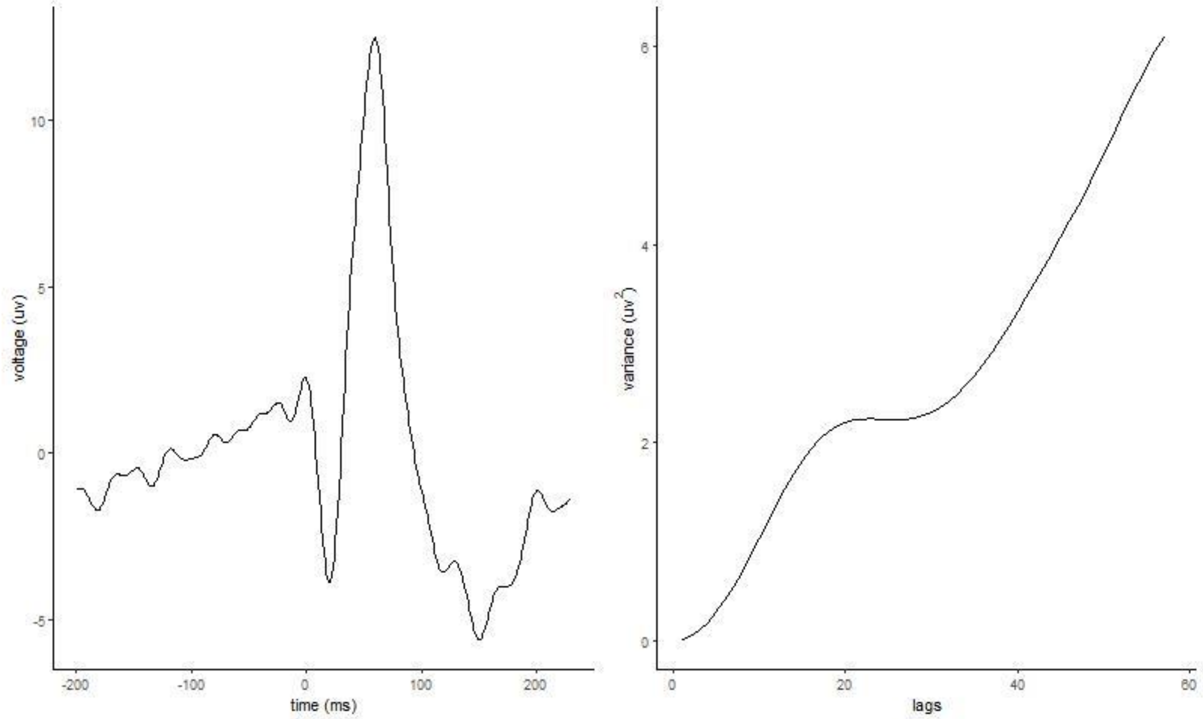


Figure 3-9: ERG and Variogram

In Figure 3-9, a sample ERG trace is shown on the left. The right panel shows the variogram over the first 60 lags. The fractal dimension can be estimated from this function (Gneiting, Ševčíková et al. 2012).

The central idea of a fractal dimension is that upon zooming in, more complexity emerges and the measure (for example length) increases. For a shape to have fractal characteristics, the relationship between scale and measure should be a power relationship such that there is a linear relationship when analysed on a log-log plot.

Although the process driving a biosignals is inherently a continuous time process, the measurement will be discrete time (sampled) with a fixed amplitude resolution from the analog to digital converter. This latter component has a resolution of 24 bits for the LKC device. Even if a biosignal has fractal properties, measuring the signal with a fixed resolution will affect the estimated fractal dimension. Entropy however was initially described and most commonly applied to discrete data although quantization artefact will also affect the estimation of this measure.

This effect is demonstrated in Figure 3-10. A sine wave was simulated and the various complexity measures used in this study were estimated over 100 increasingly fine quantization levels. The original signal is shown in the top right panel, with representative coarse and fine quantization level signals shown in the top middle and top right panels. A sine wave is a smooth curve and has a fractal dimension (FD) of 1. Its entropy depends upon the method of estimation and the number of bins. In this example, the madogram, boxcount and Higuchi fractal dimensions and the Shannon and Distribution entropies show rapid convergence. The madogram shows some aliasing artefact but the vertical scale shows very little change in the FD estimate over the quantization levels. The Katz measure does not converge. The boxcount measure converges but not to the theoretical FD of 1.0. The variogram method shows slow convergence. Figure 3-11 shows the same analysis on a chirp signal with similar results. Aliasing is not evident with this signal for the madogram method but the boxcount method still overestimates the FD. Again the Katz method does not converge. Finally, Figure 3-12 shows the quantization effect for a Gabor function; which is a sine wave windowed by a Gaussian function. The results for this function are similar showing more sensitivity to quantization for variogram FD and permutational entropy.

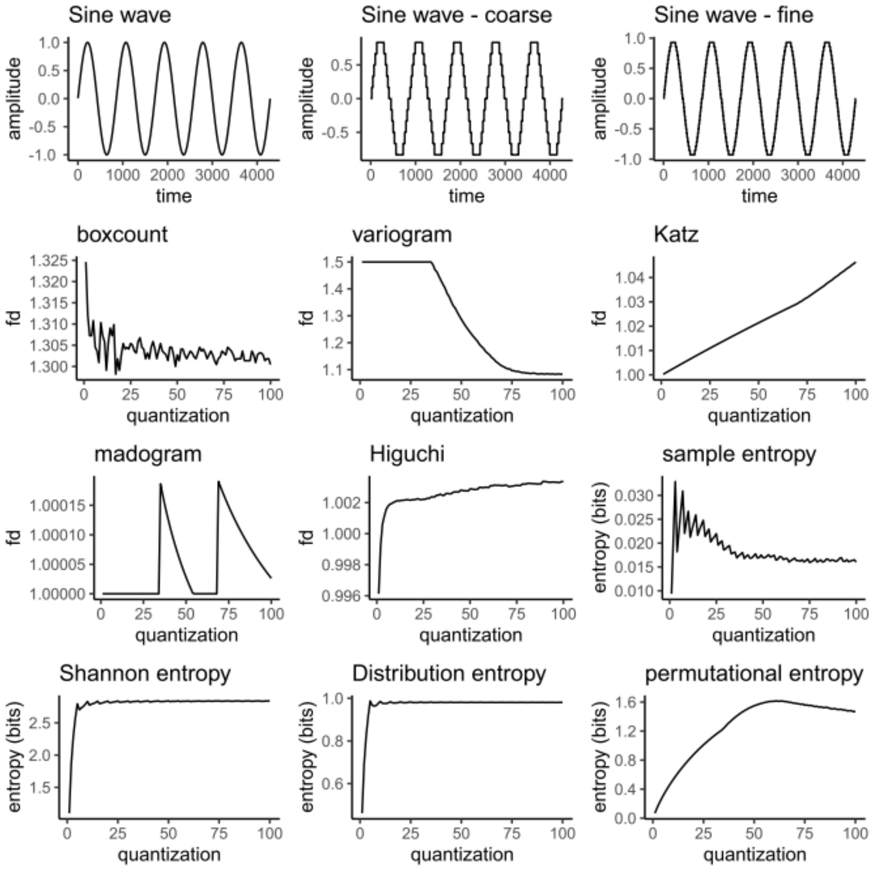


Figure 3-10: Quantization effect on sine wave signal

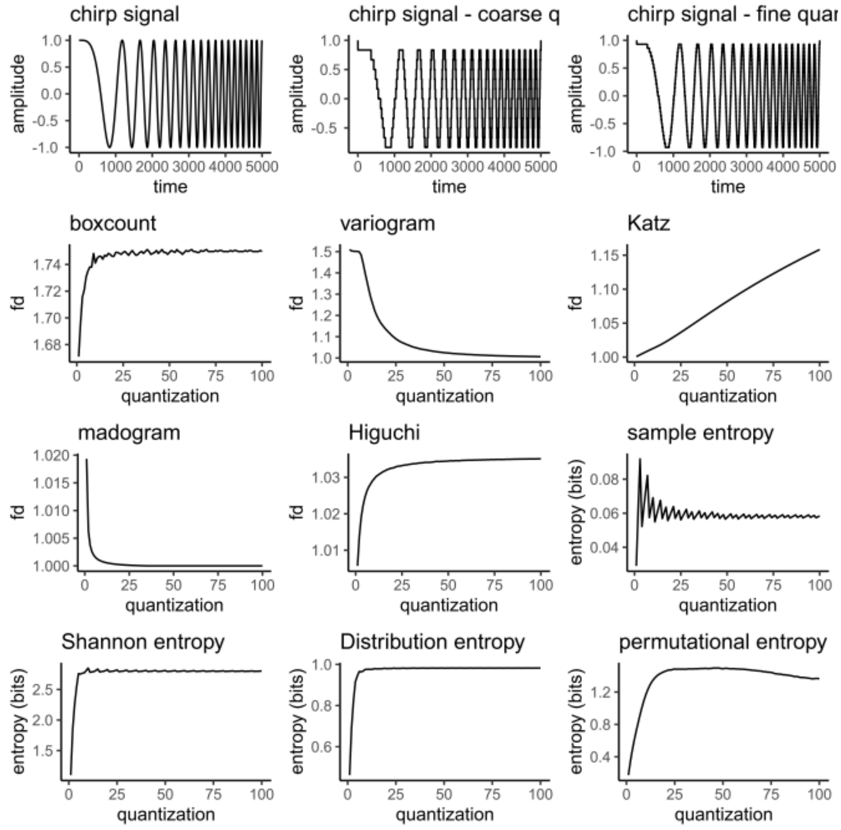


Figure 3-11: Quantization effect on chirp signal

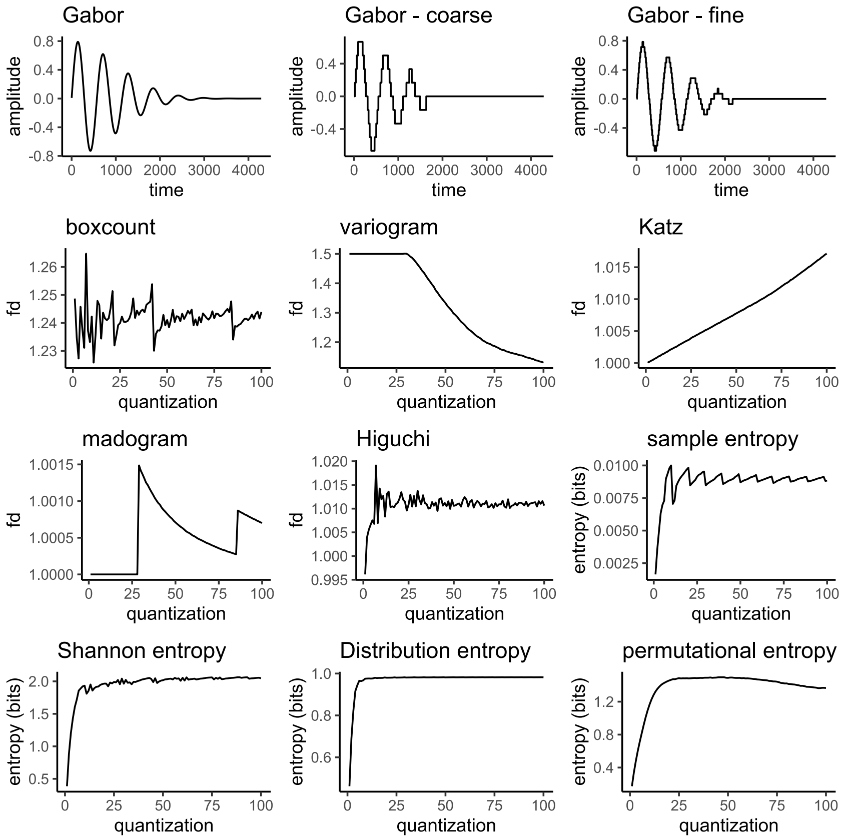


Figure 3-12: Quantization effect on Gabor signal

3.5 Methods

Participants

The study was an approved study by the Royal Victorian Eye and Ear Human Research Ethics Committee. In this study, participants were recruited from the glaucoma unit at the Royal Victorian Eye and Ear Hospital. Inclusion criteria for the glaucoma cohort were a diagnosis of glaucoma with open angles, a best-corrected visual acuity of 6/12 or better, a spherical refractive error within 5.00D and a cylindrical refractive error within 3.00D. Normal controls were recruited amongst family and friends of the glaucoma cohort. A power analysis was conducted in R based on published measures of Approximate Entropy (ApEn) extracted from the ERG in eyes with central retinal vein occlusion. Using the t-test, a significance level of 0.05 and a power of 80%, a group size of 24 was obtained.

Information from the medical record

For the glaucoma cohort, information was sourced from the medical record including the number of medications and the duration of glaucoma Parameters were captured from the visual field including mean deviation.

PhNR Recording

In this study, only left eyes were considered. PhNR responses were obtained from undilated pupils using the LKC device. ERG signals were recorded with an LKC RetEval hand held ERG device (LKC Technologies, Gaithersburg). A red flash of 4ms duration was used as the stimulus on a 10cd/m2 blue background. The device uses an inbuilt infrared camera to measure the pupil size and adjusts the flash luminance with each flash. In this study the flash energy was set to the equivalent of 1.5cd.s/m^2 through a 7 mm diameter pupil. This flash energy was chosen because the portable device could not produce a higher energy while allowing for the full range of pupil diameters. The sample rate of the device was set to 1, 900 Samples per second at 24 bits resolution, an epoch of 220ms and pre-stimulus acquisition starting at 100ms. 200 sweeps were performed for each session. Proprietary 3 pole skin electrodes were used which incorporated active, indifferent and earth electrodes. B-wave amplitude (the first positive going wave measured relative to the a-wave) and implicit time were measured with the provided device software.

Files were exported from the device at the conclusion of each testing session and the raw data extracted with the program rffbrowser (LKC Technologies, Gaithersburg). This program makes available the filtered signal expressed as mv as a function of time. It also allows each raw, unfiltered sweep to be analysed. These sweeps are in bits as a function of index into the recording vector.

The a-wave, b-wave and PhNR amplitude and implicit times were calculated from the processed waveform produced by the device and corresponded to the first negative going wave and the first positively going wave after the stimulus for the a-wave and b-wave respectively. The a-wave amplitude was measured relative to the detrended baseline and the b-wave was measured relative to the a-wave. The PhNR was defined as the first negative going wave after the b-wave.

Mathematical Analysis

Data for each recording session consisted of a matrix of 24 bit values, with 200 columns corresponding to the individual sweeps and 568 rows corresponding to the timepoints of the time series. All sweeps and filtered averages were stored in a SQL server database and processed in R via the DBI interface.

Calculations were performed in R (R Developer Core Team 2017) using the entropy (Hausser 2014) package or in Matlab. Specific details of the analytical techniques follows:

Shannon entropy was estimated from the discretised raw data in R using 20 bins. Permutational entropy was calculated for sequence lengths 3 to 9 using Matlab with the

Entropyhub library (Flood and Grimm 2021). Sample entropy was also calculated in Matlab with the same package.

Windowed entropy was calculated by estimating the Shannon entropy and SampEn within a sliding window of epoch length 150 samples. For the Shannon measure, discretization was performed for 30 bins and entropy estimated with the plug-in estimator. This was performed for each sweep and the overall entropy for each window position was determined from the mean of all of the corresponding windows. The SampEn was calculated in a similar manner with a radius of 0.2 times the standard deviation. The magnitude of the entropy change and its timing were extracted as features. DistEn was calculated as per Karmakar (Karmakar, Udhayakumar et al. 2017).

For the fractal analysis, Madogram and boxcounting fractal dimension (FD) were calculated using the Fractaldim package in R (Gneiting, Ševčíková et al. 2012). Higuchi and Katz fractal dimensions were calculated by custom code written in R.

Statistical Analysis

For each group, differences in means were calculated for age, and ERG parameters. Each parameter was tested for normality with the Shapiro-Wilk test (Shapiro and Wilk 1965). Comparisons of parameters that met the normality constraint were analysed with the two way t test except for b-wave amplitude which is known to not be normally distributed (Lee 2013). Allowance was made for the significance of comparisons using the method described by Benjamini for dependent test statistics (Benjamini and Yekutieli 2001) using the mutoss package in R (Werft 2017).

3.6 Results

3.6.1 ERG recordings

The filtered, averaged ERG recordings in the dataset were generally of high quality as shown in Figure 3-13

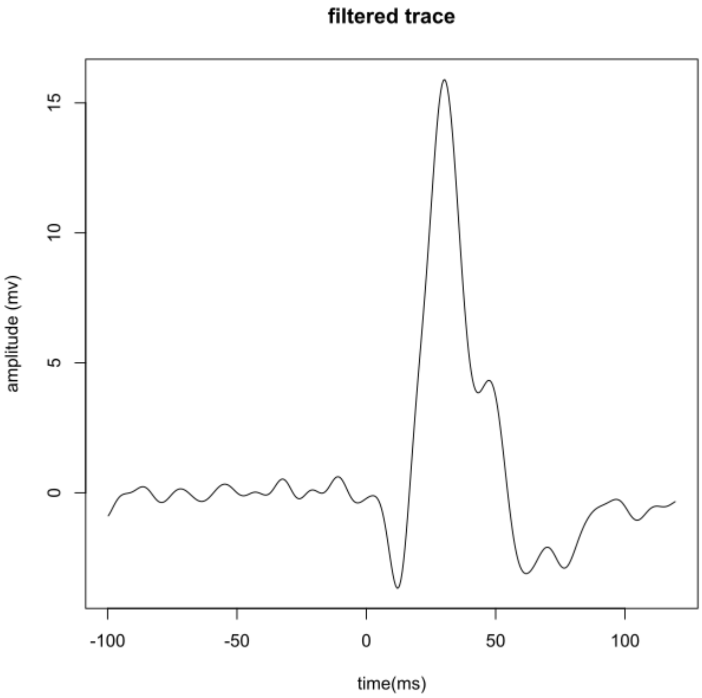


Figure 3-13: Sample average from RetEval

The unfiltered sweeps however were contaminated with a large amount of 50Hz mains frequency noise. Figure 3-14 shows a single sweep in raw analog to digital conversion (ADC) bits. This is a result of the device being hand held and not connected to any ground source. This 50Hz noise does average out with enough traces but its presence is suboptimal. A power spectrum of the sweep with the DC removed is shown in Figure 3-15 demonstrating the high level of 50Hz noise contamination. This was consistent for most of the sweeps in the dataset.

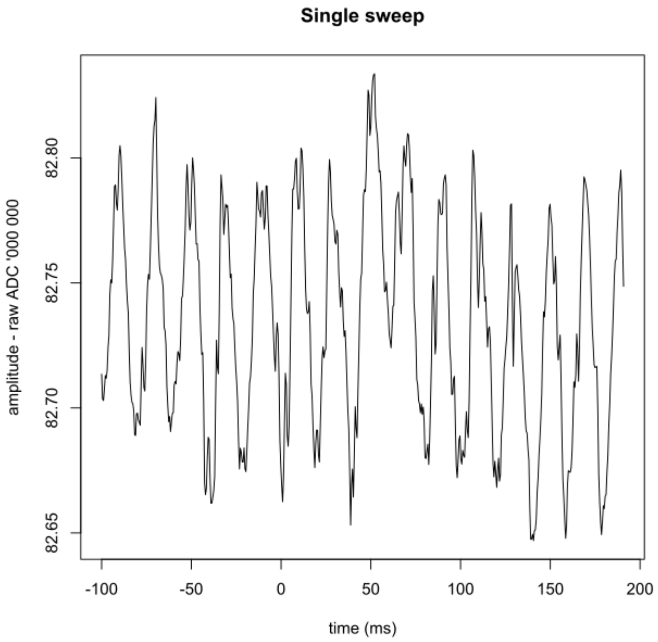


Figure 3-14: Single unprocessed sweep

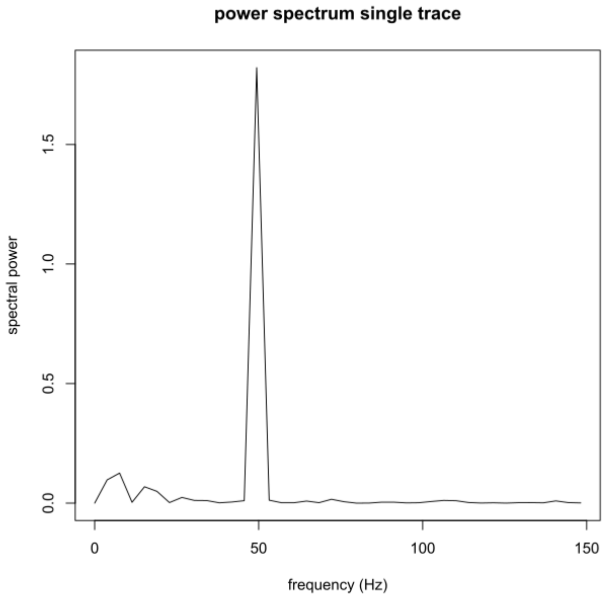


Figure 3-15: Power spectrum of single sweep.

3.6.2 ERG wave parameters and PhNR values

The results for the erg waves and PhNR for the dataset are shown in Table 3-1.

Table 3-1: ERG results PhNR = photopic negative response. Significance adjusted to allow for multiple comparisons via the method of Benjamini (Benjamini and Yekutieli 2001). Continuous variables show mean (standard deviation).

Result	**Controls**	**Glaucoma**	***P***
Test subjects	18	21	-
Age range (years)	44.3-87.6	30.8-86.3	
Age (years)	62.6 (13.9)	65.3 (14.1)	0.44
Mean deviation (dB)	-5.37 (4.73)		
b amplitude (uV)	31.1	30.7	0.66
b latency (ms)	19.3	17.1	0.26
PhNR (uV)	-5.50 (1.95)	-3.60 (2.53)	0.03

There was no significant difference in a-wave or b-wave amplitude or latency between the groups. The PhNR showed significant difference between the two groups with the measure being far more negative in the control group – that is a significant reduction in the glaucoma group compared with the controls. The variance of the measure was much higher in the glaucoma group. This is shown in Figure 3-16

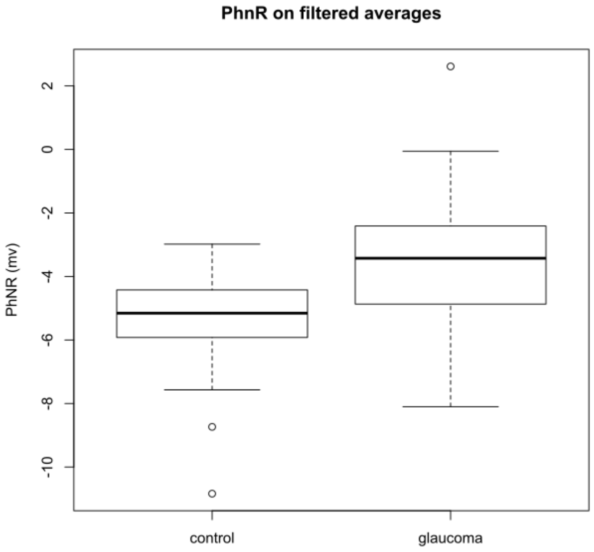


Figure 3-16: Boxplots of PhNR filtered averages.

3.6.3 Averaged Complexity Measures on the unfiltered sweeps

Table 3-2: Complexity measures raw data

Result	**Controls**	**Glaucoma**	***P***
Test subjects	18	21	-
Boxcounting FD	1.692 (0.102)	1.680 (0.100)	0.98
DistEn order 3	0.882 (0.030)	0.875 (0.029)	0.98
Higuch FDi	1.338 (0.148)	1.370 (0.150)	0.98
Katz FD	2.147 (0.113)	2.147 (0.132)	0.98
Madogram FD	1.177 (0.099)	1.182 ((0.088)	0.98
PermEn order 4	0.741 (0.106)	0.757 (0.088)	0.98
SampEn	0.680 (0.132)	0.712 (0.136)	0.98
Shannon Entropy	3.628 (0.122)	3.601 (0.125)	0.98

Notes : FD= fractal dimension, DistEn = distribution entropy order 3, PermEn = permutational entropy order 4, SampEn = Sample entropy with radius 0.2 x standard deviation. Significance adjusted for multiple comparisons via the method of Benjamini (Benjamini and Yekutieli 2001)

Table 3-3. Sensitivity, Specificity and AUC of complexity measures on averaged traces

Measure	**Sensitivity**	**Specificity**	**AUC**
Boxcounting FD	0.571	0.556	0.466
DistEn order 3	0.571	0.556	0.437
Higuchi FD	0.619	0.611	0.601
Katz FD	0.619	0.667	0.542
Madogram FD	0.857	0.333	0.524
PermEn order 3	0.762	0.444	0.54
SampEn	0.667	0.611	0.598
Shannon Entropy	0.619	0.556	0.439

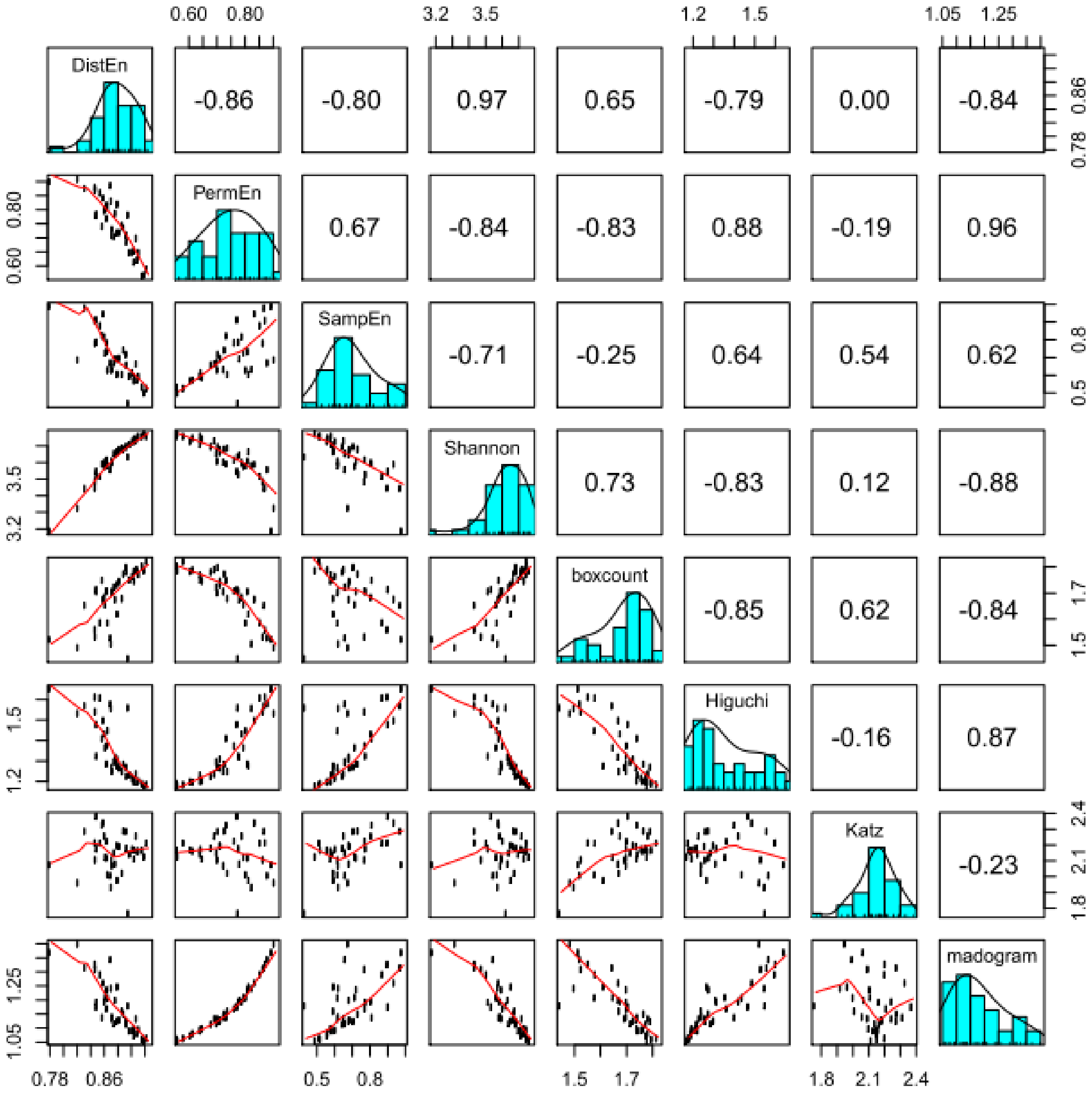


Figure 3-17: Pairs plot on raw sweep data.

Table 3-2 shows the differences between the complexity measures when calculated on each sweep and averaged. Table 3-3 shows the corresponding sensitivity, specificity and area under the receiver operating characteristic curve (AUC). The cutpoints for sensitivity and specificity were selected by maximising the sum of the sensitivity and specificity for each measure using the cutpointr library in R (Thiele and Hirschfeld 2021). There were no significant differences between the glaucoma and control cohorts on any of the complexity measures. Some of the measures correlated strongly with other measures as can be seen in the pairs plot shown in Figure 3-17.

3.6.4 Complexity measures on filtered and averaged sweeps

Figure 3-18 and Table 3-4 shows analysis on the averaged, filtered traces. Of note is that the entropies are generally lower. Significant differences were detected in the Shannon and DistEn and in the Higuchi fractal dimension after correction for multiple comparisons. The sensitivity and specificity analysis indicates that the PhNR remains a superior measure to any of the complexity measures.

Table 3-4: Complexity measures on averaged filtered traces.

Result mean (sd)	**Controls**	**Glaucoma**	***P***
Test subjects	18	21	-
PhNR (uV)	-5.50 (1.95)	-3.60 (2.53)	0.03
DistEn order 3	0.810 (0.062)	0.760 (0.032)	0.050
Shannon entropy	2.99 (0.198)	2.80 (0.259)	0.010
Higuchi FD	1.068 (0.019)	1.082 (0.017)	0.050
Katz FD	1.005 (0.003)	1.007(0.004)	0.106
Madogram FD	1.000 (0.001)	1.001 (0.001)	0.212
PermEn order 4	0.292 (0.013)	0.297 (0.016)	0.248
Box counting FDt	1.148 (0.035)	1.157 (0.061)	0.410
SampEn	0.0776 (0.023)	0.0775 (0.024)	0.618

Notes: Significance adjusted for multiple comparisons via the method of Benjamini (Benjamini and Yekutieli 2001). The measures are shown in the order corresponding to Table 3 5

Table 3-5. Sensitivity, Specificity and AUC on averaged traces.

Measure	**Sensitivity**	**Specificity**	**AUC**
PhNR	0.667	0.833	0.741
DistenEn order 3	0.944	0.571	0.741
Shannon Entropy	0.944	0.476	0.685
Higuchi FD	0.81	0.500	0.667
Katz FD	0.571	0.722	0.659
Madogram FD	0.476	0.833	0.627
PermEn4 order 4	0.619	0.722	0.614
Boxcounting FD	0.333	0.889	0.520
SampEn	0.722	0.429	0.511

Notes: Measures shown in order of decreasing Area under the receiver operating characteristic curve (AUC)

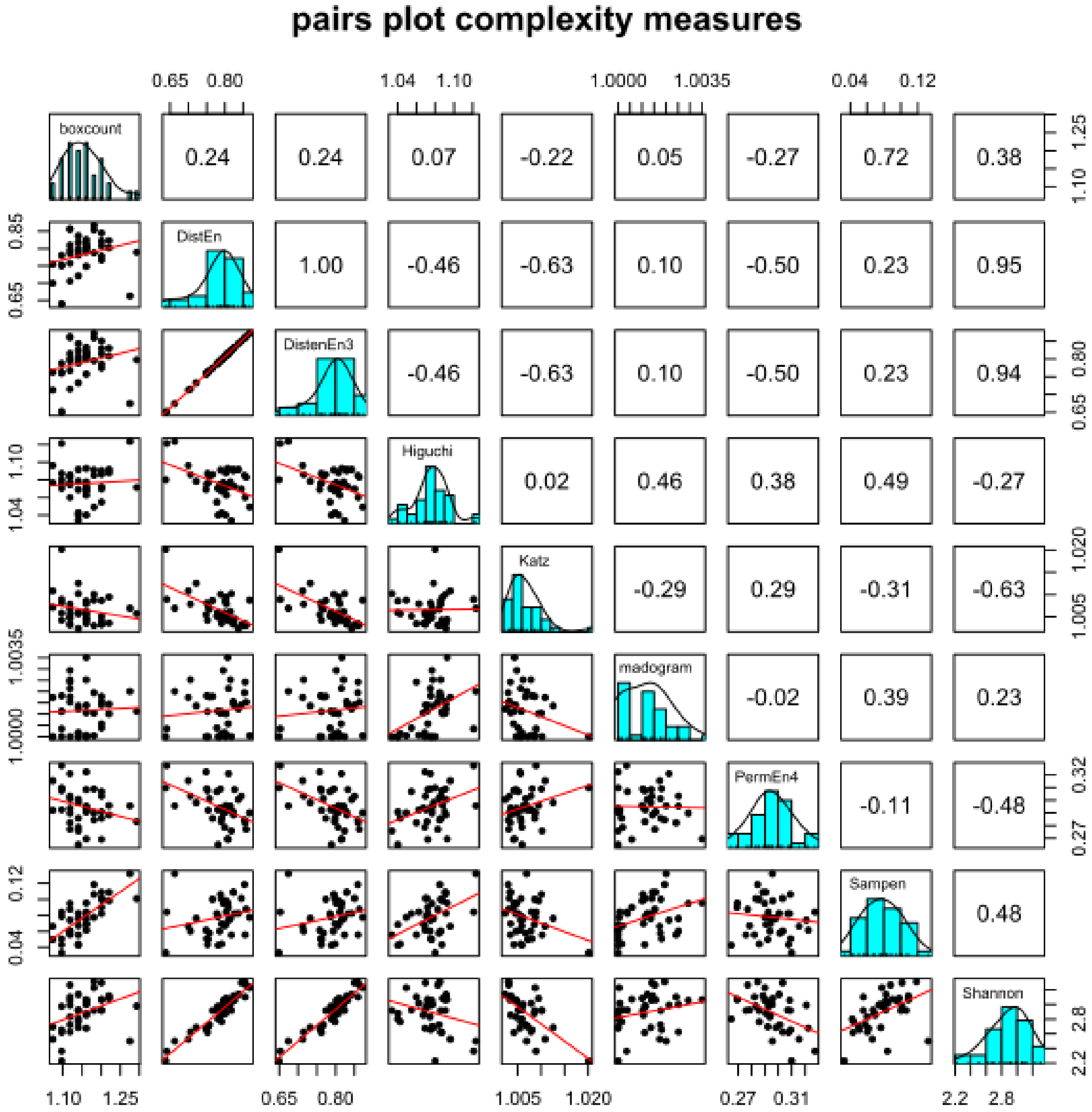


Figure 3-18: Complexity measures computed on ensemble averages.

3.6.5 Permutational Entropy

Figure 3-19 shows the permutational entropy on the averaged traces as a function of the order for both controls and glaucoma. The entropy value increased in a linear fashion with regard to sequence length. There was no significant difference between any order of permutational entropy when eyes with glaucoma were compared to those from undiseased eyes. Computational load increased very quickly as a function of permutational order reflecting the fact that the number of permutations increases as a function of the factorial of sequence length. There were no significant differences between glaucoma and control subjects for any of the permutational orders and the significance tended to fall with increasing orders.

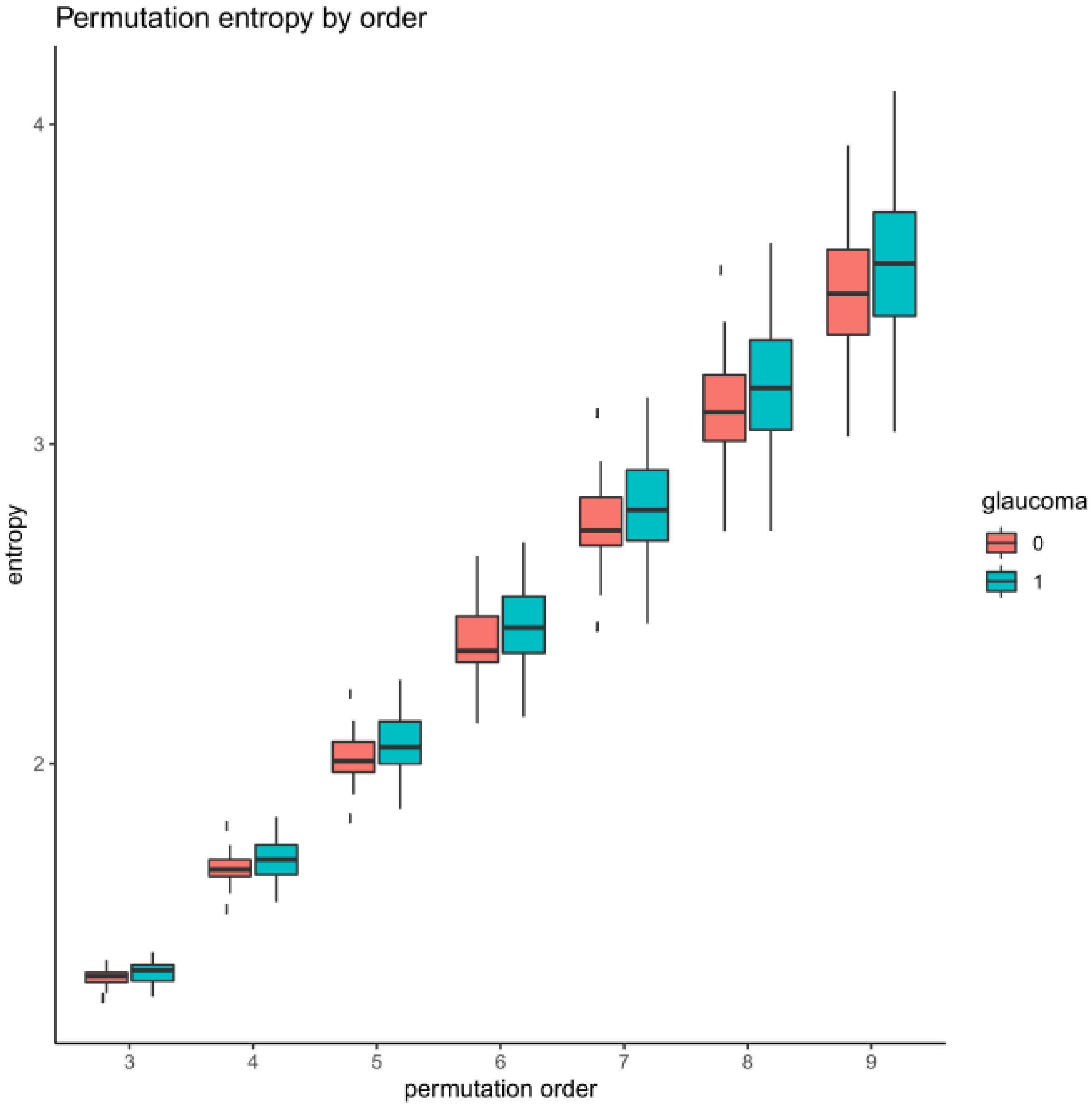


Figure 3-19: Permutation entropy by order - glaucomas vs controls.

Figure 3-20 shows the relationship of PermEn calculated on all test subjects as a function of order. There are very high correlations between the calculated PermEn values for all combinations of PermEn order.

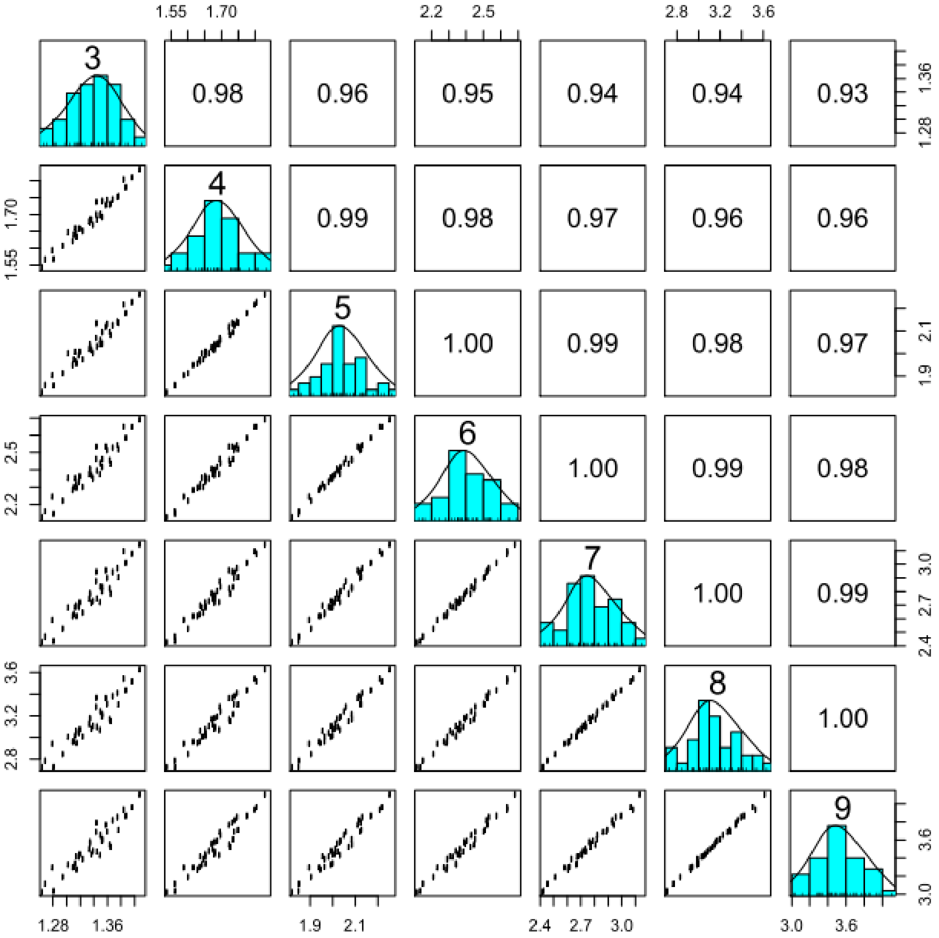


Figure 3-20: Permutation entropy on all test subjects by order. The numbers on the diagonal correspond to the order of PermEn with histograms shown for each order. The figures on the top right half are the Pearson correlation coefficients. Scatterplots are shown on the bottom left half of the diagram with the values of the calculated entropy shown.

3.6.6 Sample Entropy

SampEn was specifically developed for biological timeseries analysis and therefore remains the natural choice for analysis of the ERG. It has the advantage over many other measures of entropy of utilizing the order of values – although over a very short range (the embedding dimension plus the time lag). Configurable parameters with the algorithm include the embedding dimension (the subsequence length m to be compared with sequence m+1), the time lag and the radius. The radius is the threshold for equality of sequences. If the distance between sequences is less than the radius, the sequences are considered equal. This is usually expressed in a multiple of the standard deviation of the values and the default is 20. Figure 3-21 shows the relationship between SampEn over all filtered averages (glaucoma and controls) and a range of radii. The mean and variance decrease with increasing radius.

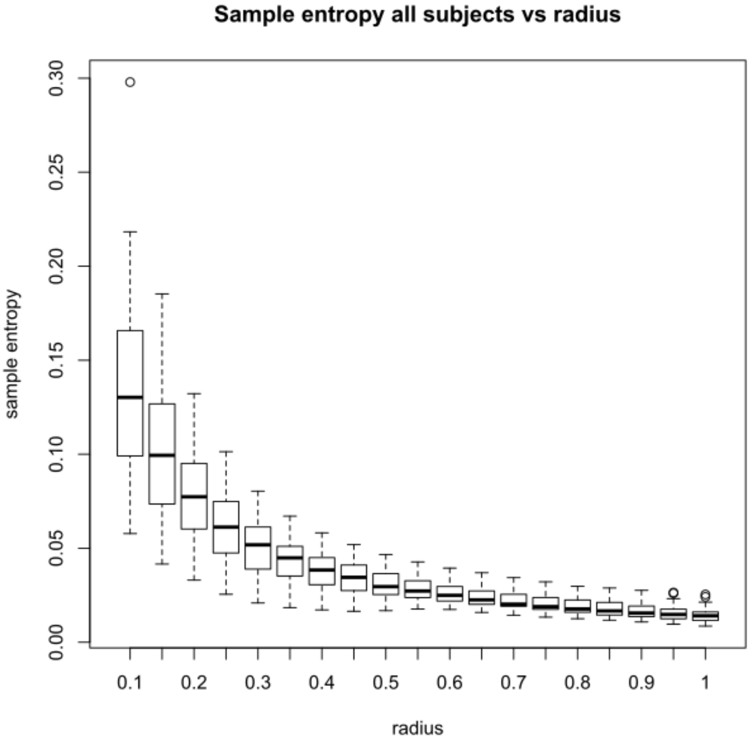


Figure 3-21: Box plots of Sample entropy (SampEn) for the entire cohort with different radii. The radius is the threshold below which the sequences specified in the algorithm are considered similar. Here it is expressed in units of multiple of the standard deviation. The graph shows a reduction in the magnitude and spread of SampEn with increasing radius

Figure 3-22 shows the SampEn as a function of radius for the glaucoma and non-glaucoma groups separately. The differences are small at all of the radii.

Figure 3-23 shows the effect of the radius on the significance of the difference in SampEn between controls and glaucoma subjects. There was an interesting biphasic difference in the function with a radius of 0.7 nearly reaching statistical significance of 0.05.

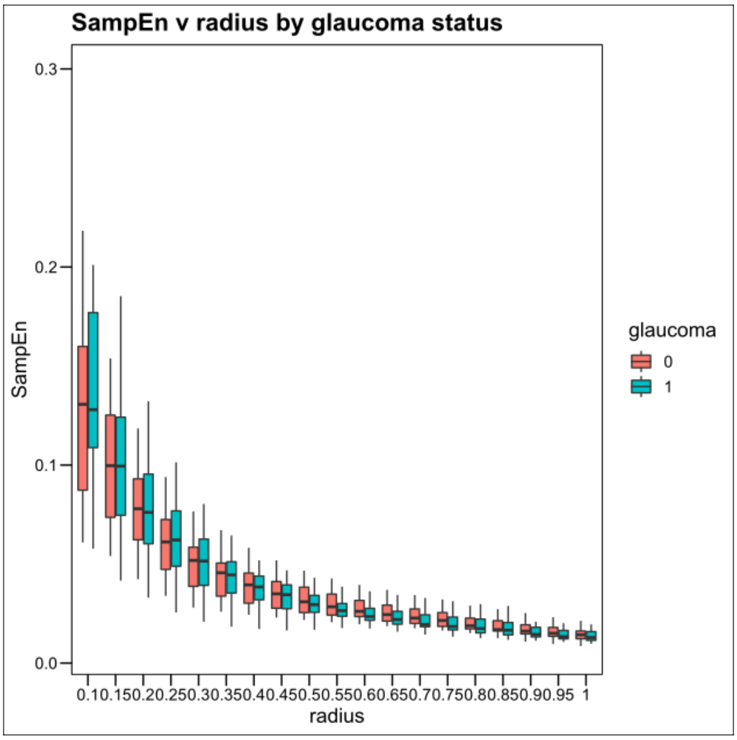


Figure 3-22: SampEn by radius and glaucoma . This is the same plot as Figure 3-21: Box plots of Sample entropy (SampEn) showing the boxplots of the glaucoma and non glaucoma cohort with different radius values – radius defined as a multiple of the standard deviation. Like the whole cohort plot, the magnitude and spread of the measurements decreases with increasing radius.

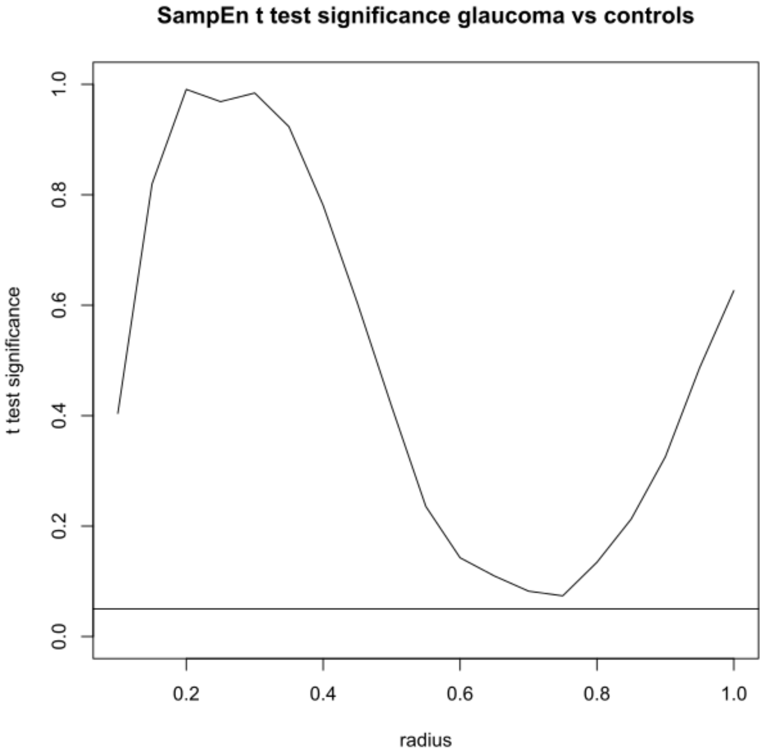


Figure 3-23: Plot of the significance of differences in sample entropy (SampEn) between the glaucoma and non-glaucoma conhorts measured by t-test calculated at the radii shown Figure 3-22. The difference between the cohorts nears significance when the radius is 0.7 times the standard deviation.

3.6.7 Fractal Dimension

Fractal dimension (FD) calculations showed strong correlations between FD estimated by the Higuchi method, Madogram and box counting methods when performed on the individual sweeps (Figure 3-17). The Katz method did not correlate with these measures. The FDs were much less well correlated when computed on the ensemble averages (Figure 3-18). This difference may relate to the quantization limit when computing FD.

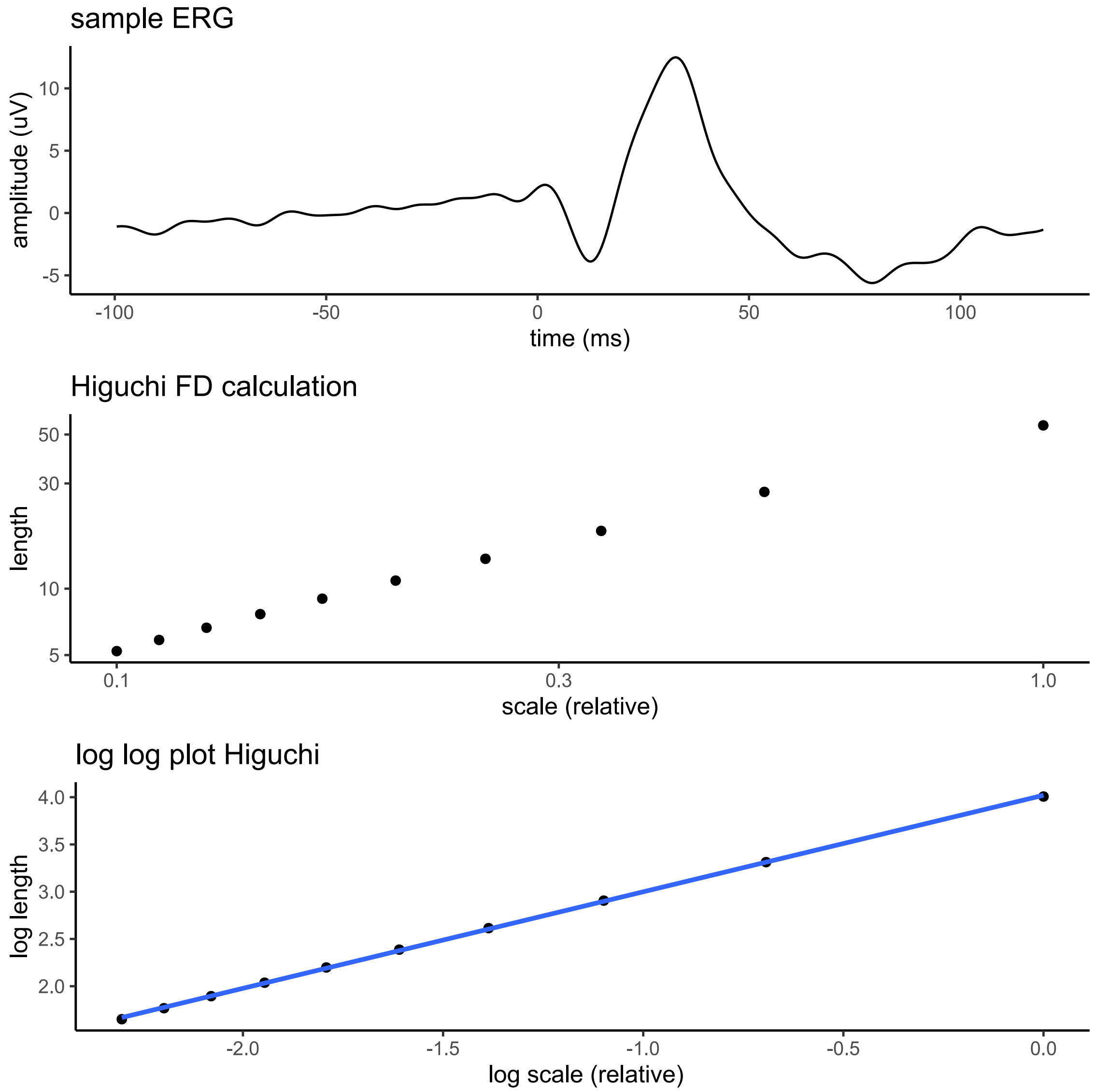


Figure 3-24: Higuchi FD calculation on a sample trace

Figure 3-24 shows the details of FD calculation on a sample ERG trace by the Higuchi method. The top pane shows the sample ERG average, plotted as a function of time. The middle pane shows the length of the curve calculated by the Higuchi method plotted as a function of scale. The axes of the plot are logarithmic. The lower pane shows the same measurements with the base 10 logarithm taken of the length and scale and replotted on a linear scale. The FD is the gradient of the line of best fit. The goodness-of-fit of the line is the extent to which the signal shows fractal properties, or can be used to determine the range of scales for which such a property exits.

It is perhaps unsurprising that there was little difference between the glaucoma and non glaucoma subjects when the signal was characterized by a single scalar parameter. This is analogous to characterizing the ERG signal by its variance or mean alone.

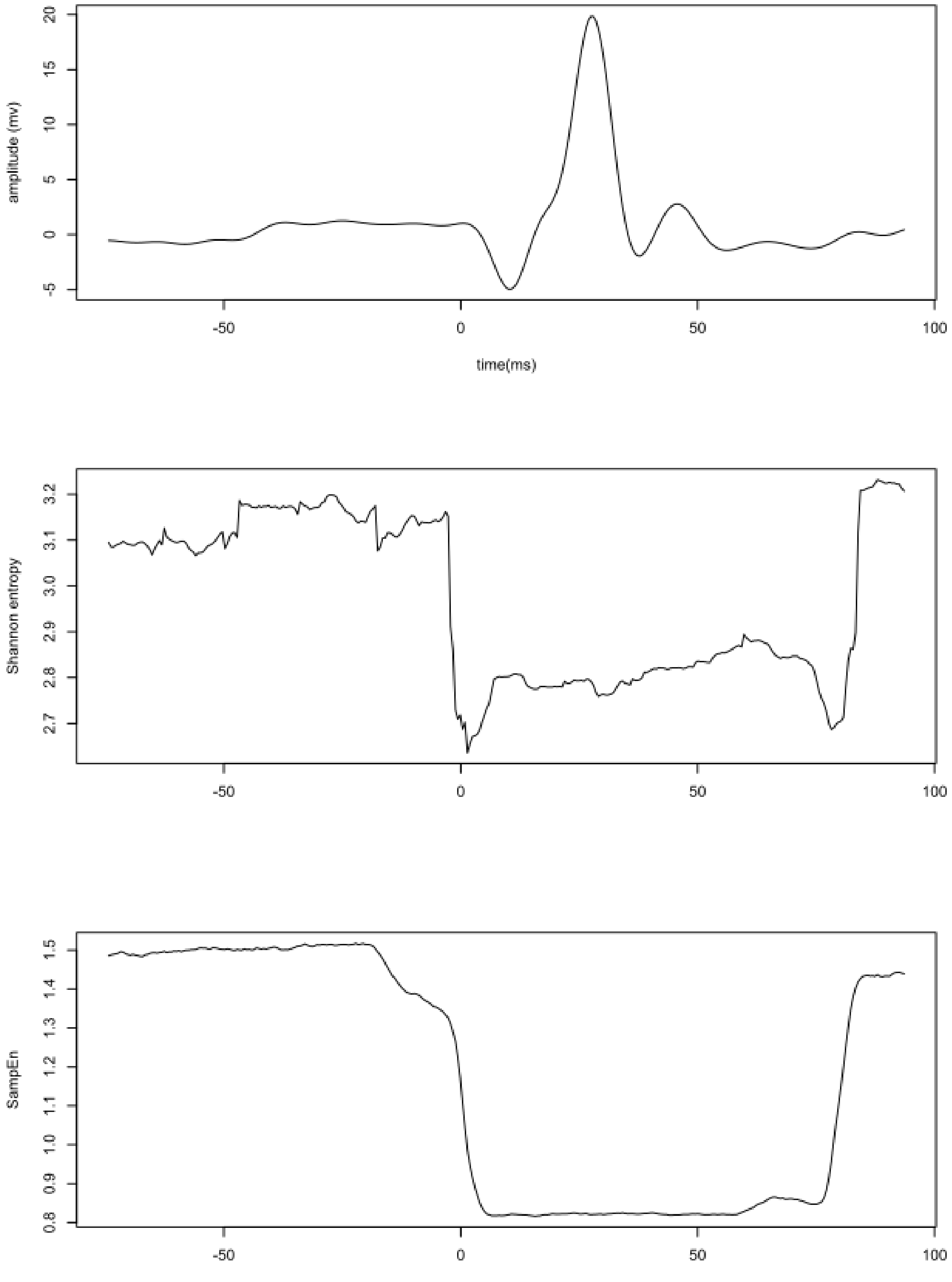


Figure 3-25: Entropy by sliding window - example subject

The entropy calculations, both with Shannon and SampEn within the sliding window showed that typically the entropy of the signal dropped for a period post stimulus and then rose again. A typical example of a control patient is shown in Figure 3-25. The timepoint corresponding to each sliding window is taken to be from the middle of the window. The upper panel shows the ERG and the two lower panels the average of the windowed entropy for all of the sweeps.

As with any sliding window technique, temporal resolution is lost in proportion to the window width. This is the reason why the entropy appears to drop before the onset of the stimulus and is related to the fact that windows centred earlier than the flash onset are influenced by changes in the window ahead of the centre point.

There was no significant difference detected in the onset of the entropy drop or in the magnitude of the change. There was however a significant difference in the terminal rise in entropy as shown in Table 3-6.

Table 3-6: Entropy changes by sliding window

Result	**Controls**	**Glaucoma**	***P***
Test subjects	18	21	
b-wave amplitude (mv)	31.1 (2.2)	30.7 (3.6)	0.66
Time of Shannon entropy rise	78.8 (5.2)	75.1 (5.1)	0.05
Time of sample entropy rise(ms)	79.8 (6.8)	74.7 (5.9)	0.04

3.7 Discussion and Summary

Complexity analysis has developed over the last 50 years (Fernández, Gómez et al. 2013) and has found utility in areas as diverse as traffic flow analysis (Kumar, Tchier et al. 2018), earthquake analysis (Hirata 1989) and bearing fault detection (Zhang, Li et al. 2012). One dimensional biosignals such as the electroencephalogram (EEG), electromyogram (EMG) and electroretinogram (ERG) can be analysed for complexity by techniques such as entropy and fractal dimension.

The electroretinogram is a massed response of the retina to a photic stimulus and all of the electrically active layers of the retina contribute to the response. The photopic negative response which peaks after the b-wave is known to have its amplitude reduced in glaucoma and is produced by the spiking of retinal ganglion cells(Viswanathan, Frishman et al. 1999). Parameters that may affect characteristics of the massed response could include the number of ganglion cells firing, changes to the firing rate or pattern or changes in the relationship of firing patterns between ganglion cells.

Much work has been done on complexity analysis of the EEG. Takahashi (Takahashi, Cho et al. 2009) using multiscale entropy measures found that photic stimulation increased the entropy in younger but not older test subjects. Muller (Müller, Lutzenberger et al. 2001) found the complexity of the EEG to be increased in subjects with Parkinson's disease compared with controls. In the area of schizophrenia, a number of groups have analysed EEG signals for complexity (Takahashi, Cho et al. 2010, Fernández, Gómez et al. 2013) and have found changes that may be reversible with medications.

In this study, a significant difference between subjects with glaucoma and controls was found in the complexity measure of the ERG as measured by both Shannon and Distribution entropy (DistEn) and by Higuchi fractal dimension. The Shannon and DistEn are highly correlated measures.

The central pathology in glaucoma is the loss of ganglion cells (Quigley 2019). Such a loss of the number of sources of the signal could account for changes in the signal's complexity. All of the glaucoma group in this study had established glaucoma and therefore substantial loss of ganglion cells. It has also been shown that the excitability of ganglion cells affected by intraocular pressure changes - initially increasing and ultimately decreasing (Risner, Pasini et al. 2018). This is manifested by changes in the mean rate of discharge and threshold of excitation. It has been shown in the rat that the on and off ganglion cells are affected by intraocular pressure to different extents (Wang, Wong et al. 2021) so this would inevitably lead to a change in the nature of the massed signal following a photic stimulus which may account for some of the changes in the complexity measures in this study. Klistorner and Graham (Klistorner and Graham 1999) showed that the ganglion cells projecting to the magnocellular layers of the thalamus were more susceptible to glaucoma and a differential loss of cells of this type may contribute changes in signal complexity.

The discharge patterns of ganglion cells may affect the complexity of the ERG signal. It is known that ganglion cells can have bursts of oscillatory activity (Kerschensteiner and Wong 2008) and that these retinal waves can lead to synchronization of ganglion cell discharges. Basar (Basar 1983) showed that neural synchronization in the central nervous system leads to reduced complexity measure parameters derived from the EEG. Although amacrine cell counts are not affected by experimental glaucoma (Kielczewski, Pease et al. 2005), it is not known whether increased synchronization occurs in ganglion cell discharge patterns in the disease.

This study demonstrates that the principles of information theory and fractal analysis can be applied to the ERG. Although some differences were found between glaucoma and non glaucoma for some measures, the significance was lower than a simple measurement of the PhNR. Some of this can be explained by the high level of 50Hz contamination in the raw data. Many of the complexity measures were highly correlated with each other and specifically, some types of fractal measures were well correlated with entropy measurements. Despite the very different theoretical underpinnings of the two measures, both seek to quantify signal complexity and therefore some correlations are to be expected.

Interpretation of the results is much more difficult that a simple measurement of the magnitude of a-wave such as the PhNR. As shown with the dependence of SampEn on radius and of Shannon entropy on bin number, the parameters would likely need careful tuning for each protocol on each type of ERG equipment.

This study has however shown an interesting result that the changes in entropy following the system perturbation by a stimulus is of shorter duration in glaucoma patients than in controls. It is known that glaucoma is a disease of the ganglion cells and that these are the only spiking cells in the retina. The cells have a background discharge which is a random stochastic point process where the probability distribution of the spike counts within an interval are distributed in a way well approximated by a Poisson process. Summed as a whole, the system is a chaotic one and is one that should be able to be modelled by information theory.

The drop in entropy following a stimulus is likely explained by increasing coherence levels from the electrically active cells in the retina following the stimulus. In this sense, the entropy drop represents an increase in signal and decrease in noise as a result of the stimulus.

An explanation for this temporal shift could be the fact that the cells most affected by glaucoma are those subserving the more peripheral field. These are the cells with the longest axons. The action potentials from these axons must travel a larger distance within the eye and this sustained duration is one possible explanation for the fact that the entropy falls faster when these cells are not properly functioning.

The ultimate goal of electrophysiology in the context of glaucoma is to derive features from the waveform that have some predictive value when fed into a classification or regression model. Although beyond the scope of this study, it may be that a complexity measure has some benefit as a predictor in a model that includes other features from the individual such as OCT and visual field features. The Shannon entropy measure can be calculated very fast without specialized software and only requires a single parameter – the number of bins. This fact makes the Shannon entropy the most likely candidate as a feature for such a model.

There are limitations to the use of fractal analysis for biological signals (Gitter and Czerniecki 1995). No biological signal is truly self-similar over an infinite range (Glenny, Robertson et al. 1991) although they may show fractal like properties over a limited range – such as the range considered in this study. The effect of temporal and amplitude resolution must be considered. The hand held portable ERG unit used in this study was convenient and yielded good quality ensemble average traces but each sweep was significantly contaminated with powerline noise and this meant that complexity calculations on individual traces did not yield satisfactory results.

The study did show a difference in the magnitude and timing of some complexity measures but ultimately these were less significant than the well established photopic negative response measure which is much more intuitive and much less computationally difficult. This study did not explore how these parameters vary with glaucoma severity. Another limitation of this study was its small sample size. The use of the skin electrodes and the lack of grounding of the RetEval system resulted in high amplitude noise and a low amplitude signal and this led to a lack of power in classifying the glaucoma and non glaucoma cohorts with any electrophysiological measure. Further work will involve developing a classifier using a larger dataset, DTL fibre electrodes and the Espion electrophysiology equipment.

The study has shown that novel features that can be extracted from the ERG are associated with the presence of glaucoma. Further studies may show that these features, in combination with other features from the history, OCT, visual field or electrophysiological tests can predict the severity or risk of progression of glaucoma and could therefore inform the clinician of a target pressure or surveillance interval.

Chapter 4

Time Frequency Analysis of the Electroretinogram

The findings of this chapter have been published in part (see Appendix)

Sarossy M, Crowston J, Kumar D, Weymouth A, Wu Z. Prediction of glaucoma severity using parameters from the electroretinogram. Scientific reports. 2021;11(1):1-9.

Sarossy M, Crowston J, Kumar D, Weymouth A, Wu Z. Time-frequency analysis of ERG with discrete wavelet transform and matching pursuits for glaucoma. Translational vision science & Technology 2022;11(10):19-19.

4.1 Abstract

Background and Objective

Severity of glaucoma is currently assessed with a combination of features of optic nerve structure derived from optical coherence tomography (OCT) and functional features derived from standard automated perimetry (SAP). Electrophysiology is an alternative and complementary way to assess function in glaucoma. This chapter examines the performance time domain and time-frequency domain feature extraction techniques applied to the electroretinogram (ERG) for the prediction the glaucoma severity.

Methods

ERGs targeting the photopic negative response (PhNR) were obtained in 103 eyes of 55 patients with glaucoma. Amplitude features were derived from the ERG with identification of the a-wave, b-wave, i-wave and the PhNR1 and PhNR2. Multiple time-frequency extraction techniques were also assessed including the discrete wavelet transform (DWT) and the matching pursuit (MP) decomposition. Texture and superpixel analysis was performed on the continuous wavelet transform of the ERG. Linear and multivariate adaptive regression spline (MARS) models were fitted using these above features to predict estimated retinal ganglion cell (eRGC) counts, a measure of glaucoma disease severity derived from standard automated perimetry and optical coherence tomography imaging.

Results

Predictive models using features from the time-frequency analyses – using both the DWT and MP – combined with amplitude markers outperformed predictive models using the markers alone with linear (P = 0.001) and MARS (P ≤ 0.011) models. For example, the proportion of variance explained (R2) by the optimal MARS model – using MP features and amplitude markers – was 0.63, compared to 0.34 for the model using the markers alone (P < 0.001).

Conclusions

Novel time-frequency features extracted from the photopic ERG substantially added to the prediction of glaucoma severity compared to using the time-domain amplitude markers alone.

4.2 Introduction

Glaucoma is a major cause of irreversible vision loss throughout the world. It is a neuropathy of the optic nerve (Weinreb, Aung et al. 2014) with progressive loss of fibres of the nerve and cell bodies of the retinal ganglion cells (RGCs). Intraocular pressure (IOP) remains the only modifiable risk factor in glaucoma, and the reduction of IOP via topical medical therapy or surgical intervention remains the main approaches for slowing the progression of the disease. Advanced glaucoma has characteristic changes of visual field and optic nerve structure, but the early changes with glaucoma can be difficult to detect. It has been estimated that at least 25% to 33% of RGCs must be lost before producing significant visual field abnormalities (Kerrigan–Baumrind, Quigley et al. 2000).

Prior to ganglion cell death, RGCs have been shown to exhibit structural and functional changes (Fry, Fahy et al. 2018). Optic nerve crush (Kalesnykas, Oglesby et al. 2012) and acute IOP elevation (Della Santina, Inman et al. 2013) studies in the mouse showed structural changes that include a reduction in the dendritic arbor area, the length of dendrites and the number of dendrites, which are correlated with the severity of the disease. Early functional changes that occur in sick RGCs include an increased excitability (Tao, Sabharwal et al. 2020), which can manifest as an increased basal or stimulated firing rate (Risner, Pasini et al. 2018)}. This may be caused by a depolarization of the resting membrane potential reducing the threshold for excitation (Calkins 2021). In primate eyes with experimental glaucoma, the RGCs become less responsive with the mean and peak spike rates falling (Weber and Harman 2005). With continued stress, the ultimate outcome for the stressed RGCs is apoptosis (Nickells 1999). Capturing early functional changes of the RGCs could thus aid in predicting glaucoma progression and provide complementary information in the early diagnosis of this condition.

The electroretinogram (ERG) (McCulloch, Marmor et al. 2015) is an electrical response measured at the cornea from a photic stimulus of the eye, and it is a direct measure of retinal function that could be used to capture early RGC dysfunction. Techniques described to measure such function include the pattern ERG (PERG), the positive scotopic threshold response (pSTR) and the Photopic Negative Response (PhNR) (Porciatti 2015). By following a cohort of people considered glaucoma suspects, Banitt (Banitt, Ventura et al. 2013) showed that changes in the PERG preceded structural changes as measured on the optical coherence tomography (OCT) by a number of years. Working in a rat model, Liu (Liu, He et al. 2017) showed that with chronic IOP elevation, pSTR dropped by 25% with no significant change in RGC density.

The photopic negative response (PhNR) of the ERG is a slow negative potential following the b-wave. First described by Viswanathan et al (Viswanathan, Frishman et al. 1999) in 1999, it has been shown to be reduced in glaucoma, both in clinical studies and in animal models (Viswanathan, Frishman et al. 2001) (Machida, Gotoh et al. 2008, North, Jones et al. 2010, Nakamura, Hangai et al. 2011, Preiser, Lagreze et al. 2013, Machida, Kaneko et al. 2015). It arises from the spiking potentials of retinal ganglion cells. It is elicited under photopic conditions by a red flash of light on a blue background and recorded from a corneal electrode relative to the lateral canthus of the eye. Previous studies have shown that there are significant correlations between the PhNR amplitude and the mean deviation (MD) of standard automated perimetry (SAP) (Kita, Holló et al. 2020) and peripapillary nerve fiber layer thickness (Cvenkel, Sustar et al. 2017). Although the PhNR can be associated with other systemic illnesses such as diabetes and hypertension (Chen, Zhang et al. 2008, Banerjee, Pandurangan et al. 2021), the PhNR may thus still be a useful measure for the objective assessment of RGC function in glaucoma.

Of note, the slow negative potential after the b-wave (the PhNR) is interrupted by a positive peak termed the i-wave of the ERG (Rosolen, Rigaudière et al. 2004). This waveform has a variable size and is poorly understood, but it is conserved in primate and non-primate species and thought to arise from cells distal to the RGCs and from the off pathway (Rangaswamy, Frishman et al. 2004). The presence of the i-wave can lead to different interpretations of the PhNR wave and therefore can affect the measurement of PhNR amplitude and latency. Recently, the PhNR has been subdivided into: (i) PhNR1, the trough after the b-wave and before the i-wave, and (ii) PhNR2, the first trough after the i-wave (Ortiz, Drucker et al. 2020) (with the latter being typically measured as the PhNR). There has been little published on the significance of the i-wave itself, although authors have noted that it increases in prominence or amplitude in the presence of RGC pathology (Rangaswamy, Frishman et al. 2004). Capturing these additional parameters of the ERG beyond the conventional PhNR measure could also help better capture RGC dysfunction.

In addition, the stimulus parameters for optimizing the PhNR for the characterization of RGC dysfunction have been studied. Sustar (Sustar, Cvenkel et al. 2009) showed that the PhNR elicited with a chromatic stimulus was better at discriminating between glaucomatous and non-glaucomatous eyes than an achromatic stimulus. Hara (Hara, Machida et al. 2020) reported that the correlations between the PhNR1 and the mean deviation were stronger for the achromatic stimulus but the converse was true for the PhNR2, where the correlations were stronger for the achromatic stimulus. It is likely that these observations relate to the changes in the i-wave between the ERGs elicited by chromatic and achromatic stimuli, which they reported as occurring earlier in the achromatic compared to chromatic stimulus, and thus had less of an impact on the PhNR I amplitude. Nonetheless, they reported that the PhNR2 of the chromatic stimulus performed the best at discriminating between eyes with and without glaucoma. However, neither of these two studies have compared the utility of these three parameters – the PhNR1, i-wave and PhNR2 – when used in combination (rather than in isolation) between the chromatic and achromatic stimuli. It is plausible that the utility of the achromatic stimulus may improve when these three parameters are considered in combination, due to its ability to more clearly demarcate the PhNR1 and i-wave.

We thus sought to examine whether the use of a fuller set of parameters (a-wave, b-wave, i-wave, PhNR1 and PhNR2) elicited using both chromatic and achromatic stimuli could better capture RGC functional loss compared with the use of the PhNR alone. This was performed by evaluating its predictive performance for a combined structural and functional measure of disease severity in glaucoma by Meideiros and colleagues (Medeiros, Lisboa et al. 2012), termed the estimated RGC (eRGC) count. Incorporating all markers into a predictive model allows for normalization of overall gain via the a- and b-wave amplitudes and the effect of the i-wave on the estimation of the PhNR via the use of the i-wave itself and the PhNR1 and PhNR2. In this study, we examined both a conventional linear predictive model and a Multivariate Adaptive Regression Splines (MARS) model. The MARS approach models the outcome variable as a linear sum of piecewise linear functions (truncated functions with knots) (Friedman 1991). As such, it can be used to model the outcome measure as a linear sum of functions that can saturate or change their significance depending on amplitude, and allow intuition about the underlying process without incorporating interaction terms. We thus sought to use MARS as it could better model potentially complex functions for each of the ERG parameters to achieve improved prediction performance and model intuition.

4.3 Staging of Glaucoma

To a clinician managing a patient with suspected or confirmed glaucoma, two critical parameters are the estimate of the extent of glaucoma damage and the estimate of vulnerability to further damage. Measures that have been described or are in common use are listed in Table 4-1.

Table 4-1: Assessment measures for glaucoma. One typical reference given for each method. BMO-MRW = Bruch's membrane opening – minimal rim width, BMO-MRA = Bruch's membrane opening – minimal rim area, PERG = pattern electroretinogram, PhNR = photopic negative response. GDX-VCC = GCC variable corneal compensation

Technology	Sub-measure	Units	Reference
Visual Field			
	Mean deviation	dB	(Ng, Sample et al. 2012)
	Visual Field Index	NA	(Sousa, Biteli et al. 2015)
	Pattern Standard Deviation	dB	(Ng, Sample et al. 2012)
OCT			
	Retinal Nerve Fibre layer thickness	um	(Sihota, Sony et al. 2006)
	Ganglion Cell layer thickness	um	(Elbendary, Abd El-latef et al. 2017)
	BMO MRW	um	(Chauhan, Danthurebandara et al. 2015)
	BMO-MRA	mm	(Gardiner, Ren et al. 2014)
Micro perimetry			
	Sensitivity	dB	(Molina-Martín, Pérez-Cambrodí et al. 2018)
ERG			
	PhNR	uV	(Viswanathan, Frishman et al. 1999)
Pattern ERG			
	Perg Ratio	NA	(Jung, Jeon et al. 2020)
Scanning Laser Polarimetry			
	GDX VCC	um	(Wang, Li et al. 2011)

OCT= optical coherence tomography, BMO-MRW = Bruch's membrane opening maximum rim width, BMO-MRA= Bruch's membrane opening maximum rim area, PhNR = photopic negative response, PERG= pattern electroretinogram, ERG=electroretinogram

Most of these measures can be used in a cross-sectional way to formulate an opinion about the current state of glaucomatous vision loss, or they can be used to monitor progression over time. In trying to synthesize the significance of these various test results, the clinician faces the challenge that the results have different units, some are on log and some on linear scales and that the relationship between these measures and underlying disease severity is often not linear – or at least not linear over the entire range of the measures.

Most machine learning models also aggregate multiple features together to produce a scalar predictor. That is, the numerical or categorical variables listed in the units column of Table 4-1 are processed to produce a scalar output such as 'disease severity' or 'progression probability' or even a binary predictor such as 1 for glaucoma and 0 for non glaucoma.

One way of conceptualizing this sort of scalar output is to produce an estimate of the count of ganglion cells. The ultimate pathological impact of glaucoma is manifested by the loss through apoptosis of ganglion cells and it is attractive and intuitive to think of the disease severity in terms of the percentage or number of cells lost. Substantive work was done Medeiros and others. (Medeiros, Lisboa et al. 2012) who, working from an earlier model of glaucoma in primates (Harwerth, Wheat et al. 2010), developed a combined measure of loss of function in glaucoma. They combined features derived from both the visual field (by standard automated perimetry) and the optic nerve (through the measurement of the thickness of the peripapillary retinal nerve fibre layer). Details of the methodology used in this study are given below.

Any measure of glaucoma that resolves to a single scalar quantity is simplistic and imperfect. Not all ganglion cells are of equal importance. Loss of a small number of ganglion cells in an area close to fixation will be devastating. Loss of a much larger number of ganglion cells dispersed throughout the peripheral field may be barely noticed. Even so, the measure is intuitive, validated and mathematically tractable and has been used for comparison in this thesis.

4.4 The frequency domain

The notion that a function of a variable can be expanded into a series of cosines of multiples of the variable is credited to and named after work of Jean-Baptiste Fourier, who published his analytic theory of heat (Fourier 1822). In it he *showed "exemples de l'usage des series trigonometrique dans la theorie de la chaleur "* translated as "examples of the usage of trigonometric series in the theory of heat." He showed the heat to be the infinite sum of a series of cosines: in the first example of his work

$$1 = a \, cos \, y + b \, cos \, 3y + c \, cos \, 5y + d \, cos \, 7 \, y \qquad 4\text{-}1$$

So that the function is modelled by an equation representing multiples of the frequency – harmonics and a coefficient for each harmonic. The decomposition became known as a Fourier series and represented the decomposition of a periodic function. In fact, the use of trigonometric series in this way was actually described earlier by Euler (Euler 1748). The principle was later applied to discrete time and non periodic functions as the Discrete Fourier Transform (DFT).

For a sampled signal represented as a time series, the DFT is given as

$$X_k = \sum_{n=0}^{N-1} x_n e^{-kn\frac{2\pi i}{N}} \qquad 4\text{-}2$$

Where $\{x_n\} = (x_0, x_1, .. x_{N-1})$ is the original time series and $\{X_k\} = (X_1, X_2, \ldots X_{N-1})$ is the DFT. X_k is a complex sequence encoding amplitude and phase for each frequency component. X_0 is the Direct current (DC) or mean of the time series. The remaining terms are repeated so the series is symmetrical about its mid point and contains N/2 frequency components. The reason for the symmetry is that the DFT contains both positive and negative frequency components. The highest frequency within the sample represents the frequency oscillating between maximum and minimum every two points. The lowest frequency resolved is the frequency corresponding to one full period of a sine wave from the start to the end of the time series.

When a sample, such as an ERG, is analysed in this way, it is said to be transformed to the *frequency domain*. The transformed function is now a function of amplitude vs frequency rather than amplitude vs time (which analogously is known as the *time domain*). A real valued time domain signal will be transformed into a complex valued frequency domain signal. The phase of each component is represented by its phase angle on the complex plane. The modulus of the frequency domain information squared is devoid of phase information and is known as the *power spectrum* of the signal. Although the complex valued frequency domain signal can be inverted via the inverse Fourier transform, the original time domain signal can not be reconstructed from the power spectrum alone. Inferences can be made about the amounts of oscillating components within the signal from its DFT but care must be taken. Where a frequency is in the signal but its period does not correspond to an exact multiple of the sample rate, the amplitude in the DFT is spread over a number of frequencies. In addition, for short signals, the imposition of a rectangular window creates distortions in the frequency domain – an effect that can be mitigated to some extent by the use of a window such as the Hamming window (Blackman and Tukey 1958) named after Richard Hamming who worked, amongst other things, on the Manhattan project. Windowing is not needed for periodic signals such as the steady state pattern ERG (Bach and Meigen 1999).

The fast Fourier transform (FFT) algorithm was published in 1965 (Cooley and Tukey 1965). It was a rediscovery of a method described by Gauss the century before. In the so-called Radix 2 method, the series is first split into two halves which are then further split and so on. This means that for this particular algorithm, the sequence must be of length 2^n. Other algorithms exist for samples not of this length. The FFT is often (but not completely correctly) used synonymously for the DFT.

The DFT has been used to analyse the electroretinogram by many authors since the FFT was described. A PubMed search revealed 146 results in September of 2018 when searching for the terms “Fourier Transform” and “Electroretinogram”.

Analysis can be performed on transient signals (such as the flash ERG) or on steady state signals such as the 30Hz flicker response. In the ERG for glaucoma, the DFT has found most utility in the analysis of the steady state pattern erg (PERG). Porciatti for example (Porciatti and Ventura 2009) used the DFT to extract the response of the eye at the fundamental frequency of the stimulus – in this case 16.28Hz. Salgrello (Salgarello, Falsini et al. 2008) found an improvement in the second harmonic of the steady state PERG in patients who had intraocular pressure improvement with drops.

4.5 The time-frequency domain

Although transformation to the frequency domain can yield great utility in signal analysis, one problem is that the location of the frequency components in time within the signal is not extracted. The analysis is best suited for signals which do not have frequencies which vary in intensity as a function of time and for signals that are stationary. In visual electrophysiology, signals such as the steady state visual evoked potential (Morgan, Hansen et al. 1996), the 30Hz flicker ERG (Wu, Burns et al. 1995) and the steady state pattern erg (Bach, Unsoeld et al. 2006) are all suitable for analysis in the frequency domain.

Where the signal is nonstationary or when it is important to know when the frequency components are happening within the signal, analysis with the DFT is not sufficient. Time frequency analysis is a class of methods where both features of time and frequency are extracted. Inevitably, there is a trade off in resolution between the two feature axes. This general type of analysis is said to transform the time-domain signal into the time-frequency domain. There is no single definition of the time-frequency domain and it will depend upon the transformation used. It can be seen that for a one-dimensional signal (represented as a vector) such as an ERG, this sort of transformation will be into a two dimensional matrix. Time-frequency analysis of the ERG was the topic of a recent PhD thesis (Gauvin 2017).

Time-frequency analyses can be classified (Sejdić, Djurović et al. 2009). In the first category, analysis is done by translating, modulating and scaling a basis function over the signal. A requirement is that the basis functions (also known as atoms) must be square-integrable – that is having finite energy. Short-time Fourier Transform, Wavelets, S-transform and Matching Pursuits are examples of this type of analysis. The second type of analysis, named Time-frequency distributions by Cohen (Cohen 1989) transforms the time domain signal with a kernel function. The Wigner distribution is an example of this method. Different kernels give different time-frequency distributions for the same signal. All analyses of this type can be considered to be the Fourier transform of the instantaneous autocorrelation function of the signal (Hlawatsch and Auger 2013).

4.5.1 Short Time Fourier Transform

The short time Fourier Transform (STFT) is the most intuitive of the time-frequency analyses. It involves sliding a window over the signal and computing the DFT for each window. Analogous to the power spectrum, the spectrogram is the squared modulus of the STFT. Mathematically, the STFT is given as equation 4-3

$$STFT_x^h(t,f) = \int_{-\infty}^{\infty} x\,(u)h(u-t)e^{-i2\pi fu}du \qquad 4\text{-}3$$

Where x(t) is the original time series, f is the frequency and h(t) is the window function. The temporal and frequency resolution of the analysis are both linked to the window length by the Heisenberg-Gabor uncertainty principle (Gabor 1946). In this analysis, the window length is fixed and does not adapt to the actual signal. As with many other time-frequency analyses, the STFT increases the dimensionality of the signal and therefore requires feature selection techniques to find meaningful information within the result. The technique has not been described in relation to the electroretinogram but was used in combination with its inverse (Yanti, Yusoff et al. 2011) as a filtering technique on the visual evoked potential. Attempts have been made to use the measure to diagnose autism (Sheikhani, Behnam et al. 2007) and stress (Sultana, Rahman et al. 2020) from the electroencephalogram although these works were pilot studies of small numbers.

4.5.2 The Matching Pursuit Algorithm

The Fourier transform, the Wavelet transform and others decompose a signal into a linear sum of orthonormal basis functions. Greedy algorithms (Tropp 2004) on the other hand work in an iterative manner finding the first matching function from a dictionary and then the next best one on the residual and so on. They are highly likely to find local rather than global minima in terms of the best-fit decomposition as measured by the least square sum of the residual. Matching Pursuit (Mallat and Zhang 1993) builds a solution one term at a time by selecting from the dictionary the function (atom) that correlates most closely with the residual signal. The STFT has atoms of a constant scale and performs well when the signal structures are localized over the same scale. However it performs poorly when the structures are much smaller or larger than the window size. The wavelet transform can deal with that situation, however may have poor temporal localization of the structures. The matching pursuit algorithm is designed to address the problems of STFT and wavelets. The first step with the algorithm is the creation of the highly redundant dictionary. The dictionary is generated by scaling and shifting a basis function in the same manner as continuous wavelets and a number of families have been described including atoms sourced from discrete wavelets(Chen and Zhang 2014), Gaussians (Ghofrani, McLernon et al. 2003), Gabors (Franaszczuk, Bergey et al. 1998) cosines, polynomials and others. The atom best matching the signal is then selected and that component removed from the signal and the algorithm proceeds to the next step. The full decomposition can be represented as a tile plot of the importance of the atoms (Wang, Guo et al. 2001). Further developments of the technique include the Orthogonal Matching Pursuit (OMP) (Tropp and Gilbert 2007), regularized orthogonal matching pursuit (ROMP) (Needell and Vershynin 2010), stagewise orthogonal matching pursuit (StOMP) (Donoho, Tsaig et al. 2012) and compressive sampling matching pursuit (CoSaMP) (Needell and Tropp 2009).

The dictionary for the matching pursuit algorithm is constructed by shifting and scaling the atoms. One or more waveforms can be used as a basis. Each atom is the same length as the signal to be processed.

The overall aim is to decompose the signal into a linear combination of the basis functions. The first atom ϕ_0 is selected by maximising the absolute value of the inner product

$$< f_0(x), \phi_0 > \tag{4-4}$$

Leaving the function as

$$f(x) =< f(x),\ \phi_0 > \ +\ R^1 f(x) \tag{4-5}$$

And then the process is repeated to find ϕ_1 and R^2 and so on.

The process is continued for either a fixed number of iterations or until a stopping criterion is reached. The final decomposition is represented as equation 4-6

$$f(x) = \sum_{n-1}^{K-1} < R^n f(x), \phi_n > + R^K f(x) \tag{4-6}$$

Some work has been done on finding the optimum dictionary for the algorithm. The dictionary can be optimized both by the selection of the generating family and by the size of the dictionary. A larger dictionary could decompose the signal with fewer coefficients but because both the coefficient and the index into the dictionary must be stored, a larger dictionary may not yield better compression as more bits are required to store the index data (Liu, Wang et al. 2004). Working with non-noisy speech signals, Sturm (Sturm and Gibson 2006) found that dictionaries based on a combination of Gabor and FOF ("Fonction d'onde Formantique") performed better than Gabor atoms or complex exponentials alone. Goodwin (Goodwin and Vetterli 1999) found dampened sinusoids superior to Gabor functions for music.

Working with the multifocal ERG in a primate model of glaucoma, Zhou and colleagues (Zhou, Rangaswamy et al. 2007) found changes in the signal in the area of the oscillatory potentials that were detected using Gabor atoms from the MP algorithm. These results were shown in both experimental glaucoma and in eyes injected with tetrodotoxin. These results mirror those of the photopic negative response and lend weight to the hypothesis that the MP algorithm may be effective in extracting features relevant to glaucoma.

4.5.3 The Cohen Class Quadratic Time Frequency Transforms

The energy of a signal can be deduced from the squared modulus of the signal or its Fourier transform:

$$E_S = \int |s(t)|^2 dt = \int |S(\omega)|^2 d\omega \qquad \text{4-7}$$

Where E_s is the energy for signal s, S is the Fourier transform of s, s(t) is the time series representation of the signal at time t and S(ω) is the frequency domain representation of the signal at frequency ω.

The energy density can be considered like a joint probability distribution : $\rho_s(t,\omega)$ such that

$$E_S = \int\int \rho_S(t,\omega) dt\, d\omega \qquad \text{4-8}$$

And like a joint probability distribution, the energy density can be described as having marginal conditions – the energy for a given time integrated along all of the frequencies and vice versa.

Cohen (Cohen 1989) showed all distributions which were covariant with constant time-frequency shifts and met marginal time and frequency conditions shown in equations 4-9and 4-10could be represented in a generalized way.as 4-11

$$\int \rho_S\,(t,\omega)\; dt \;=\; |S(\omega)|^2 \qquad \text{4-9}$$

$$\int \rho_S(t,\omega) \;=\; |s(t)|^2 \qquad \text{4-10}$$

$$E(t,\omega) = \frac{1}{4\pi^2}\int\int\int e^{-i\theta t - i\tau\omega + i\theta u}\phi(\theta,\tau) s * \left(u - \frac{1}{2}\tau\right) s\left(u + \frac{1}{2}\tau\right) du\, d\tau\, d\theta \qquad \text{4-11}$$

Where $\phi(\theta,\tau)$ is known as the kernel. The Wigner-Ville transformation is a special case of 4-11 where $\phi(\theta,\tau) = 1$. This is shown in 4-12

$$W_S(t,\omega) = \frac{1}{2\pi}\int s(t+\tau/2) s * (t-\tau/2) e^{-i\omega\tau} d\tau \qquad \text{4-12}$$

Another way to consider this is to take the instantaneous autocorrelation as

$$R_S S(t,\tau) = s(t+\tau/2) s * (t-\tau/2) \qquad \text{4-13}$$

And then take the Fourier transform. All of the Cohen Class of analysis return the signal energy and are therefore considered to be quadratic functions.

The Wigner Ville distribution gives excellent resolution in time and frequency if the integral is taken to infinity and yields a real valued function. Because of the quadratic nature of the analysis, if there are multiple frequencies in the signal, the cross product or *interference* terms can become troublesome and difficult to interpret. When considered over a finite time, the WVD becomes the Pseudo Wigner Ville distribution (PWVD) and the ϕ term becomes the window function h(τ).and can be rectangular, Gaussian, Hanning etc. This type of analysis – extracting the power relating to specific frequency bands in time – has been applied to seizure detection in the EEG (Li, Barma et al. 2019) (Guerrero-Mosquera, Trigueros et al. 2010)

4.5.4 The ZAM Transform

Zhao (Zhao, Atlas et al. 1990) described a kernel in which cross terms were minimized, the kernel had finite time support and spectral peaks were enhanced. Like the Wigner-Ville and pseudo Wigner-Ville analyses, non negativity of the spectrum was not preserved which reduced the intuition from the analysis.

The kernel function – named by the authors as the cone shaped distribution – was based on the visual principle of lateral inhibition. The cone shape referred to the shape of the kernel when considered in a three dimensional plot of amplitude vs time and τ. The authors also took special care to ensure that the kernel had finite time support.

The authors described their time frequency analysis as equation 4-11, with the kernel as

$$\phi(\theta,\tau) = \frac{sin(\pi\theta\tau)}{\pi\theta\tau} e^{(-2\pi\alpha\tau^2)} \qquad 4\text{-}14$$

The ZAM kernel was shown to be the best in resolving time-frequency information in somatosensory evoked potentials.(Hu, Luk et al. 2001)

4.5.5 Continuous Wavelet Transform

The wavelet transform was first described by Morlet and others (Goupillaud, Grossmann et al. 1984) in 1984 in the context of seismic studies for oil and gas exploration. It was a technique that tried to separate the frequency bands without “excessive loss of resolution of the time variable” remaining robust to small perturbations. The original time based signal is convolved with a series of small signals or wavelets that are scaled in width and shifted in time.

The wavelet function itself must be square integrable, have unit energy and be characterized by two parameters specifying position and dilation as 4-15. The wavelet function that generates these shifted and scaled wavelets is knowns as the mother wavelet

$$\phi_{u,s}(t) = \frac{1}{\sqrt{s}}\,\phi\left(\frac{t-u}{s}\right) \qquad 4\text{-}15$$

The wavelet transform itself is then given

$$w_x(u,s) = \int_{-\infty}^{+\infty} x\ (t)\ \phi_{(u,s)}\ (t)\ dt \qquad 4\text{-}16$$

The output of the transform is a real valued two dimensional function with axes of time and scale. In practice, the transform has a discrete realization so that the input is a time series and the output is a matrix of number of columns equal to the time series length and number of rows equal to the number of dilations performed in doing the transform. The values of each cell of the matrix comprise the wavelet coefficients.

A variety of different continuous wavelets have been described and the resolving power of the wavelet in time and frequency depends on the choice of a mother wavelet suited to the nature of the signals to be extracted. In general though, the mother wavelet must comply with the admissibility condition (Sadowsky 1996) which states that a Fourier transform of the wavelet must exist and that the transform has no DC.

Although in theory and in continuous time, there is always an inverse of the CWT, in the discrete time realization of the CWT, the inverse is not always unique and not always a perfect replica of the original function. The CWT has been used to analyse the rat electroretinogram (Forte, Bui et al. 2008), specifically to analyse the oscillatory potentials. It has also been utilized for the VEP (Tallon-Baudry, Bertrand et al. 1996) but the analyses tend to be more qualitative than quantitative.

4.5.6 Discrete Wavelet Transform

The idea of discrete wavelets had its genesis in work by Mallat (Mallat 1989). Daubechies (Daubechies 1988) invented the compactly supported orthonormal wavelets which is the basis for the modern discrete wavelet transform. The key innovation in Daubechies work was to eliminate the redundancy in the transform. This transform is quite different to the continuous wavelet version. For the DWT, the chosen wavelet has two functions – a wavelet function of integral zero which is a low pass filter and the matching scaling function with integral equal to 1 which is the high pass filter. The two filters form a quadrature mirror. The first step is to apply both of these filters to the sample and then to downsample each resultant by 2. The output of the low pass filter is then used as the input to the next step and the process is repeated. This results in the output as a tree and the total number of coefficients is equal to the number of samples in the input which necessarily must of length equal to a power of 2. Given certain conditions, the transform is invertible and the original signal can be exactly recovered.

A large number of wavelets with different properties have been generated after the work of Daubechies. Some care is needed in choosing the wavelet family or "Mother wavelet"; it is generally chosen with a form similar to the waveform structure of interest that is contained within the signal. For example, if *a priori*, the glaucoma signal is thought to be a Gaussian like wave, then the best wavelet might be one that has a similar structure. One of the constraints by the condition of orthogonality is that the scaling function cannot be symmetrical. The least asymmetrical wavelet family was designed to maximize symmetry but some asymmetry does persist. Further discussion of wavelet selection can be found below.

The biorthogonal wavelets (Cohen, Daubechies et al. 1992) are invertible but not orthogonal. The transform uses different wavelets for decomposition and reconstruction. Symmetry is achieved in the scaling function but at the expense of complexity and the introduction of redundancy. Linear phase is achieved as a result of the symmetry. There are no published studies comparing the biorthogonal and orthogonal wavelets for the electroretinogram, although some studies have noted that the biorthogonal family performs better for image compression (Rout and Bell 2004).

Multiwavelets (Downie and Silverman 1998) use a set of wavelets where the wavelet is different for each scale. This has the theoretical advantage of optimising the wavelet for the particular waveforms of interest at each scale but greatly adds to the complexity of the transformation and its analysis. (Guo, Rivero et al. 2010) . In the multiwavelet case, instead of a vector, the filter consists of a square matrix (Strela 1996). For each application of the filter, matrix multiplication is required; the signal consisting of a matrix with the number of rows equal to the number of rows and columns of the filter. To convert the signal from a vector to a matrix, preprocessing is required. The simplest preprocessing would be to duplicate the rows but this does introduce oversampling. Other techniques of preprocessing have been described for example (Xia, Geronimo et al. 1996) which have reduced the redundancy in the process.

The wavelet transformations are designed to be invertible. That is, for a wavelet transform, there should exist a corresponding inverse wavelet transform that recovers the original signal. Invertibility is important if the transform is to be used for data compression or for filtering. For example, the signal can be transformed by the wavelet transform, then processed within the wavelet domain and then resynthesized back to the original time domain format. This is what is done for wavelet denoising and filtering where the magnitude of coefficients is thresholded or specific coefficients reduced or discarded. If the constraint of invertibility is relaxed, new possibilities open up. One technique is the local wavelet maximum (Gauvin, Little et al. 2015). In this technique, the discrete wavelet transform is performed on the signal and is repeated with the signal being shifted each direction over a window. The wavelet coefficients are chosen as the maximum of the coefficients over the timeshifted range. This has the advantage of better intuition about the meaning of individual coefficients but does so at the expense of the loss of orthogonality, invertibility and conservation of energy.

An alternative approach to the local wavelet maximum approach is to ensure that the timing of waveforms is consistent is to reference the timing to the b-wave peak rather than the stimulus onset. This is the approach taken in this work and has the property that the DWT transforms remain invertible.

4.5.7 Mother wavelet selection

A number of techniques have been described for the selection of the optimum basis for the DWT. Coifman (Coifman and Wickerhauser 1992) proposed the selection of the wavelet family by minimizing the informational cost. The algorithm uses Shannon entropy to compare wavelet families. Extending this via the ratio of energy to Shannon entropy, the algorithm has been used for mother wavelet selection in the fields of wind energy (Strömbergsson, Marklund et al. 2019), roller bearing analysis and in choosing the particular decomposition level for wavelet denoising (Deák, Mankovits et al. 2017).

4.5.8 Time Frequency Feature extraction

The signal transformed into the time-frequency domain poses a problem for feature extraction. The signal – originally a time series - has been transformed into a two dimensional matrix with time and frequency encoded by the position and intensity of the coefficient encoded by the content of each cell. In the case of the continuous wavelet, discrete wavelet packet and quadratic time frequency analysis, the dimensionality is necessarily increased. In the case of the discrete wavelet transform (DWT), the dimensionality is preserved. Matching pursuit gives an opportunity to decrease the dimensionality by using a smaller, custom dictionary. The discrete wavelet and discrete wavelet packet transform are further complicated by the fact that the output is a tree structure and not a matrix.

Intuition can be difficult. Although the amplitude and location (in time and frequency) can in principle be performed, lack of consistency between individual decompositions can make this process fraught.

Decompositions into matrices can be thought of as images. An example of a decomposition -in this case the continuous wavelet transform using a Mexican Hat wavelet is shown in Figure 4-1. The top plot show the time series, the bottom plot is a heatmap plot of the absolute values of the coefficients of the output of the transformation. The heatmap shows the time axis aligned with the original time axis of the ERG signal. The vertical axis shows the period of the relevant wavelet. Higher frequency components (smaller periods) are shown lower on the plot (closer to the origin)

The times corresponding to the fastest changes within the signal contain the largest proportion of high frequency components – the fastest rate of change is the upswing and downswing of the b-wave.

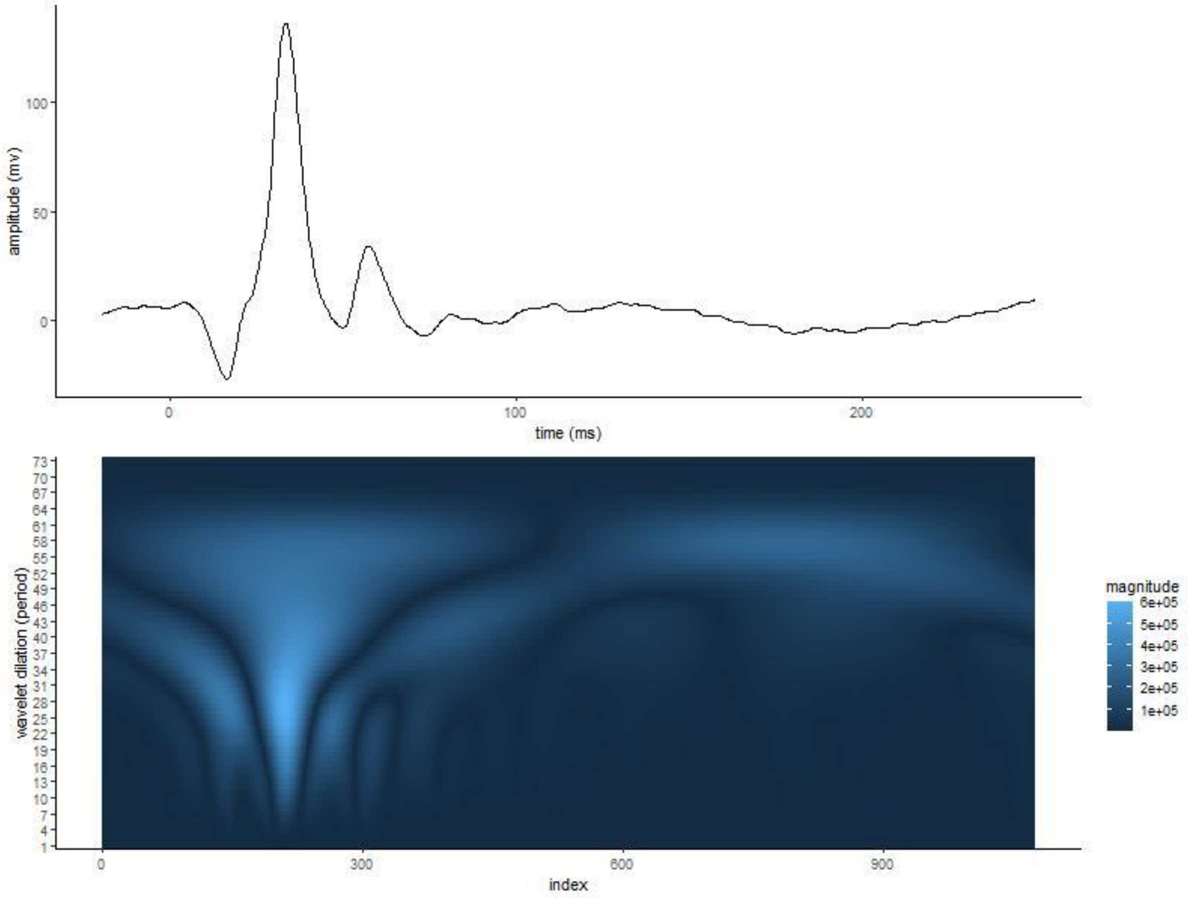


Figure 4-1: Sample ERG with heatmap of Continuous Wavelet Transform

Consideration of the output matrix as an image yields the possibility of using various feature extraction techniques developed for images. Running a technique across the ensemble of traces could enable identification of a small number of features suitable for utilisation in a predictive model. One such technique is the k-means clustering technique (Jain, Murty et al. 1999). This technique finds clusters of pixels based on a histogram of the pixels in the image. The problem with the technique in this case is that it does not take account of the position of the pixels. Clearly, clusters of matrix cells which will be the inputs to an aggregating function need to be adjoining and the position is of high importance.

An approach to incorporating intensity and locality to extract superpixels is the Simple Linear Iterative Clustering method (SLIC) (Achanta, Shaji et al. 2012). In this technique, the intensity and the location are used to compute a Euclidian distance. As the dimensions of distance and intensity are not the same, a compactness factor is specified to allow the algorithm to utilize both parameters. The algorithm can be done in a colour space or greyscale. Normalization is required to appropriately weight distances in the location and intensity direction.

Figure 4-2 compares the k-means and SLIC algorithms in segmenting a fundus image from one of the participants of the study. The top left pane shows the colour fundus image. The middle pane shows the results of k-means segmentation extracting 20 clusters. The black lines show the cluster boundaries. As the clustering decision is based only upon colour, the boundaries are extremely complex and irregular. The bottom middle panel shows the boundaries for the 13th cluster. The SLIC decomposition is shown on the top left and has much more simple boundaries. The 13th superpixel is shown in the lower right panel. Applied to a 2 dimensional matrix such as the time-frequency domain, the algorithm will extract a determined number of clusters of similarity of location and intensity. Unlike an actual image, the time-frequency domain representation has dissimilar axes in the x and y direction (time vs frequency).

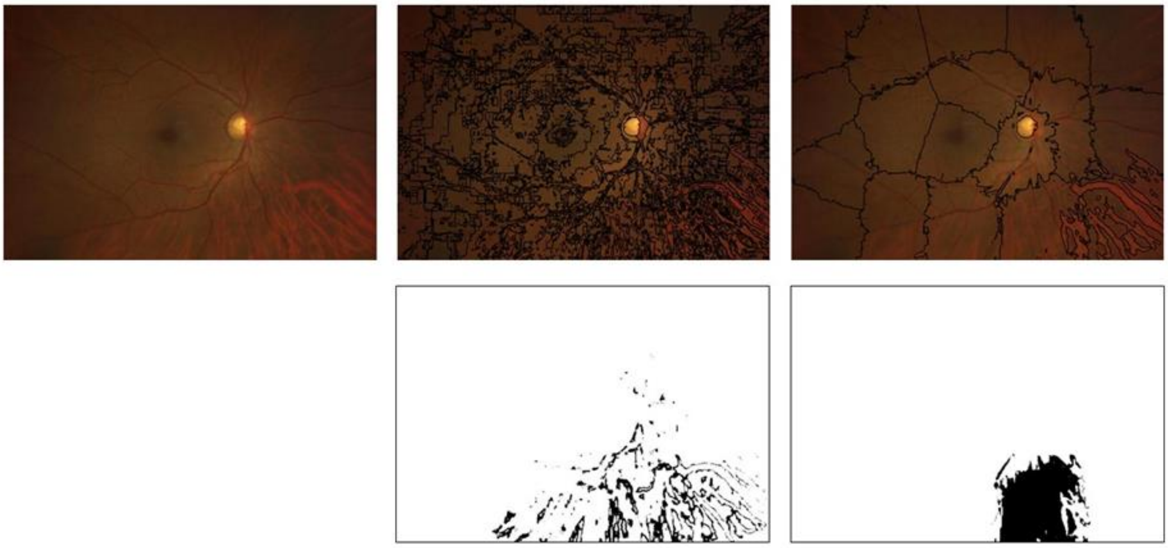


Figure 4-2: k-means vs superpixel decomposition of fundus. The top left image is a fundus photograph of one of the participants. The top middle image is a k-means clustering of 20 clusters. The black lines are the cluster boundaries. The top right image is a SLIC superpixel decomposition with 20 superpixels and a compactness of 0.3. The lower images show the 13th cluster and superpixel for the corresponding decomposition in the image above.

4.5.9 Texture Analysis

Texture analysis is another technique designed for image processing that can be used on the time-frequency analysis. A popular technique is the Grey Level Co-occurrence Matrix (GLCM) approach (Haralick, Shanmugam et al. 1973). Over a specified number of grey levels G and over a square window, a probability matrix is built for each combination of grey levels in adjacent pixels in a specified direction or the mean of a set of directions. For example, with 6 levels and a direction 0 radians, a probability matrix is made for each possible combination of adjacent pixels left to right. I.e. (0,0), (0,1), (0,2) …. (4,5),(5,5). This probability matrix $P(i,j) : i,j \in \{0,1,2,3,4,5\}$ is assigned to the pixel corresponding to the middle of the window. From each probability matrix, a number of textural features can be extracted (Xie, Shao et al. 2015) of which 8 have been used in this work:

$$Mean = \sum_{i,j=1}^{G} iP(i,j) \qquad 4\text{-}17$$

Mean stands for the average grey level of all pixels in the GLCM

$$Variance = \sum_{i=1}^{G}\sum_{j=1}^{G}(i - u^2)P(i,j) \qquad 4\text{-}18$$

Variance indicates the rate of pixel’s value changes relative to the mean u

$$Homogeneity = \sum_{i=1}^{G}\sum_{j=1}^{G}\frac{P(i,j)}{1 + (i - j)^2} \qquad 4\text{-}19$$

Homogeneity indicates the uniformity – differences are weighted by exponential decrease with increasing distance

$$Contrast = \sum_{i=1}^{G}\sum_{j=1}^{G}(i - j)^2P(i,j) \qquad 4\text{-}20$$

The contrast measures the degree of local variations. The weight assigned to each probability increases with the square of the distance so where for example P(2,2) is given no weight as these pixels in the GLCM are entirely similar.

$$Dissimilarity = \sum_{i=1}^{G}\sum_{j=1}^{G}P(i,j)|i - j| \qquad 4\text{-}21$$

Instead of the weights increasing with the square, the weights of the dissimilarity measure vary linearly with the distance in the GLCM.

$$Entropy = -\sum_{i=1}^{G}\sum_{j=1}^{G}P(i,j)log\,P(i,j) \qquad 4\text{-}22$$

The entropy as per chapter 3 of this thesis measures the information content. More disorder is associated with higher entropy

$$Angular\ Second\ Moment = \sum_{i=1}^{G}\sum_{j=1}^{G}P^2(i,j) \qquad 4\text{-}23$$

The angular second moment (ASM) is the unweighted sum of squares of the GLCM. High values occur when the texture is very orderly. The positive square root of the ASM is known as the Energy. The ASM reflects the size of the texture

$$Correlation = \sum_{i=1}^{G}\sum_{j=1}^{G}\frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i\sigma_j} \qquad 4\text{-}24$$

Where μ_i is the mean of i , μ_j is the mean of j, σ_i is the standard deviation of i and σ_j is the standard deviation of j. It measures the joint probability of the specified pixel pairs and the dependency of levels within the GLCM on those of neighbouring pixels

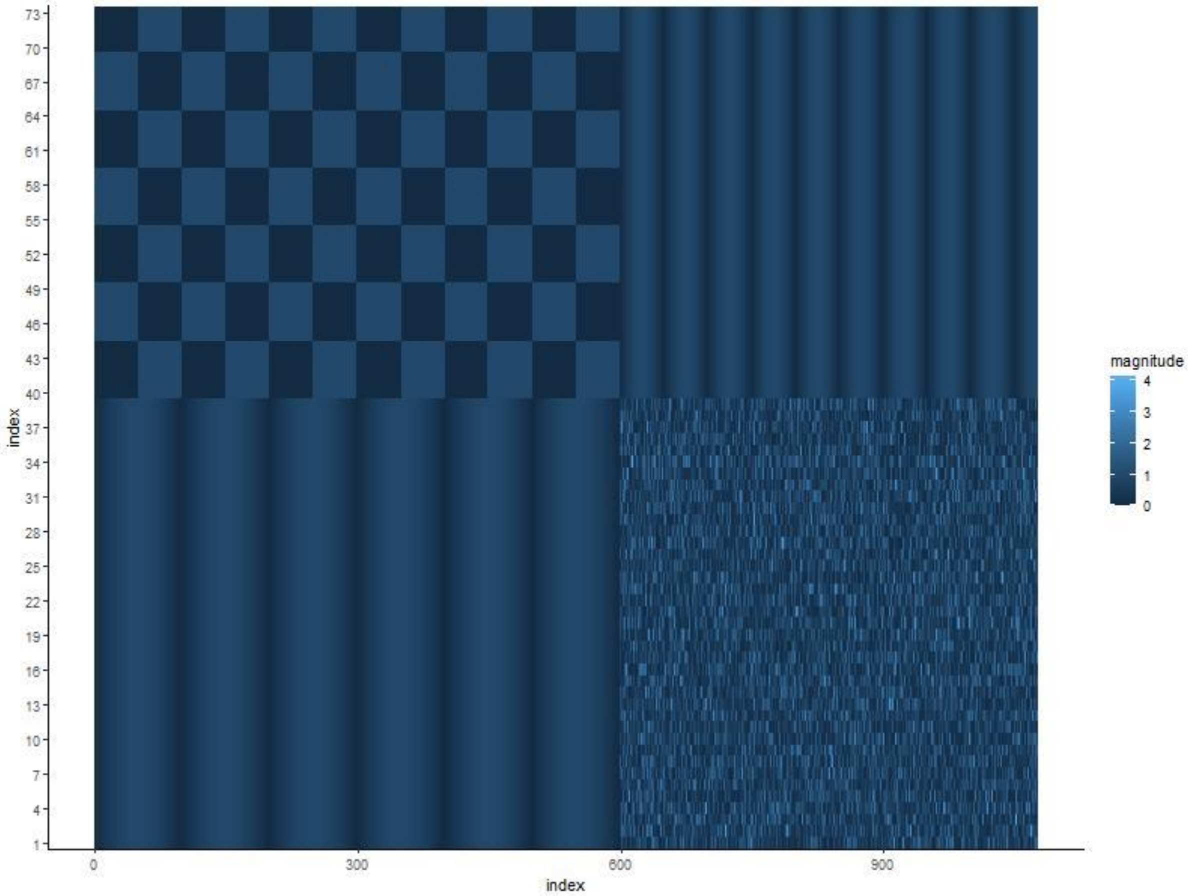


Figure 4-3: Sample texture

Figure 4-3 shows a sample texture. The upper right panel is a checkerboard pattern, the upper left a higher frequency sinusoidal grating, the lower left a lower frequency sinusoidal grating and the lower right a gaussian white noise pattern.

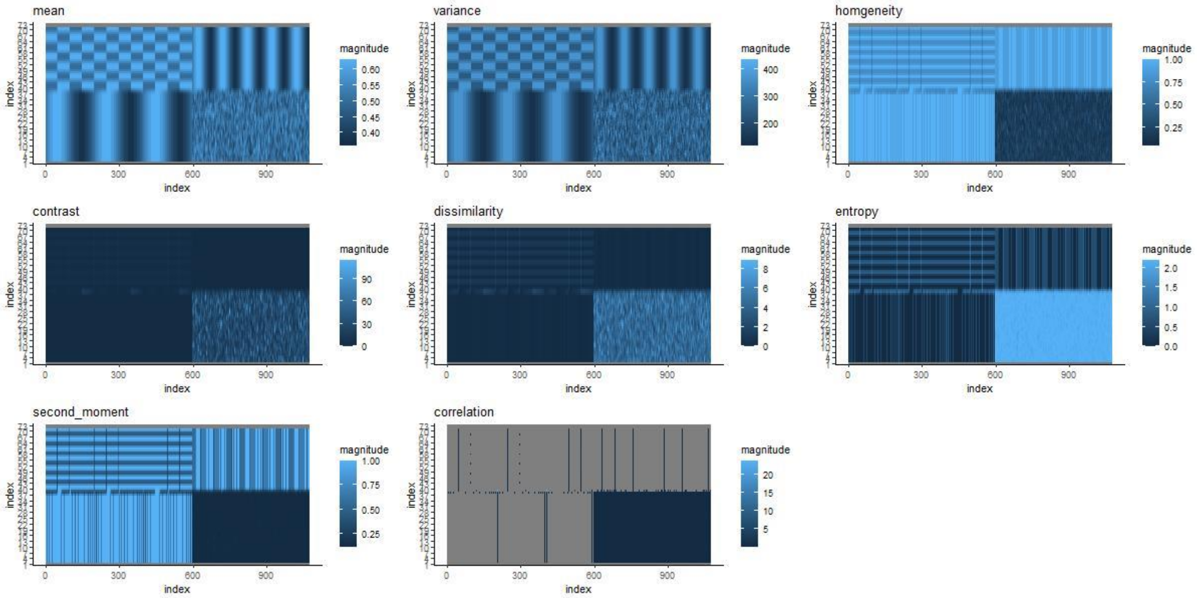


Figure 4-4: GLCM texture analysis of sample texture

Figure 4-4 shows the texture analysis of the sample texture shown in Figure 4-3. The mean is essentially just a moving average and replicates the sample texture but with less resolution. The variance is also very similar to the original. As expected, the homogeneity is greater in the more regular areas and the contrast and dissimilarity is the reverse. The entropy is strikingly greater in the Gaussian white noise patch; recall that a Gaussian distribution maximises entropy. The second moment was very low in the Gaussian patch. The correlation could not be computed for much of the sinusoidal and checkerboard areas.

Using this sort of approach on the Short term Fourier Transform of EEG signals, Şengür was able to detect periods of seizure activity within the electroencephalogram (EEG) (Şengür, Guo et al. 2016). In that work, the GLCM features were used to train a support vector machine (SVM) classifier. A number of different kernels were used including the radial basis function but no intuition about the meaning of the textural features was attempted. Bajaj and colleagues (Bajaj, Rai et al. 2017) extracted texture features from the Wigner Ville distribution (WVD) of the EEG to inform a support vector machine (SVM) classifier to detect seizures. They found the angular second moment of the signal to be the most important texture feature to inform the classifier.

4.6 Methods

The study was an approved study by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital, and it was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent prior to any study procedures being undertaken.

Participants

Participants with primary open angle glaucoma were recruited from a private ophthalmology practice, and the diagnosis of glaucoma was based on a comprehensive clinical assessment by an ophthalmologist based on characteristic optic nerve head appearance, the presence of glaucomatous visual field defects and/or neuroretinal tissue loss on OCT imaging. No participants with ocular or systemic diseases that could affect the optic nerve such (as choroidal neovascular membrane, extensive macular atrophy, diabetic retinopathy, multiple sclerosis or epiretinal membrane) were included and only adult patients over the age of 18 were eligible for inclusion in this study. Both eyes were included where glaucoma was bilateral, and only eyes with an acuity of 20/40 (LogMAR 0.3) or better were included in the study.

Information from the medical record

For the glaucoma cohort, information was sourced from the medical record including the number of medications, duration of glaucoma and parameters were captured from the visual field including mean deviation. All of the glaucoma patients had undergone prior to the test optical coherence tomography scans of macula and optic disc and had optic nerve photos taken.

Standard Automated Perimetry

All participants performed standard automated perimetry (SAP) testing using the 24-2 Swedish Interactive Threshold Algorithm (SITA) Fast protocol on the Humphrey Field Analzyer 3 (Carl Zeiss Meditec Inc.; Dublin, CA), following correction of the spherical refractive error component from subjective refraction. All visual field results had fixation losses and false negative responses of ≤ 33% and false positives of ≤ 20%.

Optical Coherence Tomography Imaging

All participants also underwent optic disc-centered OCT volume scans performed with the Cirrus HD-OCT device (Carl Zeiss Meditec Inc.; Dublin, CA) with dilated pupils. This model has superluminescent diode illumination and obtains 27,000 A-scans per second, with an overall axial resolution of 5 um. A scan centred on the optic disc was obtained consisting of 200x200 A-scans and covering an area of 6x6mm. The circumpapillary retinal nerve fibre layer (RNFL) thickness was derived from this cube from a derived 3.46mm diameter circle scan, consisting of 256 A scans with segmentation performed by the instrument. All scans had a signal strength score ≥7, and only those showing correct centration and accurate segmentation, were analysed in this study.

Estimated Retinal Ganglion Cell Counts (eRGC)

For each eye, an estimate was made of the RGC counts (eRGC) as described by Medeiros and colleagues (Medeiros, Lisboa et al. 2012). This model is based on empirical formulas developed by Harwerth and colleagues (Harwerth, Wheat et al. 2010)

From non-human primates, Harwerth developed a linear model estimating the RCG soma density (somas/mm^2) loss in dB units at location k (gl_k) at an eccentricity (ec) from the SAP visual sensitivity (s; dB) via slope (m) and intercept (b).

$$m_k = 0.054ec_k + 0.91 \qquad 4\text{-}25$$

$$b_k = -1.5ec_k - 14.8 \qquad 4\text{-}26$$

$$gl_k = (s_k - b_k)/m_k \qquad 4\text{-}27$$

Translating the non-human primate work to humans required accounting for the different length of the eye in humans vs. non-human primates and the change in visual field threshold strategy from full threshold to the faster Swedish Interactive Threshold Algorithm (SITA) (Bengtsson and Heijl 1998). For humans using SITA, the equations became

$$m_k = 0.07128ec_k + 0.91 \qquad 4\text{-}28$$

$$b_k = -1.98ec_k - 14.8 \qquad 4\text{-}29$$

$$gl_k = \frac{s_k - 1 - b_k}{m_k} + 4.7 \qquad 4\text{-}30$$

The total ganglion cell number SAPrgc is found as

$$SAPrgc = \sum_{k=1}^{K} 10^{0.1gl_k} \qquad 4\text{-}31$$

Noting that the number of ganglion cell axons entering the optic nerve from an area of the retina must be representative of the number of cell bodies, the authors also developed an estimate of axon numbers from the circumpapillary measurements of RNFL thickness measured by OCT. Recognizing a near linear loss of axons with normal aging and remodelling of the nerve fiber layer, the authors fitted the following formula relating axon density d (axons/µm^2) for a patient of age (ag) in years for a section of the RNFL scan of length px in pixels and average height (mh; µm). A 21.2 pixel length per pixel over a 10.87 mm scan length was used.

$$d = -0.007ag + 1.4 \tag{4-32}$$

And a total axon count a given simply as

$$a = 21.2\, mh\, px\, d \tag{4-33}$$

Again, to account for the effect of remodelling in glaucoma and to express the final estimate in terms of total ganglion cell axons from OCT imaging (OCTrgc), a mean deviation correction *c* is calculated from the MD in dB as

$$c = -0.26\, MD\ + 0.12 \tag{4-34}$$

And using the *d* from equation 8, the OCT model is

$$OCTrgc = 10^{0.1[10log_{10}(10870\, d\, rnfl)-c]} \tag{4-35}$$

Where *rnfl* is the mean RNFL thickness in µm from the OCT.

Medeiros and colleagues (Medeiros, Lisboa et al. 2012) finally proposed the use of a weighted mean reflecting the inverse relationship between disease severity of SAP and OCT estimates where the MD ranged from 0 to -30dB to obtain a final estimated RGC count in an eye eRGC as

$$eRGC = \left(1 + \frac{MD}{30}\right) OCT_{rgc} - \frac{MD}{30} SAP_{rgc} \tag{4-36}$$

4-36was used for the RGC count estimate in this study.

Electroretinogram recordings

Both eyes were dilated with 1% tropicamide drops prior to recording. Three strand Dawson, Trick, and Litzkow (DTL) (Dawson, Trick et al. 1979) wire electrodes were placed in the lower conjunctival fornices. Local anaesthetic drops were placed prior (1% tetracaine). Gold cup skin electrodes (Grass) were used for ground and indifferent electrodes. The indifferent electrodes were placed at the lateral canthi and the ground electrode was placed on the forehead. Impedances of all electrodes including the ground were checked using the inbuilt meter and were accepted if the impedance was less than 5kΩ.

An Espion E3 system with the ColorDome ganzfeld stimulator (Diagnosys LLC; Lowell, MA) was used to collect the ERG data. For the first step, a blue background of 10cd/m^2 (peak wavelength 465nm) was used for a preadaptation time of 2 minutes. A total of 125 flashes were then presented at a rate of 2 Hz, with a luminance of 1cd.s/m^2 using a red flash (peak wavelength 635nm) of 4ms duration. This step was compliant with the International Society for the Clinical Electrophysiology (ISCEV) extended protocol for the PhNR (Frishman, Sustar et al. 2018). Step 2 used the same photopic intensity of flash and background, but it was delivered as a white-on-white stimulus after a further 2 minutes preadaptation to the achromatic background. The stimuli were also presented at a rate of 2 Hz, with a 4 ms duration. The signals were collected with a sample rate of 4kHz (which is higher than the minimum 1kHz sample rate specified by the ISCEV protocol for recording the ERG) and an epoch length of 150ms, and with 30ms of pre-stimulus recording. Both raw data and filtered (band pass 0.3-100Hz) were recorded and signals with an amplitude of over 150uV were rejected.

Marker placement

The location of the a-wave and b-wave from the signal were determined by locating the b-wave as the global maximum and then the a-wave as the minimum in the segment between the start of the vector and the b-wave. The i-wave was located by determining the first peak after the b-wave. The PhNR1 and PhNR2 were determined to be the negative going troughs either side of the i-wave. A typical ERG averaged trace elicited from chromatic stimuli is shown in Figure 4-5

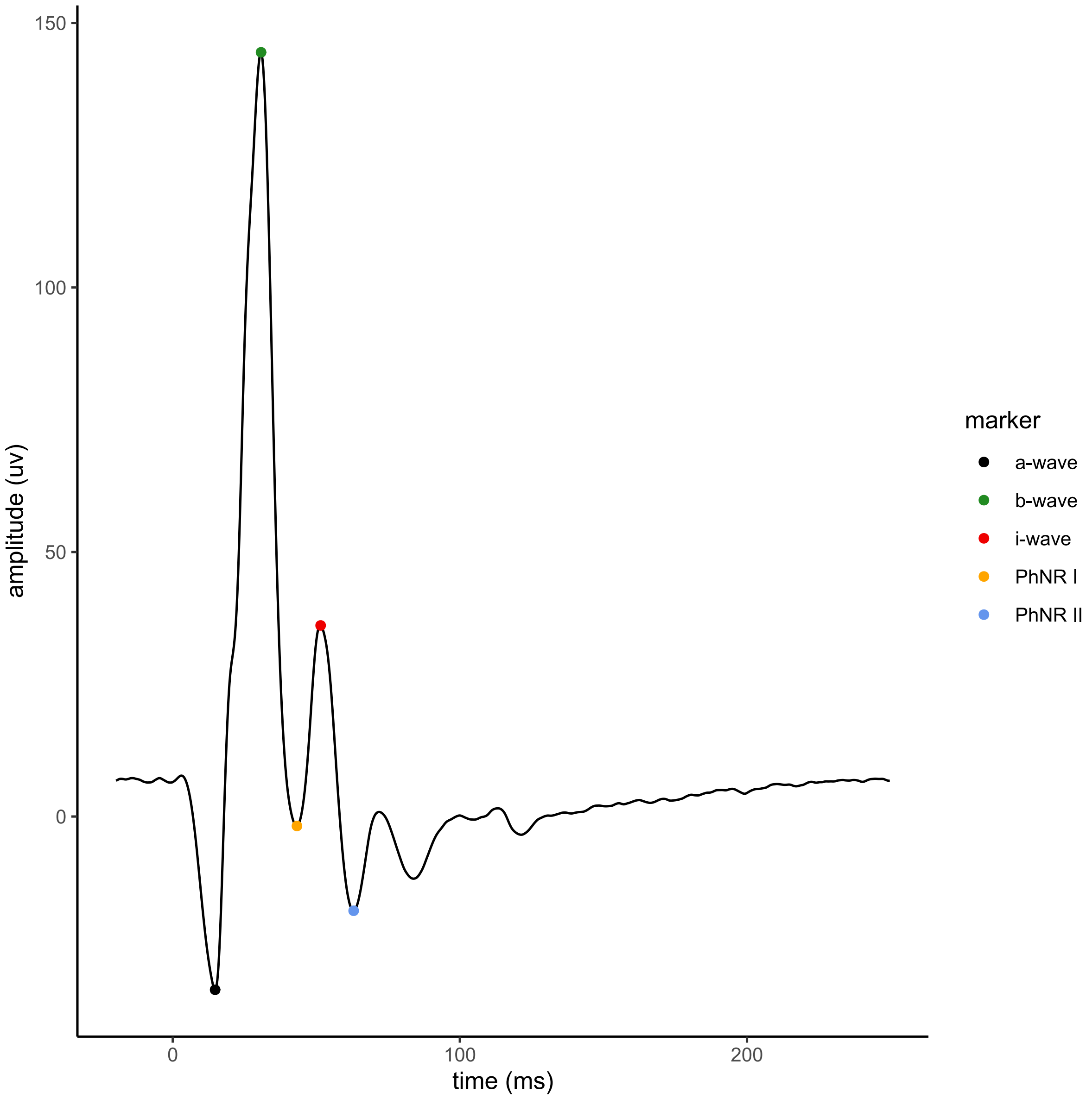


Figure 4-5: Example ERG with markers

Discrete Wavelet Transform

For this analysis all ERG trace averages from the database were selected where there had been results of at least 99 sweeps. This represented 216 averages. For each trace, the discrete wavelet transform was performed using the WMTSA package in R (Percival 2017). All available wavelets in the package were individually utilized and the results of all transforms stored in the SQL Server database. These wavelets consisted of the Haar wavelet, the Daublet family (Laba, Smith et al. 2007) d8 to d20, Symlet wavelets s2 to s20 (Chavan, Mastorakis et al. 2011), Coiflet wavelets c6 ,c12,c18 and c30 (Laba, Smith et al. 2007) and the best localized wavelets l2, l4, l6 ,l14, l18 and l20 .

The wavelets and their scaling functions are shown in Figure 4-6 and Figure 4-7 respectively. The Haar and d2 wavelets are identical. These wavelets differ in terms of orthogonality (the extent to which the coefficients are correlated), symmetry (which provides linear phase), compact support and filter order – an increased filter order of the mother wavelet increases smoothness. (Adamo, Andria et al. 2013)

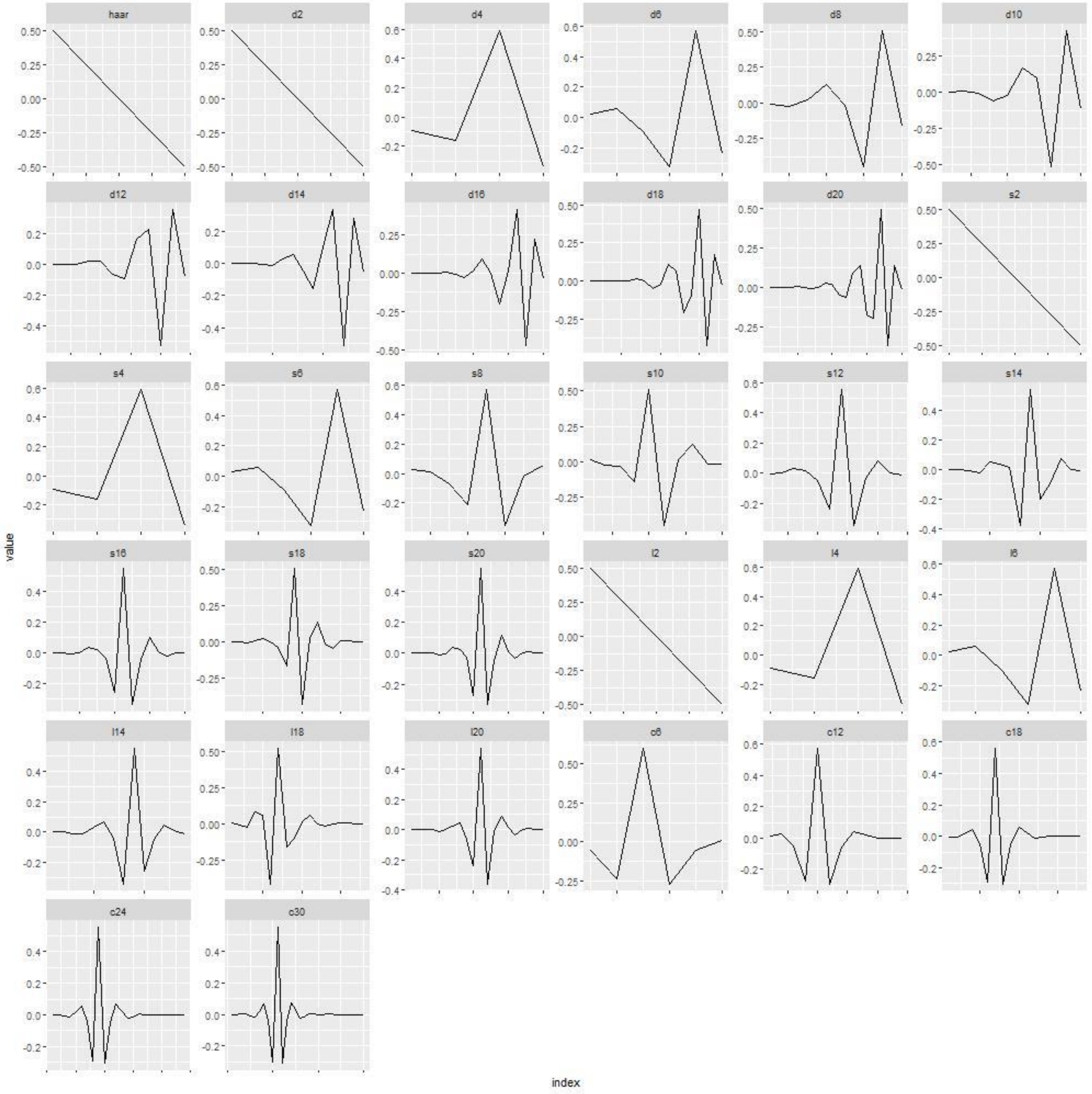


Figure 4-6: Discrete wavelets used for analysis - wavelet coefficients

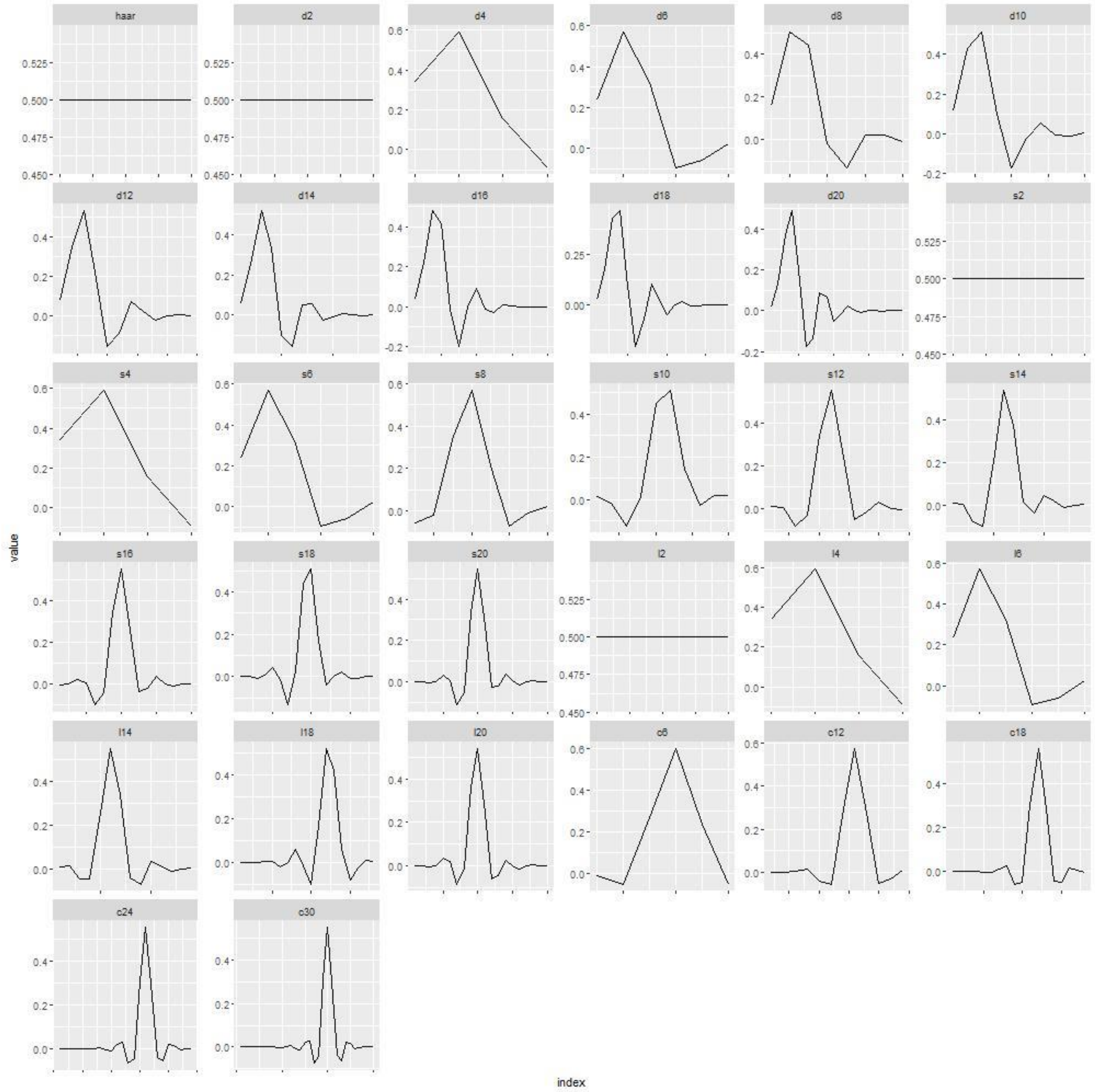


Figure 4-7: Discrete wavelets used for analysis - scaling functions

Figure 4-8 shows a typical ERG and the DWT performed in this case with the d8 wavelet (Daubechies wavelet with 8 vanishing moments). The heatmap plot above shows the absolute magnitude of the coefficients for the decomposition level crystals. The panel below shows the time domain plot. Figure 4-9 shows the original signal and each decomposition with the magnitude retaining the sign information. The horizontal axis of each subplot has been scaled to retain the same overall width for comparison.

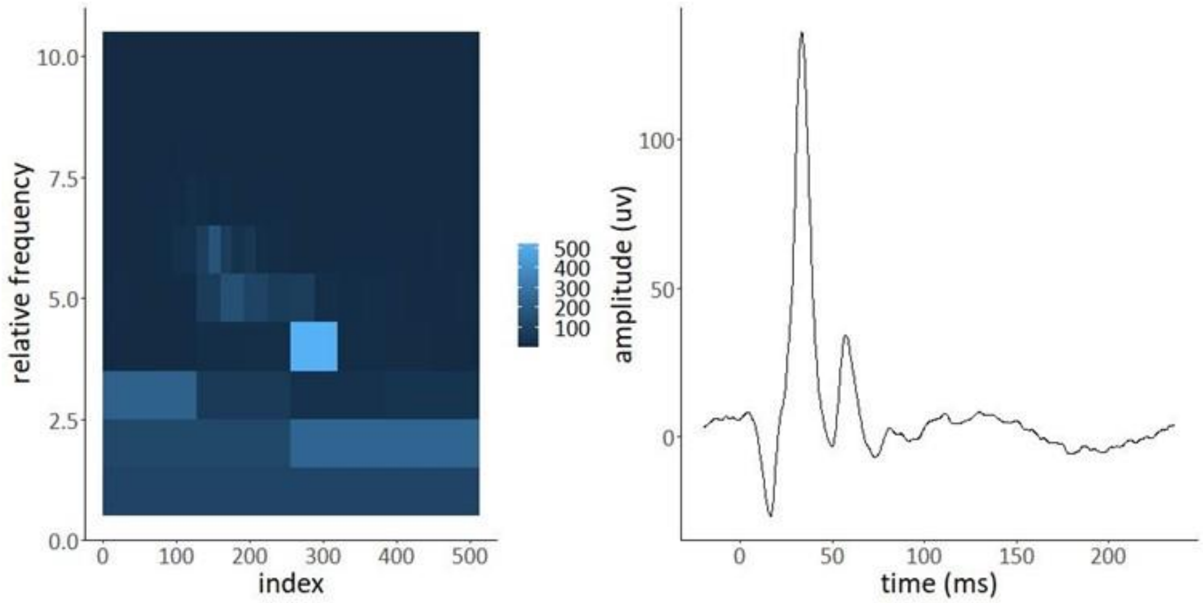


Figure 4-8: Discrete wavelet transform (DWT) of the electroretinogram (ERG). The upper panel shows the magnitude of the detail coefficients of the DWT (scalogram). The lower panel shows the ERG as a function of amplitude relative to time. Values higher up the vertical axis of the scalogram represent higher frequencies.

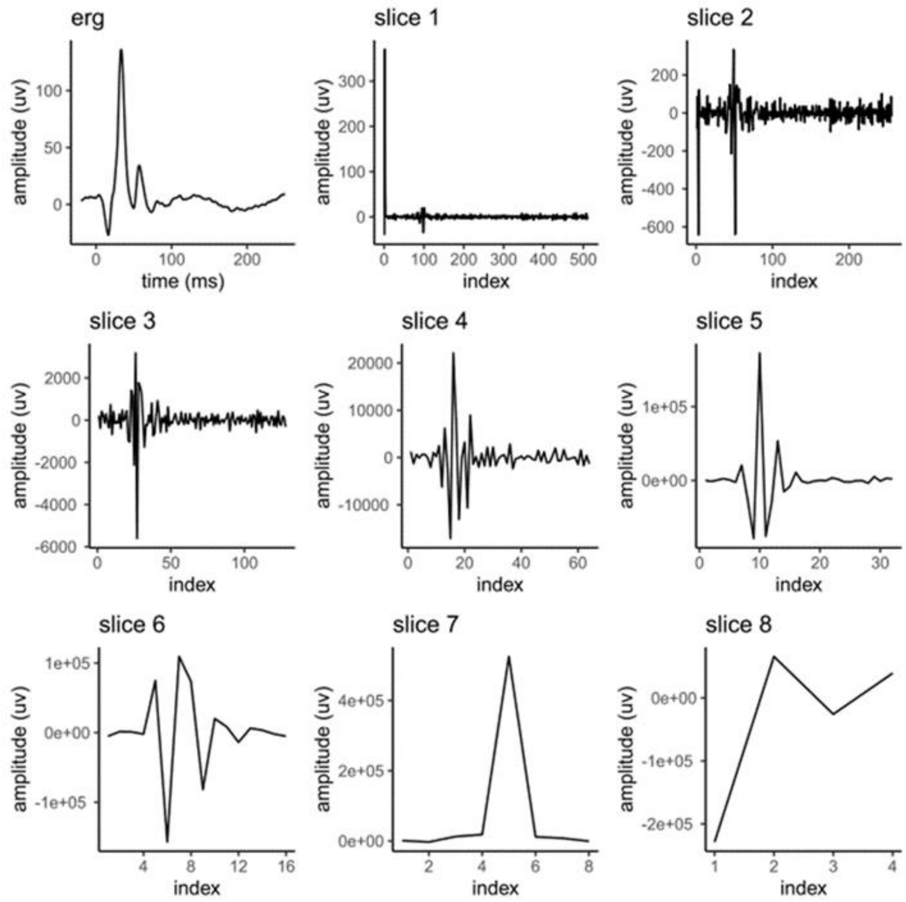


Figure 4-9: DWT decomposition of the ERG by level . The first panel represents the time domain signal. Each subsequent panel shows the detail coefficients plotted on a horizontal axis scaled to the same length as the original time domain signal. At the first decomposition level, the length of the signal is reduced by half as it is for each subsequent decomposition level.

Optimal wavelet selection was performed by minimizing the ratio of energy to Shannon entropy. The energy for each wavelet was calculated by the sum of the square of all of the wavelet coefficients. The Shannon entropy was calculated by the plug-in estimator for all of the coefficients over 20 bins. The results of this ratio calculation for each wavelet over the ensemble were summed to determine the optimum wavelet family. Bootstrap resampling of 55 samples was performed 1000 times to determine the confidence interval. Resampling was performed on participant rather than eye level to minimize bias. The wavelet selection process used the input feature matrix only and not the outcome measure – in a similar fashion to principal components analysis. Cross validation was therefore not required at this step.

To build the predictive model, the coefficients of wavelets at the sixth decomposition level (d6 crystal) were used. Each of these coefficients represented 72 samples(18ms).

For each of the wavelet families, a predictive linear model was use in R to determine correlation between the information contained in the d6 crystal and the retinal ganglion cell counts.

Matching Pursuit

The matching pursuit calculations were performed in MATLAB using the wavelet toolbox. All ERG trace averages from the database were selected where there had been results of at least 99 sweeps. This represented 216 averages. All traces were aligned so that the b-wave peak occurred at the same point (an index of 200) The Espion filtered traces were used. The data were imported into MATLAB.

Figure 4-10 shows a sample MP decomposition of the waveform shown in Figure 4-8 with the 12 atoms with the highest importance shown. The upper panel shows the atoms themselves and the lower panel shows each atom multiplied by the coefficient. The sum of the traces in the lower panel constitutes the reconstruction.

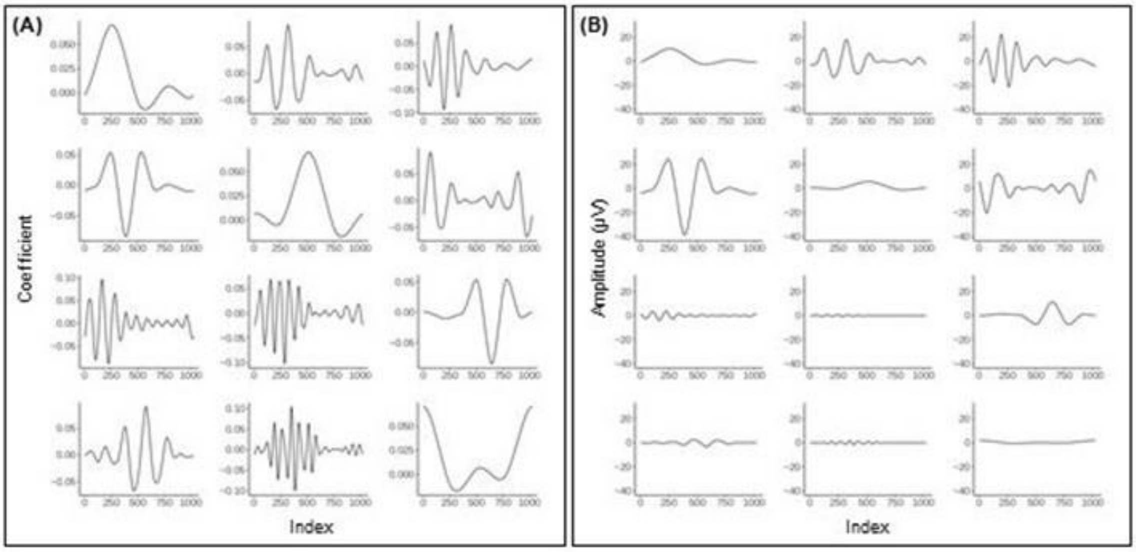


Figure 4-10: Sample matching pursuit decomposition. Panel A shows the dictionary of atoms used for the decomposition. Panel Bl shows the atoms multiplied by the coefficient. Note some of the coefficients are negative; the upper right trace of panel A has been inverted in panel B. The decomposition consists of the coefficients for each atom in the dictionary.

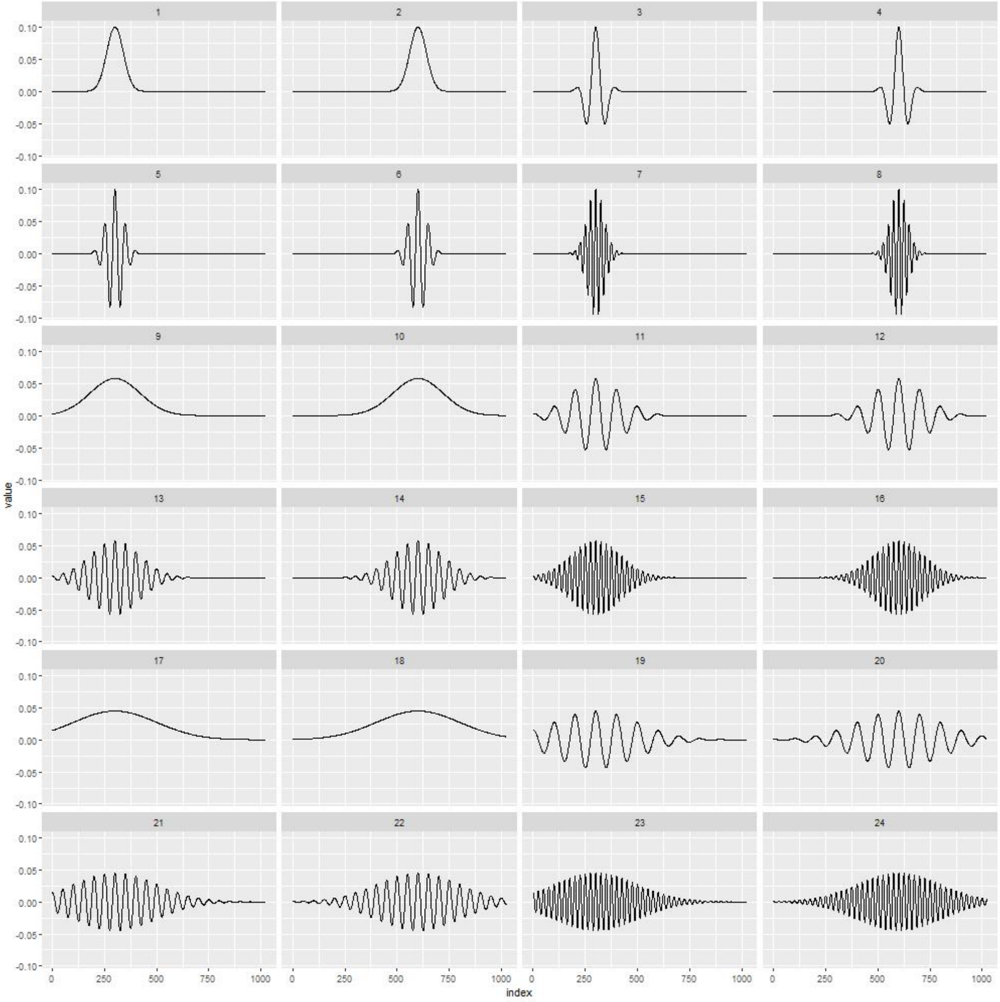


Figure 4-11: A selection of atoms from the Gabor dictionary

Dictionaries were created in MATLAB for the matching pursuit algorithm. A variety of dictionaries were tested including a custom Gabor dictionary (Chu, Narayanan et al. 2009) and all available wavelet packets from the wavelet toolbox. The generating wavelet packets from the toolbox were: Coiflet, Daubechies, Meyer, Fejer-Korovkin and symlet. Like the DWT, these resulting dictionaries had varying properties. The Feyer-Korovkin wavelet packet is said to generate more concentrated wavelet packets with less leakage of high frequency noise (Kadir, Saha et al. 2021).

The Gabor dictionary was constructed as per Chu (Chu, Narayanan et al. 2009). A selection of atoms from the Gabor dictionary is shown in Figure 4-11. A Gabor is a point by point product of a cosine and a Gaussian function and the parameters for the library include the spread of the exponential, the frequency of the cosine wave and the number of time shifts.

Optimization of the dictionary to the ensemble was performed over the choice of generating wavelet, the order, and the number of vanishing moments. For each dictionary, a smaller subdictionary of 25 atoms was selected as the top 25 atoms used within each full dictionary. The subdictionaries were compared to one another by decomposing and reconstructing each signal and finding the total root mean square error (RMSE) of the residuals over the whole ensemble. The best dictionary was considered that with the smallest RMSE. Again resampling was performed in the same manner as for DWT to determine the confidence intervals. As for the DWT, the outcome measure was not used for this step so cross validation was not required.

Continuous Wavelet Transform

The continuous wavelet transform was performed with the WMTSA package in R using both the first and second derivative Gaussian wavelet families. The matrix of coefficients was used for image analysis techniques described below.

Image based Feature Extraction

The Imager package (Barthelme 2019) was used in R to perform the SLIC superpixel extraction from the continuous wavelet and time-frequency distributions. For the CWT, the b-wave aligned waveforms used for the matching pursuit algorithm were used. For each CWT, 25 superpixels were extracted with a compactness of 0.3. The ensemble of all of the superpixel extractions were summed for each of the superpixels and this map used to determine the superpixel boundaries. As the k-means algorithm commences with a random seed, each superpixel map was checked to ensure that the superpixel numbering was consistent for the whole ensemble.

Texture based Feature Extraction

The GLCM package (Zvoleff 2020) was used to extract textural features from the CWT of the b-wave aligned ERG traces. All available features in the package were extracted. A window of 5x5 was used for the analysis with a shift of 1 and the angles 0, 45 90 and 135 degrees were specified with the algorithm summing the results for all of those angles. The features were determined to be the sum of values within the region of interest (ROI). The ROI was selected in a time period known to contain the PhNR and a range of frequencies known to be affected in glaucoma (Sarossy, Aliahmad et al. 2018). The ROI is shown as the coefficients within the black rectangle of Figure 4-19.

Predictive Modelling

Initial evaluation of the sets of features was performed by analysing correlations between the features and the modelled ganglion cell counts. The TFD and texture features sets showed very little correlation and were not investigated beyond this initial analysis.

The Caret (Kuhn 2017) package in R was used for predictive model fitting and tuning. Linear and multivariate adaptive regression splines (MARS) were fitted. Overfitting was avoided by using 10-fold cross-validation. For the MARS models, the degree was set to 1 and the maximum number of terms was set to 40. The degree is the level of interaction of terms allowed – in this case none allowed. An example second degree interaction term would be the a-wave amplitude multiplied by the b-wave amplitude. The number of terms in the model is a function of the number of input features and the number of breaks in the hinge function. Without a breakpoint, a single feature will have 2 terms: intercept and gradient. With a single breakpoint, this becomes 4 terms: intercept, the 2 gradients and the breakpoint location.

For each of the tuned 'final models', the prediction performance of the ERG for the eRGC counts was examined based on the proportion of variance explained (R^2) by the model. Each model was compared to the base case of the PhNR only model. The significance of difference was determined using a bootstrap technique (*n* = 1,000 resamples) which dealt with the within-subject correlations through resampling at the individual level.

4.7 Results

4.7.1 Characteristics of the Glaucoma Cohort

Characteristics of the glaucoma cohort are shown in Table 4-2

Table 4-2: Characteristics of the individuals and eyes with glaucoma in this study

Characteristics	(103 eyes from 55 individuals)
Individual Level	
Age (years)	75 (66, to 80)
Gender (female)	21 (38%)
Diabetes (present)	7 (13%)
Hypertension (present)	35 (64%)
Eye Level	
Refraction sphere (D)	0.00 (-1.00, to 0.50)
Visual acuity (logMAR)	0.0 (-0.1 to 0.1)
Intraocular Pressure (mm Hg)	15.0 (12.0 to 16.0)
Mean Deviation (dB)	-2.5 (-5.9 to -0.5)
Retinal Nerve Fiber Layer Thickness (µm)	73.0 (62.8 to 82.5)
eRGC ('000s)	601 (470 to 753)

Continuous statistics presented as median and interquartile range, categorical statistics as number and percentage. logMAR = logarithm of the minimum angle of resolution; eRGC = estimated retinal ganglion cell count

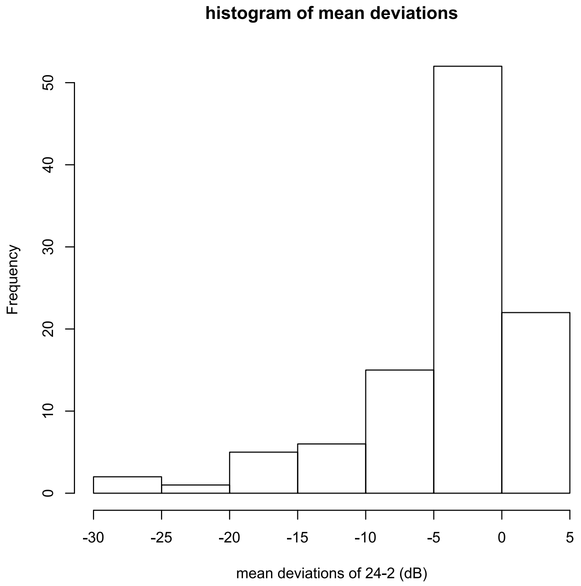


Figure 4-12: Histogram of mean deviation

A histogram of the distribution of mean deviations of the eyes in the cohort is shown in Figure 4-12. The majority of the eyes had mild to moderate glaucoma and for some the glaucoma was not evident by analysis of the MD alone.

Figure 4-13 shows a histogram of the distribution of the average RNFL thickness measurements. As this is a linear measurement compared with the log nature of the MD, the distribution is closer to a normal distribution compared to the lognormal shape of the MD. The histogram of the eRGC measure is shown in Figure 4-14. In this study both eyes were included. The correlation coefficient between the eRGC of the right and left eye in subjects where both eyes were tested was 0.61. Bias was avoided in the bootstrap resampling by resampling at a subject rather than eye level.

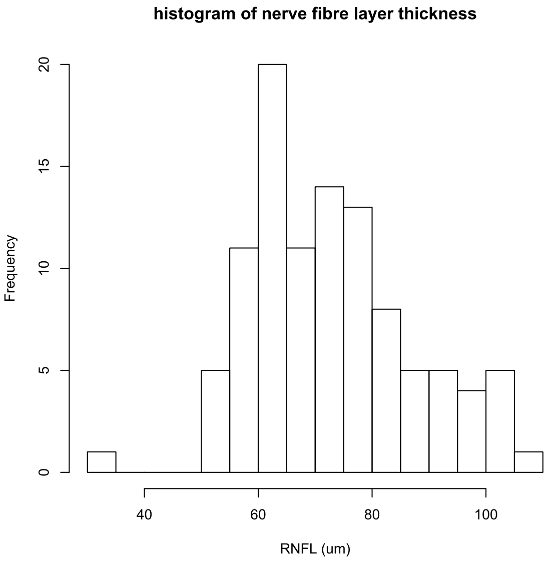


Figure 4-13: Histogram of RNFL thickness

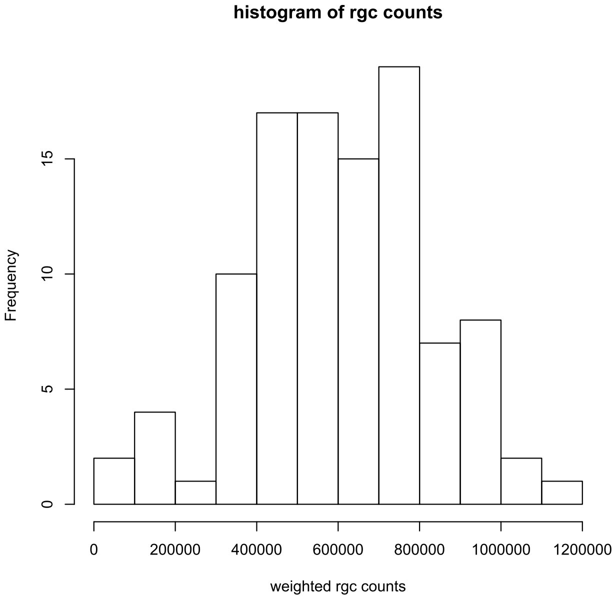


Figure 4-14: Histogram of weighted RGC count

4.7.2 Mother Wavelet Selection

Figure 4-15 shows the result of the energy to entropy ratio calculations over the entire ensemble of eyes. The wavelet with the highest value (and the optimum by this technique) is the Coiflet 8 wavelet. This was the wavelet used for the modelling of glaucoma severity.

At a sampling rate of 4kHz, the 6th decomposition level represents the frequency band centred on 20Hz

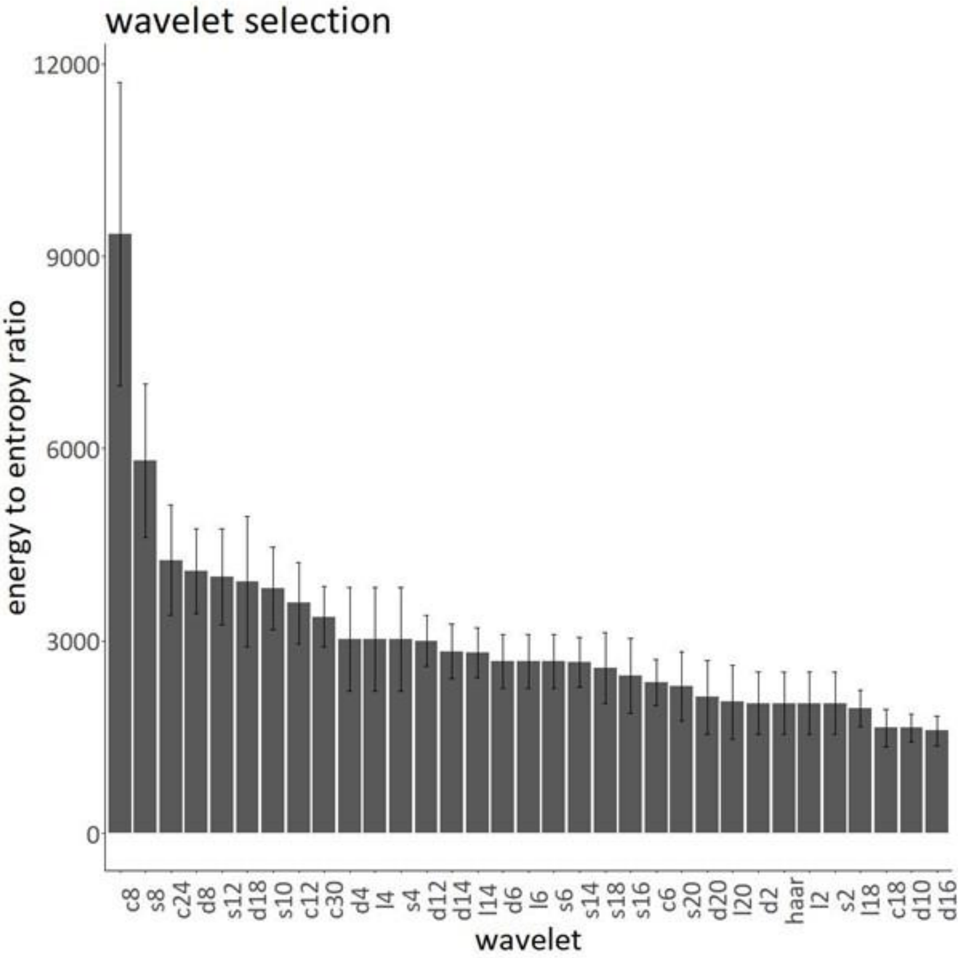


Figure 4-15: Energy to entropy ratio for all of the wavelets in descending order. Wavelets are identified by family and length. C = Coiflet, s = symlet, d= Daubechies, l= best localized. Higher values indicate more optimum discrete wavelet for the decomposition. The error bars show 95% confidence intervals for the bootstrap resampling.

4.7.3 Superpixel decomposition of Continuous Wavelet Transform

Figure 4-16 shows the ensemble sum of all continuous wavelet transforms (CWT) as a heatmap plot for superpixel 21. These sums were used to form the decomposition map shown in Figure 4-17. Each geographical area in this plot represents a single superpixel. The decomposition is performed by applying this mask to the CWT. That is the coefficients within the CWT that correspond to the map of superpixel 21 are extracted as the sum of those coefficients. The transformed output, like the matching pursuit decomposition, is composed of an index (the superpixel number) and the summed coefficients from that area.

Like the textural features extraction but unlike the matching pursuit and discrete wavelet transforms, reconstruction of the original signal is not possible from the transform.

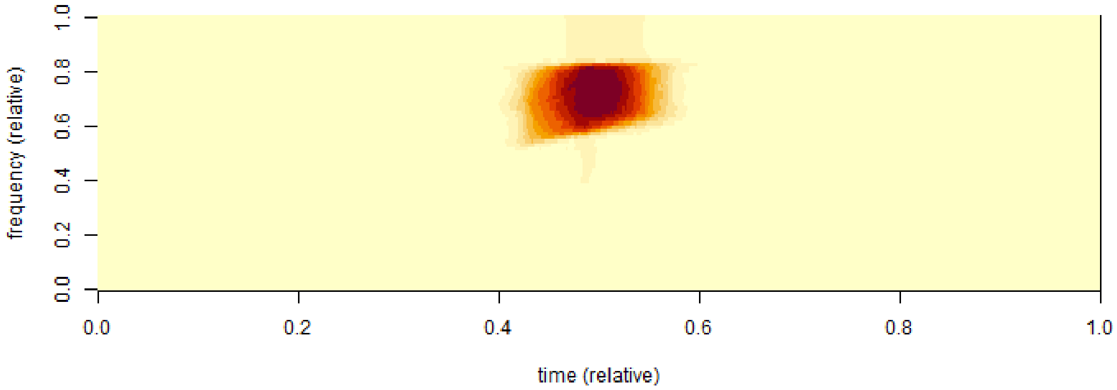


Figure 4-16: Ensemble sum for superpixel 21.

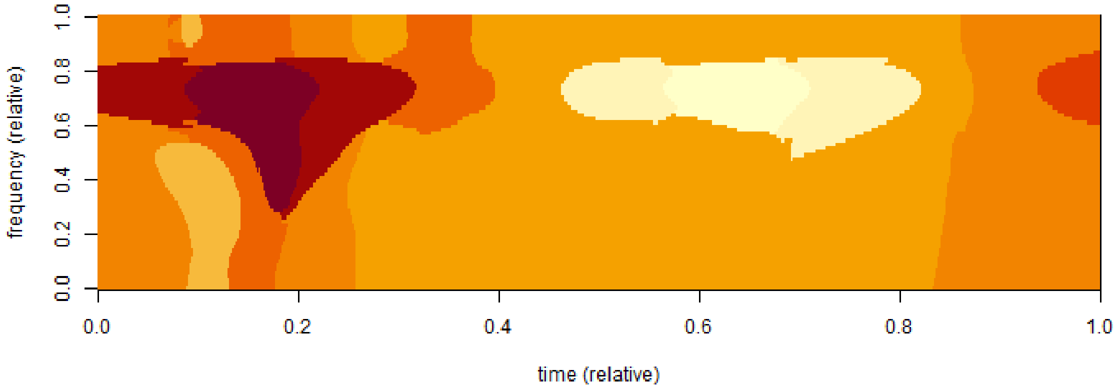


Figure 4-17: SLIC decomposition mask.

4.7.4 Textural Features of the Continuous Wavelet Transform of the ERG

The GLCM feature maps for the sample waveform shown in Figure 4-5 are illustrated in Figure 4-18. The region of interest (ROI) used for the calculations is shown in Figure 4-19. The value for the parameter was determined to be the sum of the parameters within the ROI.

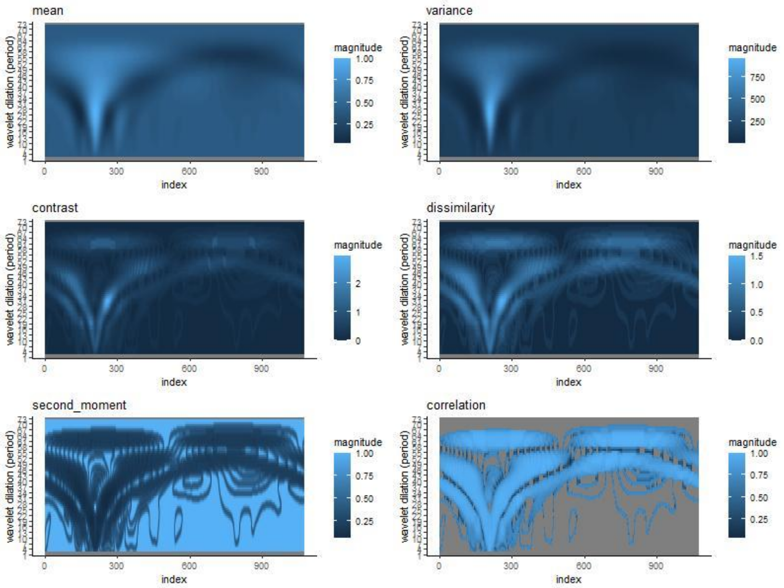


Figure 4-18: GLCM textural features for the sample waveform.

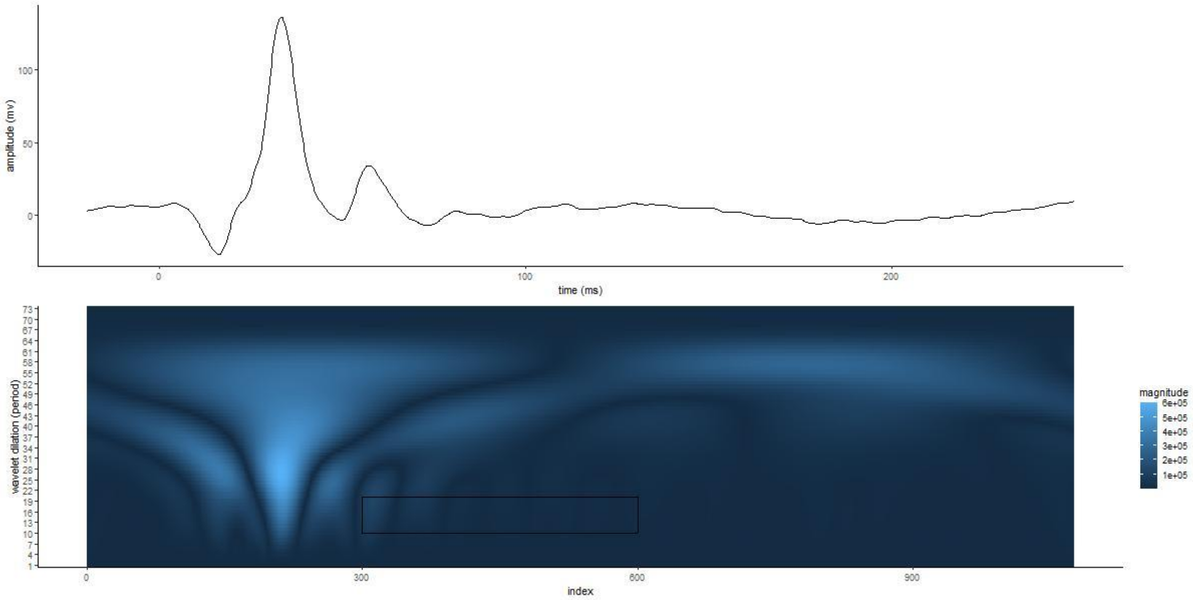


Figure 4-19: Region of interest.

Table 4-3 shows the correlation between the GLCM features from the ROI and the glaucoma severity as measured by the eRGC. Echoing the results of chapter 3 of this thesis, a small but significant correlation was found between the eRGC and the entropy of the texture of the signal. Again, there was decreased entropy with more severe glaucoma (lower eRGC).

Table 4-3: Correlation of eRGC and GLCM measures.

texture measure	Correlation with eRGC	p value
mean	-0.051	0.49
variance	-0.043	0.56
contrast	0.128	0.08
homegeneity	-0.131	0.07
dissimilarity	0.131	0.07
entropy	0.134	0.05

4.7.5 Matching Pursuit Dictionary Selection

The Gabor dictionary performed poorly and was not considered for further analysis. The results for the matching pursuit dictionary selection are shown in Figure 4-20. Optimization was undertaken across mother wavelet, order of wavelet and number of vanishing moments. For all of the wavelet packet generators except the Fejer-Korovkin, the order of wavelet packet made little difference. There were small differences between the mother wavelets but for all families, the best results were obtained with 6 vanishing moments. The best overall dictionary was that created by the symlet of order 6 (wpsym6) with 6 vanishing moments.

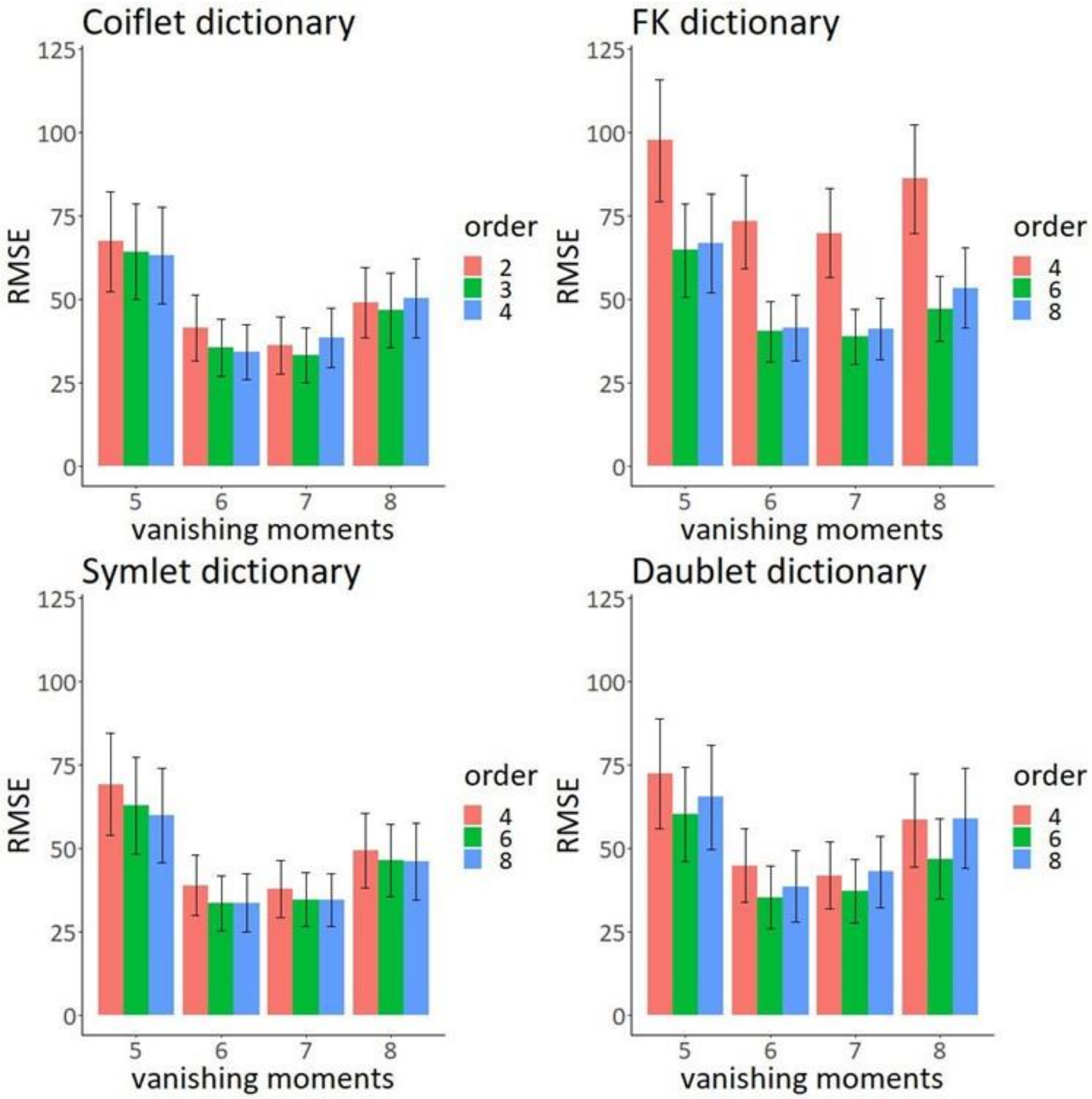


Figure 4-20: Matching Pursuit Dictionary Optimization. RMSE=root mean square error, FK = Fejer-Korovkin wavelet. The bar graphs show the residual sum of squares from reconstruction of the ensemble traces for different mother wavelet packets (by panel), vanishing moments (abscissa) and order of wavelet packet (by colour). Error bars represent 95% confidence intervals from the thousand fold bootstrap resample.

An example of a stepwise reconstruction is shown in Figure 4-21 The top panel shows the original signal. Lines have been placed at the a-wave and b-wave peaks for reference. These lines have been duplicated on the bottom pane and the signal reconstructed from the custom dictionary. Note that the custom dictionary is the 12 best atoms on average from the matching pursuit algorithm across all of the traces. Here, the signal is reconstructed and the process illustrated from the 5th to 12th iteration. In this particular case with this trace and dictionary, the reconstructed wave shows good concordance for the a-wave and PhNR but the fit for the b-wave peak is not as good.

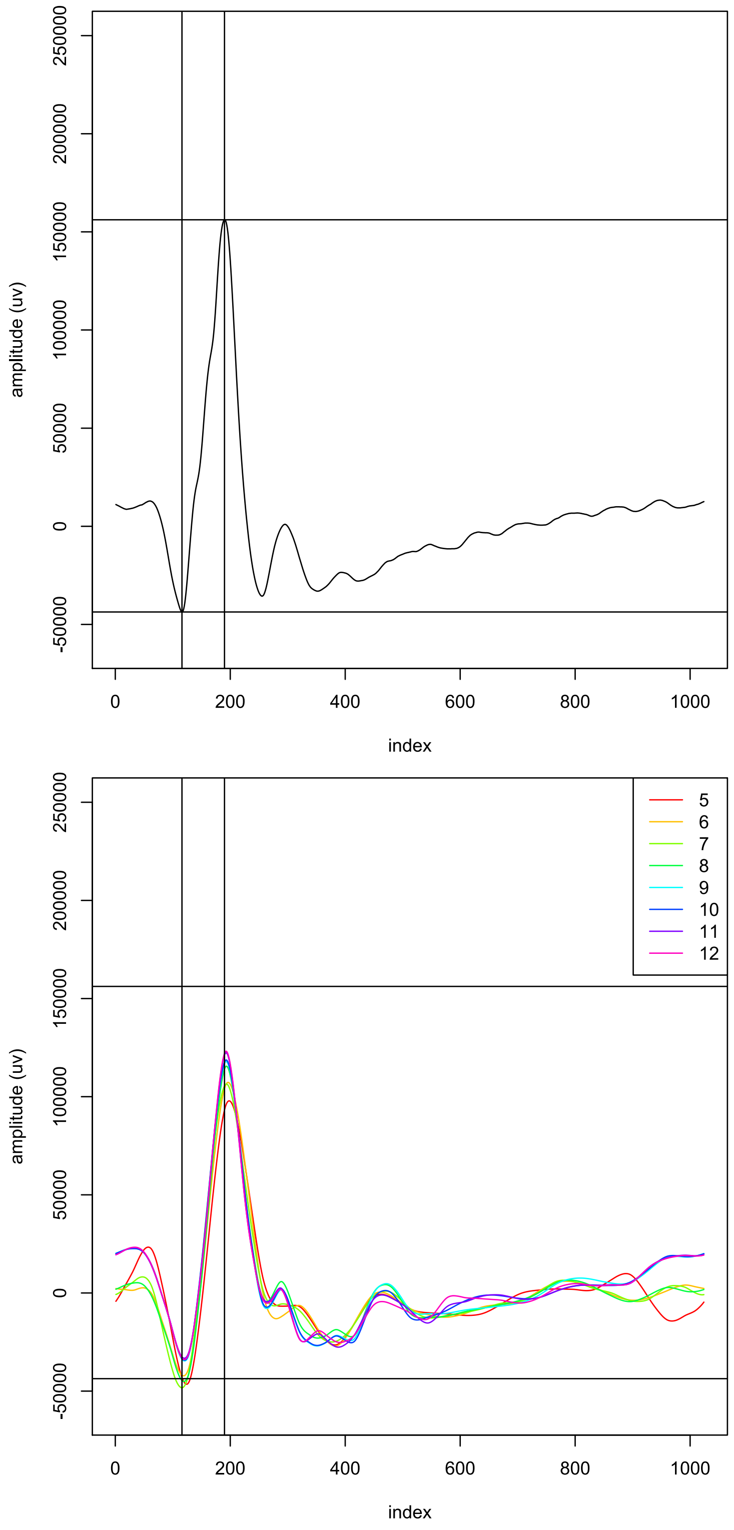


Figure 4-21: Matching pursuit reconstruction.

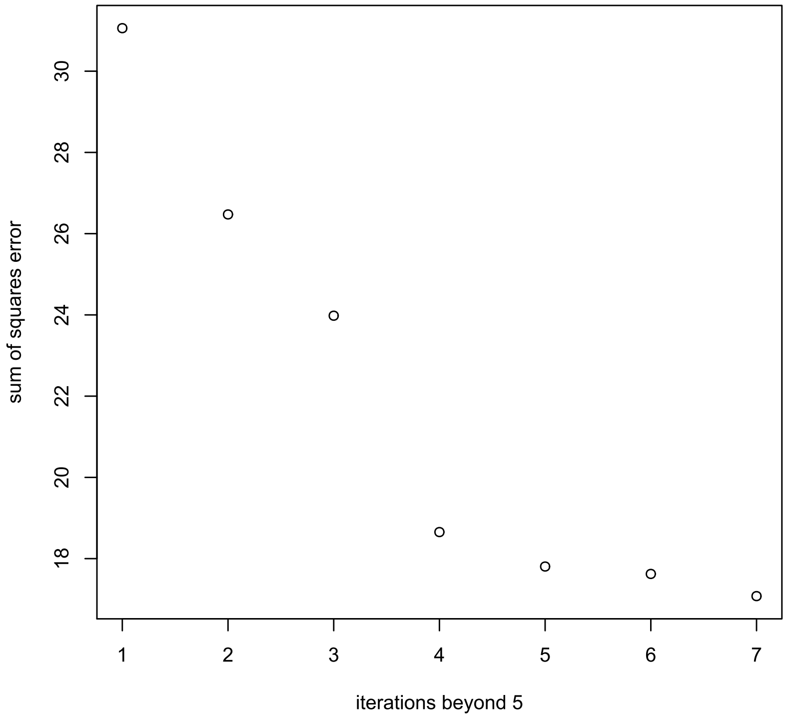


Figure 4-22: Order of MP decomposition sum of squares error

The asymptotic process of fitting the atoms is shown in Figure 4-22 with the sum of squares error dropping at a decreasing rate with additional atoms. A full decomposition with the 1024 atom dictionary would yield a near perfect fit but at the cost of near 1024 parameters. Analogous to the utilisation of the wave amplitudes, this decomposition yields a set of features that is much smaller than the length of the original signal vector.

4.7.6 Models with conventional features

Table 4-4: Predictive performance conventional features

			P-Value for Comparison Between				
	R^2		PhNR vs. Markers		Linear vs. MARS	Chromatic vs. Achromatic	
	Linear	MARS	Linear	MARS		Linear	MARS
Chromatic							
PhNR	0.09	0.11	0.04	0.04	0.26	0.32	0.40
Markers	0.22	0.33			0.04	0.24	0.24
Achromatic							
PhNR	0.16	0.18	0.04	0.04	0.26	-	-
Markers	0.31	0.45			0.04	-	-

PhNR = photopic negative response. Significance adjusted for multiple comparisons with the method of Benjamini (Benjamini and Yekutieli 2001)

The PhNR/B ratio gave poor predictive performance with R^2 values ranging from 0.028 for the linear chromatic model through to 0.073 for the MARS chromatic model. Table 4-4 summarizes the predictive ability of the different models using the ERG for the eRGC measure for the PhNR alone and the full set of markers. The predictions of eRGC were significantly better with the models utilising the full feature set from the ERG, compared with those using the PhNR alone ($P \geq 0.02$ for all). The predictive performance of the MARS models was also significantly better than the linear models using the full set of markers ($P = 0.01$ for both), but not when using the PhNR alone ($P \geq 0.14$ for both). Finally, the models derived from ERG recordings with achromatic stimuli performed better at predicting eRGC counts than ones using chromatic stimuli, but it this difference did not reach statistical significance ($P \geq 0.19$ for all).

4.7.7 Models with time-frequency features

Table 4-5: Predictive model performance time-frequency features

			***P*-Value for Comparison Between**				
	R^2		**PhNR vs. TFA**		**Linear vs. MARS**	**Chromatic vs. Achromatic**	
	Linear	**MARS**	**Linear**	**MARS**		**Linear**	**MARS**
Chromatic							
Markers	0.21	0.34			0.021	0.235	0.244
Superpixels	0.46	0.43	<0.001	<0.001	0.390	0.334	0.024
MP	0.35	0.45	0.035	0.146	0.101	0.259	0.008
DWT	0.25	0.57	0.334	0.010	<0.01	0.277	0.424
MP + markers	0.43	0.63	0.001	<0.001	0.005	0.005	0.146
DWT+ markers	0.41	0.53	0.001	0.021	0.083	0.083	0.369
Achromatic							
Markers	0.31	045			0.025		
Superpixels	0.51	0.71	<0.001	<0.001	0.021		
MP	0.44	0.72	0.050	<0.001	<0.001		
DWT	0.33	0.55	0.390	<0.001	<0.001		
MP + markers	0.50	0.73	0.010	<0.001	<0.001		
DWT+ markers	0.47	0.56	0.019	0.118	0.122		

MP=matching pursuit, DWT=discrete wavelet transform. Significance adjusted for multiple comparisons with dependent statistics by the method of Benjamini (Benjamini and Yekutieli 2001)

Table 4-5 show the final model fit performance for the time-frequency feature sets. MARS and Linear models showed highly significantly better performance in predictive ability than the PhNR alone. MARS generally performed better than linear and achromatic stimulus yielded a better predictive performance than chromatic with the MARS models for all models but the PhNR alone. For the linear models, the achromatic stimulus performed significantly better only for the models using the discrete wavelet transform features.

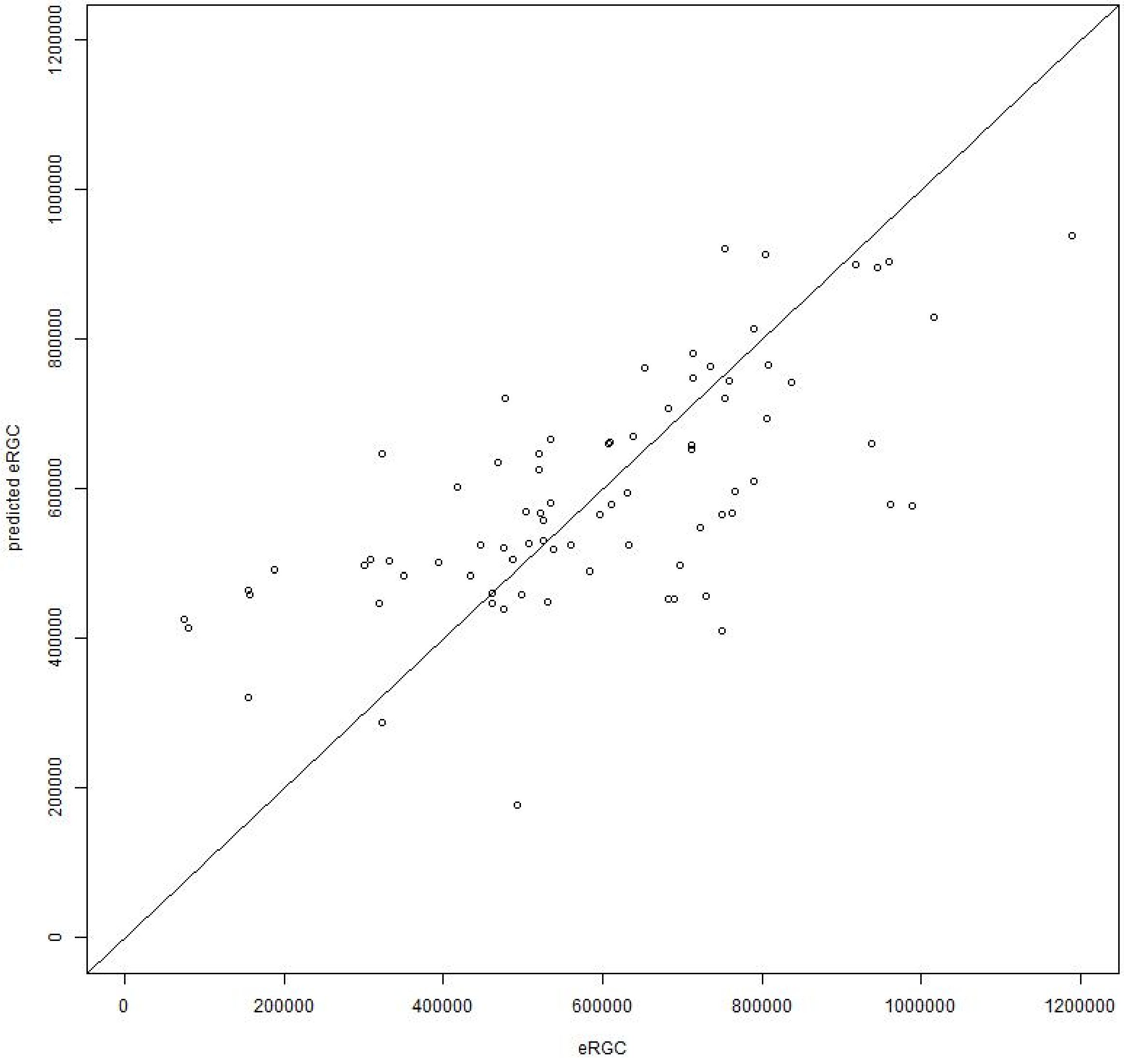


Figure 4-23: Model v actual rgc count linear achromatic marker

Figure 4-23 illustrates the fit of the 5 marker linear model compared to the actual weighted ganglion cell counts. The line represents the line of best fit. What is interesting is that the line of best fit has a negligible intercept and a near perfect gradient of 1. The fit shows fair homoscedasticity as shown in Figure 4-24.

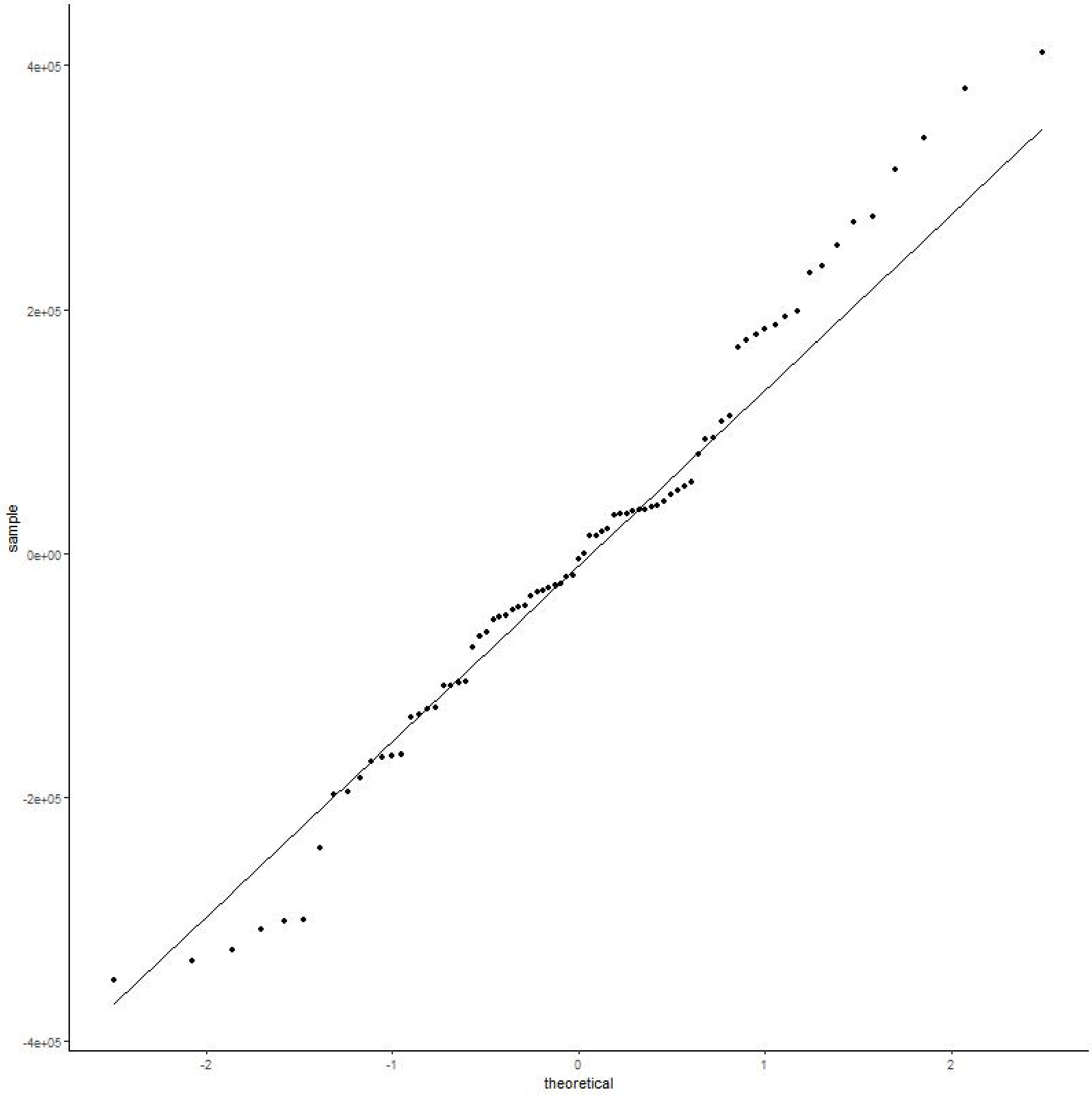


Figure 4-24: QQ plot of linear model fit

Our previously described normalized PhNR measurement performed poorly in predicting the ganglion cell soma counts with the linear model. The addition of the extra markers did not add significantly to the model performance over inclusion of the b-wave alone.

4.8 MARS model fits

There is a plethora of models that can be fitted, tuned and cross validated to the feature matrix. Most of the models offer poor intuition and are black-box like. Multivariate Adaptive Regression Splines (Friedman 1991) offers reasonable intuition and moderately low complexity. The technique fits a linear combination of predictor hinge functions so that the coefficient applied to a predictor varies piecewise according to its magnitude. It can dramatically increase the degrees of freedom of a model as the parameters include the coefficient and hinge point for each predictor. There is at this point, no accepted way of evaluating model fit by Akaiki information criterion (AIC) (Akaiki 1973) or similar criteria.

Fitting these models is implemented in R in the Earth Package (Milborrow 2017). This title is a play on words viz. the planet Mars. To explain the process, an illustration will be made for the markers model of the achromatic stimulus.

The summary of the model from R is shown in Enclosure 4-1. The text states that 4 of 5 predictors were used and the hinge functions show that the b-wave marker was pruned from the model. The termination or stopping condition in this example was when the number of terms had reached 21. The default for the tuning process is the {larger of (20 or 2 times the number of features) plus 1} up to a maximum of 200. The enclosure indicates that the stop condition was triggered when the number of terms reached 21.

Enclosure 4-1: Model output from R for achromatic marker model

```
Call: earth(x=matrix[79,5], y=c(937500,789300...), keepxy=TRUE, degree=1,
nprune=40)
[
            coefficients
(Intercept)       940162.91
h(a- -30478.5)      -44.60
h(a- -24662.6)       41.89
h(i-25604.1)         13.14
h(P2- -29860.3)     -10.00
h(-22892.4-P1)      -12.98

Selected 6 of 16 terms, and 4 of 5 predictors (nprune=40)
Termination condition: Reached nk 21
Importance: a, P2, P1, i, b unused
Number of terms at each degree of interaction: 1 5 (additive model)
```

As per Enclosure 4-1, the ganglion cell model has an intercept of 940 162 ganglion cells and the only term that can increase the count is the i-wave term. Larger a-wave absolute values (more negative) reduce the ganglion cell count less but the effect has two gradients and saturates. Conversely, more negative PhNR1 wave amplitudes reduce the ganglion cell count more. Higher i-wave amplitudes increase the number of predicted ganglion cell somas after a threshold.

The effect of the PhNR2 wave amplitude is semi-intuitive. This is referenced to the baseline and more negative values reduces the ganglion cell count by less (ie increases the total soma count predicted by the model). This effect is up to a ceiling of around -29mV. The effect of this term on soma count predictions continues even through positive values of the marker. The process driving the negative going wave underlying the PhNR1 and PhNR2 has physiological plausibility in that it corresponds to the time when the ganglion cells would be firing.

We will consider an individual case from the achromatic stimulus data. The ERG marker parameters are as follows:

marker	Amplitude (nv)
a	-2 295
b	107 761
PhNR1	17 601
i	7 539
PhNR2	8 957

For each of the terms, the effect on the ganglion cell estimate is shown on the corresponding graph in Figure 4-25. The value in nV is projected by the red line to the graph intersection and the value of the term is read off the vertical axis at the level of the blue line. The result summing the terms and the intercept is 287, 902.

In this example, the PhNR2 and i-wave have no effect, the a-wave and the PhNR2 terms contribute to give an overall estimate of RGCs of 287, 902, being similar to the estimate of 322, 678 RGCs for this example determined from the visual field and OCT data.

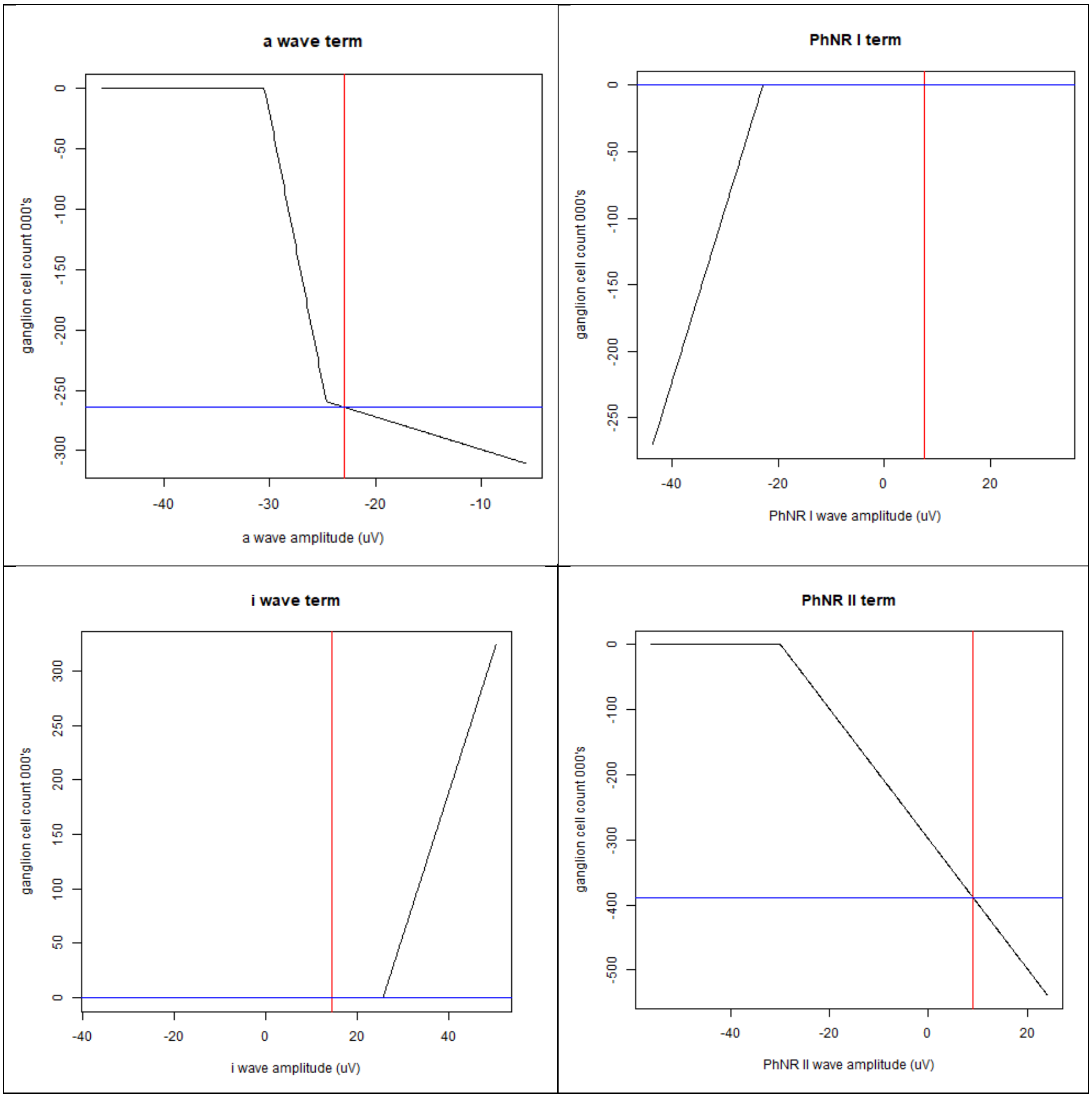


Figure 4-25: MARS model - achromatic stimulus

4.9 Effect of Diabetes and Hypertension

Table 4-6: Effect of diabetes and hypertension on the eRGC

Characteristics	*present*	*absent*	*significance*
Diabetes			
number	12	91	
eRGC ('000s)	531.32 (187.1)	612.56 (225.8)	0.18
Hypertension			
number	64	39	
eRGC ('000s)	573.44 (225.5)	651.76 (216.2)	0.08

eRGC = estimated retinal ganglion cell count mean (sd). Significance calculated by 2 tailed t test

There was a small but non-significant difference in the eRGC count between eyes of patients with and without diabetes and systemic hypertension. This is shown in Table 4-6. In both cases, the presence of the systemic disease was associated with a reduction in the eRGC. The benefit of including the presence or absence of these conditions within the predictive models was tested. For all of the MARS models, the two systemic disease parameters were automatically tuned out of inclusion – that is the final models did not include terms for these two conditions. For each of the linear models, the Aikake Information Criterion (AIC) was tested with the parameters included or excluded. For every model, the AIC was higher with the two systemic parameters included, indicating that models including these features had a lower goodness-of-fit then models which did not include them. These findings suggest that these two features do not affect the electroretinogram independently of the association with glaucoma.

4.10 Discussion and Summary

Glaucoma is a disease of dysfunction and death of retinal ganglion cells (Chrysostomou, Trounce et al. 2010). The pathophysiology at a cellular level is complex but there has been much recent progress in the field (Tribble, Vasalauskaite et al. 2019). Manifestations of the initial dysfunction include changes in the mitochondria (Williams, Harder et al. 2017), dendrites (Williams, Howell et al. 2013) and synapses (Berry, Qu et al. 2015). Ultimately ongoing stress leads to apoptosis (Reichstein, Ren et al. 2007) from which there can be no recovery. There is much interest (Fry, Fahy et al. 2018) in the potential to identify this pre terminal dysfunction in retinal ganglion cells as it is possible that, with decrease in metabolic stress, the cells could recover.

For a clinician managing a patient suspected of having glaucoma there are a small number of critical decisions. First, whether the patient has glaucoma and needs treatment. In statistical learning, this is posed as a classification problem: given the feature set, what is the probability that the individual falls into one of two or more classes – generally glaucoma versus non-glaucoma. In reality this is a relatively trivial problem in that in the borderline cases where there is a potential delay in diagnosis, the actual loss of opportunity to treat will rarely lead to a catastrophic and permanent loss because the clinician will review the individual and retest in an appropriate timeframe. There could be a subset of patients who have an underlying vulnerability and who could be susceptible to sudden and catastrophic loss but this is not generally known

Once a decision to treat glaucoma has been made, there are relatively few alternatives available to the clinician. Caloric restriction (Guo, Kimura et al. 2016) , resveratrol (Luna, Li et al. 2009) and vitamin B3 supplements (Hui, Tang et al. 2020) have been proposed and have some potential. Neuroprotection (Yadav, Sharma et al. 2020, Kumar, Ramakrishnan et al. 2021) and agents that directly affect the retinal microcirculation (Park, Kwon et al. 2011) are another alternative mode of treatment. The overwhelmingly most important method of treatment however is reduction of the intraocular pressure.

There are many methods to reduce the intraocular pressure including a range of topical eye drops, surgical and laser procedures. All methods have the potential for complications and adverse effects as well as financial cost so that the clinician needs to determine the cost/benefit ratio of the treatment and an appropriate review period. In practice, for most patients any target intraocular pressure is achievable although the cost in risk, side effects and monetary cost is often too great and the decision revolves around the assumption of an appropriate risk and cost target intraocular pressure.

Appropriate investigations are therefore essential in making both of these key clinical decisions. The OCT is a structural test. It can in theory detect sick ganglion cells via changes in the size of the somas, the interstitial spaces or the synaptic zones (Kalesnykas, Oglesby et al. 2012). It is, however, generally thought that OCT changes represent permanent loss of ganglion cell somas.

Perimetry is a functional test and therefore is affected by the combination of ganglion cell loss and dysfunction. The difference between the structural loss and the functional loss could in theory predict the extent to which recovery is possible. Some dysfunctional cells however have no potential for recovering even partial function. Measurement of the visual field through psychophysical experiments does however interpose a large degree of neurological processing between the site of pathology (the ganglion cell) and the measurement of the response (patient pushing a button).

Electrophysiological testing does have the potential to more directly measure the function of an individual class of cells and in doing so measure the problem closer to its source. The photopic negative response has been identified as being correlated with glaucoma and indeed the results in this chapter show this to be the case. This is manifested by the reasonable R squared values of a linear model using the PhNR as the only predictor.

The PhNR has physiological plausibility too in that the timing of the waveform corresponds with the time that the ganglion cells would be spiking. Chapter 2 of this thesis explores techniques that can enhance the repeatability of the response.

At first glance, it would seem implausible that the a-wave or b-waves could be affected by glaucoma and that therefore they would have limited predictive value in determining the disease severity. These waves are known to come from the photoreceptors and bipolar cells respectively – cells not directly affected by glaucoma. With the linear modelling, a moderate increase in R squared was observed by the inclusion of the b-wave – more so for the chromatic than the achromatic stimulus. This was of course at the expense of an additional degree of freedom in the model. Inclusion of the full suite of markers improved the prediction still further but again at the cost of additional degrees of freedom.

There are several potential explanations for this phenomenon. First, the b-wave amplitude could be used as a normalizing factor – accounting for overall gain changes between eyes or between sessions. This approach has been used before by measuring the PhNR amplitude from the b-wave peak or by inclusion of the a-wave and b-wave marker voltages in the construction of a composite ratio (Wu, Hadoux et al. 2016). In general, these composite measures have increased reliability at the expense of correlation with glaucoma severity.

Structural and functional changes in the ganglion cell and outer plexiform layer could have effects on the physiology of the cellular processes generating parts of the signal before the b-wave and so that the possibility remains that the ERG could have detectible changes before the PhNR peak and even before the b-wave peak.

The central hypothesis of this chapter is that the ERG contains information relevant to the determination of the overall state of ganglion cell function. The ganglion cells carry a maintained discharge and light stimulation affects the rate and pattern of this discharge. Eliminating the action potential with tetrodotoxin was shown to affect the ERG post b-wave peak in a similar way to experimental glaucoma. It is likely that the changes to the firing patterns overall and the relative spatial distribution of this firing will change the ERG in the time, time frequency and frequency domain.

In this chapter, I have explored a number of different approaches to extract features from the ERG via a time-frequency approach and to project these results onto an output scale which is an overall index of ganglion cell function. Even with linear modelling, the inclusion of extra features from the ERG traces yielded significantly better predictive ability than the use of the PhNR amplitude alone. It is likely that a single process generates the negative going wave after the b wave recognises as the PhNR and a different process producing the upgoing i-wave overlaps the response; dividing this into PhNR1 and PhNR2. A MARS modelling approach allows the inclusion of all three markers and corrects for collinearity between them. A time-frequency approach such as demonstrated in this chapter could isolate the two processes even more efficiently.

With a more sophisticated model such as MARS, use of the more novel features resulted in substantially better predictive ability. The potential criticism of overfitting is negated here by 10 fold cross-validation and parameter tuning. Some computational complexity exists in implementing any of these techniques although open source libraries for Python are available.

In one sense, a perfectly predictive model where the ERG can flawlessly replicate the estimate of retinal ganglion cell soma counts is not the optimum outcome. If the same estimate of the state of glaucoma can be obtained by less invasive and complex means (ie Standard Automated Perimetry and OCT) then there is no predictive value in the ERG for this application.

On the other hand, if the ERG provides information which is only partially correlated with the SAP or OCT, then the possibility exists that it will in fact provide additional useful information. This cannot be determined from a cross sectional study like this one but would rather need a longitudinal study and the fitting of a model of what was known compared with what ultimately happened.

Although higher predictive performance was obtained by the time frequency models compared with markers alone and also by MARS compared with linear models, the higher complexity models are much less intuitive.

The importance of this work is therefore a demonstration that information relevant to the diagnosis and management of glaucoma exists within the ERG signal. Black box classifiers and regressors can be built using complex features and models but ultimately the clinical application will be limited as clinicians may not be able to understand the model or its recommendations.

Another limitation of this particular study is that although very good predictive performance was demonstrated for the dataset used in the study, it is not clear whether the results can be generalized to another dataset or the population in general.

The elegance of the approach of Meideros in defining an index of ganglion cell function from a combination of objective parameters is extended by this work. Models were developed and the predictions of ganglion cell counts correlate well with the numbers determined by the combination of perimetry and OCT. Importantly, all of the correlations have a trivial intercept so that if the Meideros model predicts zero ganglion cells then the ERG models would predict the same.

The limitations of this study include the relatively small sample size and the relatively small range of glaucoma severities. Further work could include a repeatability study to explore whether time-frequency features have test-retest repeatability in individual patients and how these compare to the repeatability of conventional ERG amplitude measurements.

This study was cross-sectional in nature and a valuable extension to this work would be to explore if the features derived from the ERG change over time with the progress or treatment of glaucoma in individual patients. As the aim of the glaucoma physician is to prevent the progression of glaucoma, these types of natural history studies are always difficult to do.

Finally, even better predictive performance might be attainable by a mixture of different types of features such as ERG amplitudes, latencies and various features extracted through different time frequency techniques. In that case though, intuition and interpretation would be even more difficult.

In summary, the experimental work has shown that severity of glaucoma as measured by a combined structure-function measure of ganglion cells can be modelled with conventional amplitude measurements from the ERG and that predictive performance can be further enhanced with time-frequency feature extraction. The work makes a compelling case that information about the severity of glaucoma exists within the electroretinogram and that further work may yield a way of processing the ERG to provide the glaucoma physician useful clinical information in the management of individual patients.

Chapter 5

Summary and Future Directions

5.1 Introduction

The key clinical decisions in glaucoma are the early diagnosis and the assessment of the adequacy of treatment. Currently, these decisions are made on the basis of structural and functional assessments of the eye – most importantly changes in standard automated perimetry (SAP) and in the thickness of the peripapillary retinal nerve fibre layer (RNFL) of the Optical Coherence Tomography (OCT) scan (Stein, Khawaja et al. 2021). The two measures have the potential to differ and in doing so uncover a structure-function gap that can indicate the potential for recovery (Ahmed, Waisbourd et al. 2021). The electroretinogram is an objective test of retinal function and has the potential to offer the glaucoma physician extra relevant clinical information in making decisions about the need for or adequacy of treatment. The Photopic Negative Response (PhNR) is a measure of ganglion cell activity (Viswanathan, Frishman et al. 2001) which is elicited without the need for high acuity or prolonged dark adaptation (Frishman, Sustar et al. 2018). Small portable devices are available that can measure the response without the need for a dedicated laboratory (Tang, Hui et al. 2018). The PhNR however does suffer from a perceived lack of repeatability (Tang, Edwards et al. 2014) and there have been no published methods to combine the electrophysiological and psychophysical measures of function in glaucoma. Addressing these issues with the ERG could lead to the wider adoption of the technique for the day-to-day management of patients with glaucoma.

5.2 Summary of findings

This thesis sought to improve the utility of the PhNR by improving its repeatability, exploring novel features that can be extracted from the signal and by demonstrating a link between the measure and an objective measure of glaucoma severity – the estimated retinal ganglion cell count (eRGC).

Chapter 2 examined adaptive filtering of the electroretinogram (ERG) signal. Of the methods explored to reduce baseline drift and wander, the complete ensemble empirical mode decomposition (CEEMDAN) was found to have the best performance. This technique has the additional benefit of being derived from an open-source C++ library, making implementation very simple, as the function can be called from a variety of programming environments and implemented in real time.

Chapter 3 then explored the use of features derived from information and chaos theory to classify eyes into glaucoma or non-glaucoma. The basis for the expected change in complexity was that glaucoma reduced the number of spiking ganglion cells and this should result in a reduction in the complexity of the ERG signal after the b-wave. The study showed that the Shannon entropy and the Higuchi fractal dimension were the most promising features when calculated on ensemble averages. No complexity features showed significant differences between eyes with and without glaucoma when calculated from individual sweeps and then averaged. This might be explained by the large amount of 50Hz noise present on each sweep when collected with skin electrodes and the portable ERG device. No complexity measure was superior to the PhNR in classifying the eyes into glaucoma and non-glaucoma.

Finally, **Chapter 4** used a variety of features derived from the ERG to model the severity of glaucoma using the eRGC as the outcome measure. In the time domain, using a full set of amplitude markers proved superior in modelling disease severity than using just the PhNR alone. This result was true regardless of model type or stimulus/background colour. Multi-adaptive regression splines (MARS) models performed significantly better than linear models when using the amplitude markers, but not when using the PhNR alone. Both chromatic and achromatic stimuli performed similarly. A variety of novel features based on time-frequency techniques were explored. Predictive models using features based on the discrete wavelet transform (DWT) showed the best balance between performance, explainability, and reproducibility. These models were superior to models based on time domain features alone and had R squared values of up to 0.57, compared to a 0.11 for a model using the PhNR alone.

5.3 Future directions

This thesis explored ways to improve the repeatability and utility of the ERG for the diagnosis and severity assessment of glaucoma. Perhaps one of the most important clinical decisions in glaucoma is the detection of progression. A cross-sectional study such as this one cannot determine whether the new techniques described will be helpful for this. Progression detection is vitally important because it indicates that the current treatment is inadequate and must be modified to obtain an intraocular pressure at which progression will not occur. Although the most rigorous measure of progression would be comparison of serial measures of severity, this requires higher levels of confidence in the severity measure than can currently be obtained. As such, many statistical models have been described that can detect glaucoma progression from features such as individual point sensitivities from the SAP or measures of RNFL, which do not require an explicit determination of glaucoma severity. Whether eRGC will be a suitable measure for detection of disease progression remains to be seen.

The utility of the PhNR and the other features described in this thesis could be assessed in their ability to detect progression in a longitudinal study. Because the glaucoma physician is always attempting to minimize or eliminate the risk of progression, such studies would require large numbers of participants and those participants would need to be followed for a prolonged period of many years.

The novel denoising techniques may have utility well beyond the electroretinogram and the photopic negative response. Figure 5-1 shows the application of CEEMDAN denoising to an electro-oculogram (EOG) sweep measured from the author. Dramatic improvement in the baseline stability can be seen.

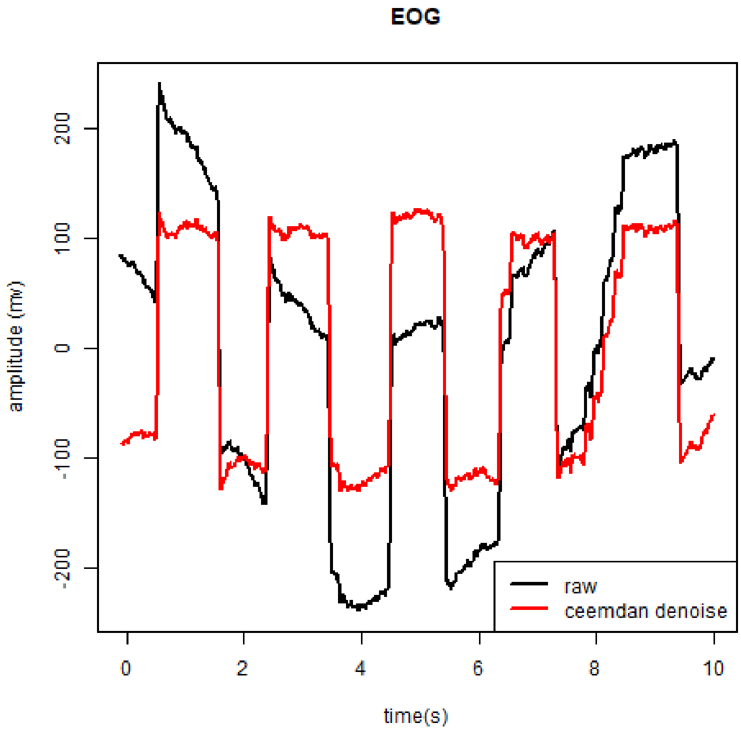


Figure 5-1: CEEMDAN denoising of the electro-oculogram

The rigorous optimization of mother wavelet selection for the DWT transform may also have applications outside of glaucoma when using the electroretinogram. The technique of energy to entropy ratio is ambivalent to the presence of glaucoma and features extracted by this technique could have utility for the detection or monitoring of other diseases such as toxic or inflammatory retinopathy.

5.4 Conclusion

In summary, this body of work has described a new, efficient, and effective method of removing baseline drift and wander from the ERG that might show utility in other areas of electrophysiology. A new approach to modelling glaucoma severity informed by time-frequency features suggests that there is much information about the underlying pathophysiology of the disease within the ERG. This may help the development of new algorithms to detect disease progression with the potential to save vision, reduce observation frequency or both.

References

Abásolo, D., et al. (2006). "Entropy analysis of the EEG background activity in Alzheimer's disease patients." Physiological measurement **27**(3): 241.

Abásolo, D., et al. (2006). "Analysis of EEG background activity in Alzheimer's disease patients with Lempel–Ziv complexity and central tendency measure." Medical engineering & physics **28**(4): 315-322.

Achanta, R., et al. (2012). "SLIC superpixels compared to state-of-the-art superpixel methods." IEEE transactions on pattern analysis and machine intelligence **34**(11): 2274-2282.

Adamo, F., et al. (2013). "A comparative study on mother wavelet selection in ultrasound image denoising." Measurement **46**(8): 2447-2456.

Akaiki, H. (1973). Information theory and an extension of the maximum likelihood principle. 2nd International Symposium on Information Theory. B. N. P. F. Csáki. Budapest, Akadémia Kiadó**:** 267-281.

Aldebasi, Y. H., et al. (2004). "S-cone, L+ M-cone, and pattern, electroretinograms in ocular hypertension and glaucoma." Vision research **44**(24): 2749-2756.

Alvarez-Ramirez, J., et al. (2009). "A DFA approach for assessing asymmetric correlations." Physica A: Statistical Mechanics and its Applications **388**(12): 2263-2270.

Andrade, A. O., et al. (2006). "EMG signal filtering based on empirical mode decomposition." Biomedical Signal Processing and Control **1**(1): 44-55.

Ang, G. S., et al. (2020). "Selective laser trabeculoplasty versus topical medication as initial glaucoma treatment: the glaucoma initial treatment study randomised clinical trial." British Journal of Ophthalmology **104**(6): 813-821.

Antos, A. and I. Kontoyiannis (2001). "Convergence properties of functional estimates for discrete distributions." Random Structures & Algorithms **19**(3-4): 163-193.

Arnalich-Montiel, F., et al. (2009). "Performance of glaucoma progression analysis software in a glaucoma population." Graefe's archive for clinical and experimental ophthalmology **247**(3): 391-397.

Arnold, M., et al. (1998). "Adaptive AR modeling of nonstationary time series by means of Kalman filtering." IEEE Transactions on Biomedical Engineering **45**(5): 553-562.

Artes, P. H., et al. (2014). "Visual field progression in glaucoma: what is the specificity of the guided progression analysis?" Ophthalmology **121**(10): 2023-2027.

Aschenbrandt, T. (1883). "Die physiologische Wirkung und Bedeutung des Cocain. muriat. auf den menschlichen Organismus." DMW-Deutsche Medizinische Wochenschrift **9**(50): 730-732.

Aulhorn, E. (1975). "Ueber die AUtomatisierung der perimetrischen Untersuchung. Die Wahl der Untersuchungsstrategie aus augenärtztlicher Sicht." Biomed Tech (Stuttg) **20**.

Aung, T., et al. (2001). "Acute primary angle-closure: long-term intraocular pressure outcome in Asian eyes." American journal of ophthalmology **131**(1): 7-12.

Awoyesuku, E. and A. Onua (2019). "Non-contact Tonometry Versus Pachymetry Corrected Intraocular Pressure: Any Difference? A Case for Pachymetry during Glaucoma Screening." Current Journal of Applied Science and Technology: 1-7.

Bach, M., et al. (2013). "ISCEV standard for clinical pattern electroretinography (PERG): 2012 update." Documenta Ophthalmologica **126**(1): 1-7.

Bach, M. and T. Meigen (1999). "Do's and don'ts in Fourier analysis of steady-state potentials." Documenta Ophthalmologica **99**(1): 69-82.

Bach, M., et al. (2006). "Pattern ERG as an early glaucoma indicator in ocular hypertension: a long-term, prospective study." Investigative ophthalmology & visual science **47**(11): 4881-4887.

Bajaj, V., et al. (2017). "Time-frequency image based features for classification of epileptic seizures from EEG signals." Biomedical Physics & Engineering Express **3**(1): 015012.

Bandt, C. and B. Pompe (2002). "Permutation entropy: a natural complexity measure for time series." Physical review letters **88**(17): 174102.

Banerjee, A., et al. (2019). "Comparison between broadband and monochromatic photopic negative response in full-field electroretinogram in controls and subjects with primary open-angle glaucoma." Documenta Ophthalmologica **138**(1): 21-33.

Banerjee, A., et al. (2021). "Comparison of broadband and monochromatic photopic negative response in eyes of patients with diabetes with no diabetic retinopathy and different stages of diabetic retinopathy." Indian Journal of Ophthalmology **69**(11): 3241.

Banister (1622). A Worthy Treatise of the Eyes. London.

Banitt, M. R., et al. (2013). "Progressive loss of retinal ganglion cell function precedes structural loss by several years in glaucoma suspects." Investigative ophthalmology & visual science **54**(3): 2346-2352.

Barlow, H. and W. Levick (1969). "Changes in the maintained discharge with adaptation level in the cat retina." The Journal of physiology **202**(3): 699-718.

Barnard, A. R., et al. (2011). "Specific deficits in visual electrophysiology in a mouse model of dominant optic atrophy." Experimental eye research **93**(5): 771-777.

Barraco, R., et al. (2011). "ERG signal analysis using wavelet transform." Theory in Biosciences **130**(3): 155-163.

Barthelme, S. (2019). imager: Image Processing Library Based on 'CImg'.

Bartlett, J. and C. Frost (2008). "Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables." Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology **31**(4): 466-475.

Basar, E. (1983). "Toward a physical approach to integrative physiology. I. Brain dynamics and physical causality." American Journal of Physiology-Regulatory, Integrative and Comparative Physiology **245**(4): R510-R533.

Bebie, H., et al. (1976). "Static perimetry: strategies." Acta ophthalmologica **54**(3): 325-338.

Beckerman, H., et al. (2001). "Smallest real difference, a link between reproducibility and responsiveness." Quality of Life Research **10**(7): 571-578.

Bengtsson, B., et al. (2005). "Effects of input data on the performance of a neural network in distinguishing normal and glaucomatous visual fields." Investigative ophthalmology & visual science **46**(10): 3730-3736.

Bengtsson, B. and A. Heijl (1998). "SITA Fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma." Acta Ophthalmologica Scandinavica **76**(4): 431-437.

Bengtsson, B. and A. Heijl (2008). "A visual field index for calculation of glaucoma rate of progression." American journal of ophthalmology **145**(2): 343-353.

Bengtsson, B., et al. (1997). "A new generation of algorithms for computerized threshold perimetry, SITA." Acta Ophthalmologica Scandinavica **75**(4): 368-375.

Benjamini, Y. and D. Yekutieli (2001). "The control of the false discovery rate in multiple testing under dependency." Annals of statistics: 1165-1188.

Berninger, T. A. and G. B. Arden (1988). "The pattern electroretinogram." Eye **2**: S257.

Berry, R. H., et al. (2015). "Synapse loss and dendrite remodeling in a mouse model of glaucoma." PloS one **10**(12): e0144341.

Bezerianos, A., et al. (2003). "Time-dependent entropy estimation of EEG rhythm changes following brain ischemia." Annals of biomedical engineering **31**(2): 221-232.

Binns, A. M., et al. (2011). "The relationship between stimulus intensity and response amplitude for the photopic negative response of the flash electroretinogram." Documenta Ophthalmologica **122**(1): 39-52.

Blackman, R. B. and J. W. Tukey (1958). "The measurement of power spectra from the point of view of communications engineering—Part I." Bell System Technical Journal **37**(1): 185-282.

Blanco-Velasco, M., et al. (2008). "ECG signal denoising and baseline wander correction based on the empirical mode decomposition." Computers in biology and medicine **38**(1): 1-13.

Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement." lancet **1**(8476): 307-310.

Bois-Reymond, D. (1849). Untersuchungen uber theirische Electricitaet. Berlin, G. Reimer.

Bolz, S., et al. (2008). "K+ currents fail to change in reactive retinal glial cells in a mouse model of glaucoma." Graefe's archive for clinical and experimental ophthalmology **246**(9): 1249-1254.

Bornas, X., et al. (2006). "Sample entropy of ECG time series of fearful flyers: preliminary results." Nonlinear dynamics, psychology, and life sciences **10**(3): 301-318.

Bowd, C., et al. (2012). "Predicting glaucomatous progression in glaucoma suspect eyes using relevance vector machine classifiers for combined structural and functional measurements." Investigative ophthalmology & visual science **53**(4): 2382-2389.

Bretschneider, F. and J. R. De Weille (2018). Introduction to electrophysiological methods and instrumentation, Academic Press.

Brisseau, M. (1709). Traité de la cataracte et du glaucoma. Paris.

Bruhn, J., et al. (2001). "Shannon entropy applied to the measurement of the electroencephalographic effects of desflurane." Anesthesiology: The Journal of the American Society of Anesthesiologists **95**(1): 30-35.

Brulé, J., et al. (2007). "Evidence of a possible impact of the menstrual cycle on the reproducibility of scotopic ERGs in women." Documenta Ophthalmologica **114**(3): 125-134.

Buckley, S., et al. (1994). "Acute angle closure glaucoma: relative failure of YAG iridotomy in affected eyes and factors influencing outcome." British Journal of Ophthalmology **78**(7): 529-533.

Bui, B. V., et al. (2013). "Relationship between the magnitude of intraocular pressure during an episode of acute elevation and retinal damage four weeks later in rats." PloS one **8**(7): e70513.

Butterworth, S. (1930). "On the theory of filter amplifiers." Wireless Engineer **7**(6): 536-541.

Calkins, D. J. (2021). "Adaptive responses to neurodegenerative stress in glaucoma." Progress in retinal and eye research: 100953.

Cantor, G. (1883). Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Leipzig, B.G. Teubner.

Carvalho, V. R., et al. (2020). "Evaluating five different adaptive decomposition methods for EEG signal seizure detection and classification." Biomedical Signal Processing and Control **62**: 102073.

Castiglioni, P. (2010). "Letter to the Editor: What is wrong in Katz's method? Comments on: A note on fractal dimensions of biomedical waveforms." Computers in biology and medicine **40**(11-12): 950-952.

Cerutti, S., et al. (1987). "Analysis of visual evoked potentials through Wiener filtering applied to a small number of sweeps." Journal of biomedical engineering **9**(1): 3-12.

Cervantes-De la Torre, F., et al. (2013). Fractal dimension algorithms and their application to time series associated with natural phenomena. Journal of Physics: Conference Series, IOP Publishing.

Chang, K.-M. (2010). "Arrhythmia ECG noise reduction by ensemble empirical mode decomposition." Sensors **10**(6): 6063-6080.

Chauhan, B. C. and C. F. Burgoyne (2013). "From clinical examination of the optic disc to clinical assessment of the optic nerve head: a paradigm change." American journal of ophthalmology **156**(2): 218-227. e212.

Chauhan, B. C., et al. (2015). "Bruch's membrane opening minimum rim width and retinal nerve fiber layer thickness in a normal white population: a multicenter study." Ophthalmology **122**(9): 1786-1794.

Chavan, M. S., et al. (2011). Implementation of SYMLET wavelets to removal of Gaussian additive noise from speech signal. Proceedings of Recent Researches in Communications, Automation, Signal Processing, Nanotechnology, Astronomy and Nuclear Physics: 10th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications (EHAC’11), Cambridge.

Chen, H., et al. (2008). "The photopic negative response of flash ERG in nonproliferative diabetic retinopathy." Documenta Ophthalmologica **117**(2): 129-135.

Chen, W.-t., et al. (2006). "Characterization of surface EMG signals using improved approximate entropy." Journal of Zhejiang University Science B **7**(10): 844-848.

Chen, W., et al. (2007). "Characterization of surface EMG signal based on fuzzy entropy." IEEE Transactions on neural systems and rehabilitation engineering **15**(2): 266-272.

Chen, Z. and X. Zhang (2014). "Noise and vibration assessment of permanent-magnet synchronous motors based on matching pursuit." Journal of Vibroengineering **16**(4): 1831-1841.

Cheng, H., et al. (2007). "The relationship between visual field and retinal nerve fiber layer measurements in patients with multiple sclerosis." Investigative ophthalmology & visual science **48**(12): 5798-5805.

Chiu, S. J., et al. (2010). "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation." Optics express **18**(18): 19413-19428.

Chrysostomou, V. and J. G. Crowston (2013). "The photopic negative response of the mouse electroretinogram: reduction by acute elevation of intraocular pressure." Investigative ophthalmology & visual science **54**(7): 4691-4697.

Chrysostomou, V., et al. (2010). "Mechanisms of retinal ganglion cell injury in aging and glaucoma." Ophthalmic Research **44**(3): 173-178.

Chu, S., et al. (2009). "Environmental sound recognition with time–frequency audio features." IEEE Transactions on Audio, Speech, and Language Processing **17**(6): 1142-1158.

Chui, C. K. and M. D. van der Walt (2015). "Signal analysis via instantaneous frequency estimation of signal components." GEM-International Journal on Geomathematics **6**(1): 1-42.

Clausius, R. (1856). "X. On a modified form of the second fundamental theorem in the mechanical theory of heat." The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **12**(77): 81-98.

Cocchiarella, L., et al. (2001). Guides to the Evaluation of Permanent Impairment, American Medical Association.

Cohen, A., et al. (1992). "Biorthogonal bases of compactly supported wavelets." Communications on pure and applied mathematics **45**(5): 485-560.

Cohen, B. A. and A. Sances (1977). "Stationarity of the human electroencephalogram." Medical and Biological Engineering and Computing **15**(5): 513-518.

Cohen, L. (1989). "Time-frequency distributions-a review." Proceedings of the IEEE **77**(7): 941-981.

Cohen, W. (1957). "Spatial and textural characteristics of the Ganzfeld." The American journal of psychology **70**(3): 403-410.

Coifman, R. R. and M. V. Wickerhauser (1992). "Entropy-based algorithms for best basis selection." IEEE Transactions on Information theory **38**(2): 713-718.

Colominas, M. A., et al. (2014). "Improved complete ensemble EMD: A suitable tool for biomedical signal processing." Biomedical Signal Processing and Control **14**: 19-29.

Colotto, A., et al. (2000). "Photopic negative response of the human ERG: losses associated with glaucomatous damage." Investigative ophthalmology & visual science **41**(8): 2205-2211.

Cooley, J. W. and J. W. Tukey (1965). "An algorithm for the machine calculation of complex Fourier series." Mathematics of computation **19**(90): 297-301.

Cooper, E. A., et al. (2013). "Assessment of OLED displays for vision research." Journal of vision **13**(12): 16-16.

Cubbidge, R. and J. Wild (2001). "The influences of stimulus wavelength and eccentricity on short-wavelength pathway isolation in automated perimetry." Ophthalmic and Physiological Optics **21**(1): 1-8.

Cuenca, Á., et al. (2019). "A remote control strategy for an autonomous vehicle with slow sensor using kalman filtering and dual-rate control." Sensors **19**(13): 2983.

Cugini, P., et al. (2001). "Is a reduced entropy in heart rate variability an early finding of silent cardiac neurovegetative dysautonomia in type 2 diabetes mellitus?" Journal of Clinical and Basic Cardiology **4**(4): 289-294.

Cursiefen, C., et al. (2001). "The negative response of the flash electroretinogram in glaucoma." Documenta Ophthalmologica **103**(1): 1-12.

Cvenkel, B., et al. (2017). "Ganglion cell loss in early glaucoma, as assessed by photopic negative response, pattern electroretinogram, and spectral-domain optical coherence tomography." Documenta Ophthalmologica **135**(1): 17-28.

Cvenkel, B., et al. (2022). "Monitoring for glaucoma progression with SAP, electroretinography (PERG and PhNR) and OCT." Documenta Ophthalmologica **144**(1): 17-30.

Damji, K. F., et al. (1999). "Selective laser trabeculoplasty vargon laser trabeculoplasty: a prospective randomised clinical trial." British Journal of Ophthalmology **83**(6): 718-722.

Daniel, C. (1959). "Use of half-normal plots in interpreting factorial two-level experiments." Technometrics **1**(4): 311-341.

Daubechies, I. (1988). "Orthonormal bases of compactly supported wavelets." Communications on pure and applied mathematics **41**(7): 909-996.

Davila, C. E. and M. S. Mobin (1992). "Weighted averaging of evoked potentials." IEEE Transactions on Biomedical Engineering **39**(4): 338-345.

Dawson, W. W., et al. (1979). "Improved electrode for electroretinography." **18**(9): 988-991.

De Boer, J. F., et al. (2017). "Twenty-five years of optical coherence tomography: the paradigm shift in sensitivity and speed provided by Fourier domain OCT." Biomedical optics express **8**(7): 3248-3280.

Deák, K., et al. (2017). "Optimal Wavelet Selection for the Size Estimation of Manufacturing Defects of Tapered Roller Bearings with Vibration Measurement using Shannon Entropy Criteria." Strojniski Vestnik/Journal of Mechanical Engineering **63**(1).

Deering, R. and J. F. Kaiser (2005). The use of a masking signal to improve empirical mode decomposition. Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on, IEEE.

Delgado, M. F., et al. (2019). "Management of glaucoma in developing countries: challenges and opportunities for improvement." ClinicoEconomics and outcomes research: CEOR **11**: 591.

Della Santina, L., et al. (2013). "Differential progression of structural and functional alterations in distinct retinal ganglion cell types in a mouse model of glaucoma." Journal of Neuroscience **33**(44): 17444-17457.

Demours (1818). Traité des maladies des yeux, avec planches coloriées représentant ces maladies d'après natures suivi de la description de l'Oeil Humain. Paris, L'Auteur & Crochard.

Dewar, J. (1877). "The Physiological Action of Light1." Nature **15**: 433.

Donders, F. (1855). "Ueber die sichtbaren Erscheinungen der Blutbewegung im Auge." Archiv für Ophthalmologie **2**(2): 75-105.

Donoho, D. L., et al. (2012). "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit." IEEE Transactions on Information theory **58**(2): 1094-1121.

Downie, T. and B. Silverman (1998). "The discrete multiple wavelet transform and thresholding methods." IEEE transactions on signal processing **46**(9): 2558-2561.

Dragomiretskiy, K. and D. Zosso (2013). "Variational mode decomposition." IEEE transactions on signal processing **62**(3): 531-544.

Drance, S. M. (1972). "The glaucomatous visual field." The British journal of ophthalmology **56**(3): 186.

Drasdo, N., et al. (2001). "The s-cone PHNR and pattern ERG in primary open angle glaucoma." Investigative ophthalmology \& visual science **42**(6): 1266-1272.

Drasdo, N., et al. (2007). "The length of Henle fibers in the human retina and a model of ganglion receptive field density in the visual field." Vision research **47**(22): 2901-2911.

Edmunds, B., et al. (2001). "The National Survey of Trabeculectomy. II. Variations in operative technique and outcome." Eye **15**(4): 441-448.

Elbendary, A. M., et al. (2017). "Diagnostic accuracy of ganglion cell complex substructures in different stages of primary open-angle glaucoma." Canadian Journal of Ophthalmology **52**(4): 355-360.

Esgiar, A. N., et al. (2002). "Fractal analysis in the detection of colonic cancer images." IEEE Transactions on information Technology in Biomedicine **6**(1): 54-58.

Euler, L. (1748). Introductio in analysin infinitorum, MM Bousquet.

Evangelho, K., et al. (2019). "Pathophysiology of primary open-angle glaucoma from a neuroinflammatory and neurotoxicity perspective: a review of the literature." International ophthalmology **39**(1): 259-271.

Fang, L., et al. (2017). "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search." Biomedical optics express **8**(5): 2732-2744.

Fankhauser, F., et al. (1977). "Some aspects of the automation of perimetry." Survey of ophthalmology **22**(2): 131-141.

Fernández, A., et al. (2013). "Complexity and schizophrenia." Progress in Neuro-Psychopharmacology and Biological Psychiatry **45**: 267-276.

Ferry, N. H. a. M. (2017). vmd: Variational Mode Decomposition.

Finkelstein, D., et al. (1968). "Human electroretinogram near the absolute threshold of vision." Investigative ophthalmology & visual science **7**(2): 214-218.

Flood, M. W. and B. Grimm (2021). "EntropyHub: An open-source toolkit for entropic time series analysis." PloS one **16**(11): e0259448.

Folgar, F. A., et al. (2014). "Lateral and axial measurement differences between spectral-domain optical coherence tomography systems." Journal of Biomedical Optics **19**(1): 016014.

Forte, J. D., et al. (2008). "Wavelet analysis reveals dynamics of rat oscillatory potentials." Journal of neuroscience methods **169**(1): 191-200.

Foster, P. J., et al. (2002). "The definition and classification of glaucoma in prevalence surveys." British Journal of Ophthalmology **86**(2): 238-242.

Fourier, J. (1822). Theorie analytique de la chaleur, par M. Fourier, Chez Firmin Didot, père et fils.

Franaszczuk, P. J., et al. (1998). "Time–frequency analysis using the matching pursuit algorithm applied to seizures originating from the mesial temporal lobe." Electroencephalography and clinical neurophysiology **106**(6): 513-521.

Fredette, M.-J., et al. (2008). "Reproducibility of pattern electroretinogram in glaucoma patients with a range of severity of disease with the new glaucoma paradigm." Ophthalmology **115**(6): 957-963.

Friedman, J. H. (1991). "Multivariate adaptive regression splines." The annals of statistics: 1-67.

Frishman, L., et al. (2018). "ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram." Documenta Ophthalmologica: 1-5.

Frishman, L., et al. (2001). The evolution of the photopic negative response and possible clinical applications. Vision Science and its Applications, Optica Publishing Group.

Fry, L. E., et al. (2018). "The coma in glaucoma: retinal ganglion cell dysfunction and recovery." Progress in retinal and eye research.

Gabor, D. (1946). "Theory of communication. Part 1: The analysis of information." Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering **93**(26): 429-441.

Gaffney, A. J., et al. (2014). "Measurement of cone dark adaptation: a comparison of four psychophysical methods." Documenta Ophthalmologica **128**(1): 33-41.

Garcia, G.-G. P., et al. (2019). "Accuracy of Kalman filtering in forecasting visual field and intraocular pressure trajectory in patients with ocular hypertension." JAMA ophthalmology **137**(12): 1416-1423.

Garcia, G.-G. P., et al. (2019). "Using Kalman filtering to forecast disease trajectory for patients with normal tension glaucoma." American journal of ophthalmology **199**: 111-119.

Gardiner, S. K., et al. (2014). "A method to estimate the amount of neuroretinal rim tissue in glaucoma: comparison with current methods for measuring rim area." American journal of ophthalmology **157**(3): 540-549. e542.

Gauvin, M. (2017). TIME-FREQUENCY ANALYSIS OF THE HUMAN PHOTOPIC ELECTRORETINOGRAM: METHOD, NORMATIVE DATA AND, McGill University.

Gauvin, M., et al. (2015). "Functional decomposition of the human ERG based on the discrete wavelet transform." Journal of vision **15**(16): 14-14.

Ghofrani, S., et al. (2003). Comparing Gaussian and chirplet dictionaries for time-frequency analysis using matching pursuit decomposition. Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No. 03EX795), IEEE.

Giangiacomo, A., et al. (2006). "Diagnosing glaucoma progression: current practice and promising technologies." Current opinion in ophthalmology **17**(2): 153-162.

Gibbs, J. W. (1948). Elementary Principles in Statistical Mechanics. New Haven, Yale University Press.

Gilles, J. (2013). "Empirical wavelet transform." IEEE transactions on signal processing **61**(16): 3999-4010.

Gilles, J. (2021). Empirical Wavelet Transforms, Matlab Central File Exchange.

Gitter, J. A. and M. J. Czerniecki (1995). "Fractal analysis of the electromyographic interference pattern." Journal of neuroscience methods **58**(1-2): 103-108.

Glass, E., et al. (1995). "Simulations for FASTPAC and the standard 4-2 dB full-threshold strategy of the Humphrey Field Analyzer." Investigative ophthalmology & visual science **36**(9): 1847-1854.

Glenny, R. W., et al. (1991). "Applications of fractal analysis to physiology." Journal of Applied Physiology **70**(6): 2351-2367.

Gneiting, T., et al. (2012). "Estimators of fractal dimension: Assessing the roughness of time series and spatial data." Statistical Science: 247-277.

Goldmann, H. (1955). "Un nouveau tonometre d'applanation." Bull Soc Ophtalmol Fr **67**: 474-478.

Goldmann, H. and T. Schmidt (1957). "Über applanationstonometrie." Ophthalmologica **134**(4): 221-242.

Goodwin, M. M. and M. Vetterli (1999). "Matching pursuit and atomic signal models based on recursive filter banks." IEEE transactions on signal processing **47**(7): 1890-1902.

Gotch, F. (1903). "The time relations of the photo-electric changes in the eyeball of the frog." The Journal of physiology **29**(4-5): 388-410.

Gotoh, Y., et al. (2004). "Selective Loss of the Photopic Negative Response in Patients With OpticNerve Atrophy." Archives of ophthalmology **122**(3): 341-346.

Goupillaud, P., et al. (1984). "Cycle-octave and related transforms in seismic signal analysis." Geoexploration **23**(1): 85-102.

Gouras, P. and R. D. Gunkel (1962). "The resonant frequencies of rod and cone electroretinograms." Investigative ophthalmology & visual science **1**(1): 122-126.

Gouras, P. and C. Mackay (1988). "DETECTING EARLY POSTCHIASMATIC VISUALLY EVOKED-RESPONSES." Clinical vision sciences **3**(2): 119-124.

Granit, R. (1933). "The components of the retinal action potential in mammals and their relation to the discharge in the optic nerve." The Journal of physiology **77**(3): 207-239.

Grassberger, P. and I. Procaccia (1983). "Estimation of the Kolmogorov entropy from a chaotic signal." Physical review A **28**(4): 2591.

Groneberg, A. and C. Teping (1980). Topodiagnostik von Sehstörungen durch Ableitung retinaler und kortikaler Antworten auf Umkehr-Kontrastmuster. Plastische Chirurgie der Lider und Chirurgie der Tränenwege, Springer**:** 409-415.

Guerrero-Mosquera, C., et al. (2010). "New feature extraction approach for epileptic EEG signal detection using time-frequency distributions." Medical & biological engineering & computing **48**(4): 321-330.

Guillemeau, J. (1585). Traité des maladies de l'oeil : qui sont en nombre de cent treize, ausquelles il est suject. Paris, Chez Charles Massé, au Clos Bruneau à l'enseigne de la Pyramide.

Guo, L., et al. (2010). "Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks." Journal of neuroscience methods **193**(1): 156-163.

Guo, X., et al. (2016). "Caloric restriction promotes cell survival in a mouse model of normal tension glaucoma." Scientific Reports **6**(1): 1-10.

Gur, M. and Y. Zeevi (1980). "Frequency-domain analysis of the human electroretinogram." JOSA **70**(1): 53-59.

Ha, A., et al. (2019). "Optic Disc Microhemorrhage in Primary Open-Angle Glaucoma: Clinical Implications for Visual Field Progression." Investigative ophthalmology & visual science **60**(6): 1824-1832.

Haffmanns, J. (1862). "Beitrage zur Kenntniss des Glaukoms. III. Resultate. Aus dem Hollandischen deutsch bearbeitet von M Schmidt." Arch Ophthalmol **8**: 143-178.

Hara, Y., et al. (2020). "Comparisons of photopic negative responses elicited by different conditions from glaucomatous eyes." Japanese journal of ophthalmology **64**(2): 114-126.

Haralick, R. M., et al. (1973). "Textural features for image classification." IEEE Transactions on Systems, man, and cybernetics(6): 610-621.

Hartline, H. K. (1925). "The electrical response to illumination of the eye in intact animals, including the human subject; and in decerebrate preparations." American Journal of Physiology-Legacy Content **73**(3): 600-612.

Harwerth, R., et al. (2010). "Linking structure and function in glaucoma." Progress in retinal and eye research **29**(4): 249-271.

Hausdorff, F. (1918). "Dimension und ausseres Mass." Math. Ann. Mathematische Annalen **79**(1-2): 157-179.

Hausser, J., Strimmer, K. (2014). entropy: Estimation of Entropy, Mutual Information and Related Quantities.

Havlicek, L. L. and N. L. Peterson (1976). "Robustness of the Pearson correlation against violations of assumptions." Perceptual and Motor Skills **43**(3_suppl): 1319-1334.

Hébert, M. and P. Lachapelle (1999). "Reproducibility of ERG responses obtained with the DTL electrode." Vision research **39**(6): 1069-1070.

Heijl, A., et al. (1987). A package for the statistical analysis of visual fields. Seventh International Visual Field Symposium, Amsterdam, September 1986, Springer.

Hellmuth, T. and M. Welle (1998). "Simultaneous measurement of dispersion, spectrum, and distance with a fourier transform spectrometer." Journal of Biomedical Optics **3**(1): 7-12.

Helmholtz, H. (1851). Beschreibung eines Augen-Spiegels zur Untersuchung der Netzhaut in lebenden Auge. Berlin, A Forstner.

Helmholtz, H. v. (1884). "Principien der Statik monocyclischer Systeme."

Helske, J. L., Perttu (2017). Rlibeemd : Ensemble empirical mode decomposition (EEMD) and its complete variant (CEEMDAN).

Hess, R. F. and C. L. Baker Jr (1984). "Human pattern-evoked electroretinogram." Journal of Neurophysiology **51**(5): 939-951.

Higuchi, T. (1990). "Relationship between the fractal dimension and the power law index for a time series: a numerical investigation." Physica D: Nonlinear Phenomena **46**(2): 254-264.

Hippocrates, A. F. (2016). "Aphorisms."

Hirata, T. (1989). "A correlation between the b value and the fractal dimension of earthquakes." Journal of Geophysical Research: Solid Earth **94**(B6): 7507-7514.

Hlawatsch, F. and F. Auger (2013). Time-frequency analysis, John Wiley & Sons.

Holder, G. E. (2001). "Pattern electroretinography (PERG) and an integrated approach to visual pathway diagnosis." Progress in retinal and eye research **20**(4): 531-561.

Holmgren, F. (1870). "Om Retinaströmmen, Upsala Läkaref." Förh **4**: 419-455.

Holzinger, A., et al. (2012). On applying approximate entropy to ECG signals for knowledge discovery on the example of big sensor data. International conference on active media technology, Springer.

Hood, D. and J. Li (1997). "A technique for measuring individual multifocal ERG records." Trends Opt Photon **11**: 280-293.

Hood, D. C., et al. (2007). "Structure versus function in glaucoma: an application of a linear model." Investigative ophthalmology & visual science **48**(8): 3662-3668.

Hooge, F. (1976). "1/f noise." Physica B+ C **83**(1): 14-23.

Hsieh, M.-H., et al. (2020). "Fourier analysis of circumpapillary retinal nerve fiber layer thickness in optical coherence tomography for differentiating myopia and glaucoma." Scientific Reports **10**(1): 1-9.

Hu, Y., et al. (2001). "Comparison of time-frequency distribution techniques for analysis of spinal somatosensory evoked potential." Medical and Biological Engineering and Computing **39**(3): 375-380.

Huang, D., et al. (1991). "Optical coherence tomography." Science **254**(5035): 1178-1181.

Huang, N. E., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, The Royal Society.

Hui, F., et al. (2020). "Improvement in inner retinal function in glaucoma with nicotinamide (vitamin B3) supplementation: A crossover randomized clinical trial." Clinical & experimental ophthalmology **48**(7): 903-914.

Ishikawa, H., et al. (2005). "Macular segmentation with optical coherence tomography." Investigative ophthalmology & visual science **46**(6): 2012-2017.

Iyer, J., et al. (2020). "Toward a new definition of glaucomatous optic neuropathy for clinical research." Current opinion in ophthalmology **31**(2): 85-90.

Iyer, J. V., et al. (2020). "Defining glaucomatous optic neuropathy using objective criteria from structural and functional testing." British Journal of Ophthalmology.

Jain, A. K., et al. (1999). "Data clustering: a review." ACM computing surveys (CSUR) **31**(3): 264-323.

Johnson, J. B. (1928). "Thermal agitation of electricity in conductors." Physical review **32**(1): 97.

Johnson, M. A. and R. W. Massof (1982). "The photomyoclonic reflex: an artefact in the clinical electroretinogram." British Journal of Ophthalmology **66**(6): 368-378.

Jones, M. N. and D. J. Mewhort (2004). "Case-sensitive letter and bigram frequency counts from large-scale English corpora." Behavior research methods, instruments, & computers **36**(3): 388-396.

Julier, S. J. and J. K. Uhlmann (1997). New extension of the Kalman filter to nonlinear systems. Signal processing, sensor fusion, and target recognition VI, International Society for Optics and Photonics.

Jung, K. I., et al. (2020). "Pattern electroretinograms in preperimetric and perimetric glaucoma." American journal of ophthalmology **215**: 118-126.

Kadir, R., et al. (2021). "LDPC coded hybrid discrete cosine transform and Fejér–Korovkin wavelet transform-based SC-FDMA for image communication." Array **12**: 100107.

Kalesnykas, G., et al. (2012). "Retinal ganglion cell morphology after optic nerve crush and experimental glaucoma." Investigative ophthalmology & visual science **53**(7): 3847-3857.

Kalman, R. E. (1960). "A new approach to linear filtering and prediction problems." Journal of basic Engineering **82**(1): 35-45.

Karlberg, B., et al. (1968). "Electroretinography during short-term intraocular tension rise." Acta ophthalmologica **46**(4): 742-748.

Karmakar, C., et al. (2017). "Stability, consistency and performance of distribution entropy in analysing short length heart rate variability (HRV) signal." Frontiers in physiology **8**: 720.

Katz, M. J. (1988). "Fractals and the analysis of waveforms." Computers in biology and medicine **18**(3): 145-156.

Keel, S., et al. (2018). "Prevalence of glaucoma in the Australian National Eye Health Survey." British Journal of Ophthalmology: bjophthalmol-2017-311786.

Kelvin, W. T. B. (1890). Mathematical and physical papers, University Press.

Kerr, J. F., et al. (1972). "Apoptosis: a basic biological phenomenon with wideranging implications in tissue kinetics." British journal of cancer **26**(4): 239-257.

Kerrigan–Baumrind, L. A., et al. (2000). "Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons." Investigative ophthalmology & visual science **41**(3): 741-748.

Kerschensteiner, D. and R. O. Wong (2008). "A precisely timed asynchronous pattern of ON and OFF retinal ganglion cell activity during propagation of retinal waves." Neuron **58**(6): 851-858.

Kielczewski, J. L., et al. (2005). "The effect of experimental glaucoma and optic nerve transection on amacrine cells in the rat retina." Investigative ophthalmology & visual science **46**(9): 3188-3196.

Kim, Y., et al. (2015). "The instability of the Pearson correlation coefficient in the presence of coincidental outliers." Finance Research Letters **13**: 243-257.

Kirkiewicz, M., et al. (2016). "Photopic negative response of full-field electroretinography in patients with different stages of glaucomatous optic neuropathy." Documenta Ophthalmologica **132**(1): 57-65.

Kita, Y., et al. (2020). "RETeval Portable Electroretinogram Parameters in Different Severity Stages of Glaucoma." Journal of glaucoma **29**(7): 572-580.

Klistorner, A. I. and S. L. Graham (1999). "Early magnocellular loss in glaucoma demonstrated using the pseudorandomly stimulated flash visual evoked potential." Journal of glaucoma **8**(2): 140-148.

Koch, K., et al. (2004). "Efficiency of information transmission by retinal ganglion cells." Current Biology **14**(17): 1523-1530.

Koller, C. (1884). "On the use of cocaine for producing anaesthesia on the eye." The Lancet **124**(3197): 990-992.

Koller, C. (1884). Ueber die Verwundung des cocain zur anästhesirung am auge.

Kolmogorov, A. N. (1985). "A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces." Trudy Matematicheskogo Instituta imeni VA Steklova **169**: 94-98.

Kreuz, A. C., et al. (2018). "Macular and multifocal PERG and FD-OCT in preperimetric and hemifield loss glaucoma." Journal of glaucoma **27**(2): 121-132.

Kuhn, M. (2017). caret: Classification and Regression Training.

Kumar, D., et al. (2018). "An efficient computational technique for fractal vehicular traffic flow." Entropy **20**(4): 259.

Kumar, S., et al. (2021). "Neuroprotection of the inner retina also prevents secondary outer retinal pathology in a mouse model of glaucoma." Investigative ophthalmology & visual science **62**(9): 35-35.

Laba, M., et al. (2007). "Influence of wavelet type on the classification of marsh vegetation from satellite imagery using a combination of wavelet texture and statistical component analyses." Canadian Journal of Remote Sensing **33**(4): 260-265.

Lang, A. and A. Puusa (1981). "Dual influence of temperature on compound nerve action potential." Journal of the Neurological Sciences **51**(1): 81-88.

LaRocca, F., et al. (2011). "Robust automatic segmentation of corneal layer boundaries in SDOCT images using graph theory and dynamic programming." Biomedical optics express **2**(6): 1524-1538.

Le, J. T., et al. (2019). "Ab interno trabecular bypass surgery with iStent for open-angle glaucoma." Cochrane Database of Systematic Reviews(3).

Lee, A., et al. (2019). "Case report: vitamin A deficiency and nyctalopia in a patient with chronic pancreatitis." Optometry and Vision Science **96**(6): 453-458.

Lee, B., et al. (2010). "Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry." Physiological measurement **31**(12): 1585.

Lee, M. H., et al. (2015). "Vitamin A Deficiency Presenting with 'Itchy Eyes'." Case reports in ophthalmology **6**(3): 427-434.

Lee, M. H. B. M., Sarossy M. G. (2013). Skewness of the distribution of normal ERG values and implications for calculation of the normal range. 51st ISCEV Symposium. Chonqing China, Springer. **127:** 13.

Lee, M. H. S. M. G. W. Z. T. J. (2016). Empirical mode decomposition denoising of the ERG to improve teh reliability of the photopic negative response. 54th ISCEV Symposium. Singapore, Springer. **133:** 22-21.

Leone, F. C., et al. (1961). "The folded normal distribution." Technometrics **3**(4): 543-550.

Levick, W., et al. (1983). "Performance of cat retinal ganglion cells at low light levels." The Journal of general physiology **82**(3): 405-426.

Lewis, F. L., et al. (2017). Optimal and robust estimation: with an introduction to stochastic control theory, CRC press.

Leydhecker, G. (1950). "The electroretinogram in glaucomatous eyes." The British journal of ophthalmology **34**(9): 550.

Li, B., et al. (2005). "The decline of the photopic negative response (PhNR) in the rat after optic nerve transection." Documenta Ophthalmologica **111**(1): 23-31.

Li, J. W., et al. (2019). "Brain rhythm sequencing using EEG signals: A case study on seizure detection." IEEE access **7**: 160112-160124.

Li, P., et al. (2015). "Assessing the complexity of short-term heartbeat interval series by distribution entropy." Medical & biological engineering & computing **53**(1): 77-87.

Li, Y., et al. (2015). "An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis." Measurement **63**: 330-345.

Lim, J. H., et al. (2013). "Detrended fluctuation analysis and Kolmogorov–Sinai entropy of electroencephalogram signals." Physics letters A **377**(38): 2542-2545.

Liu, H. H., et al. (2017). "Reversal of functional loss in a rat model of chronic intraocular pressure elevation." Ophthalmic and Physiological Optics **37**(1): 71-81.

Liu, M., et al. (2011). "Dendritic changes in visual pathways in glaucoma and other neurodegenerative conditions." Experimental eye research **92**(4): 244-250.

Liu, Q., et al. (2004). "Size of the dictionary in matching pursuit algorithm." IEEE transactions on signal processing **52**(12): 3403-3408.

Luna, C., et al. (2009). "Resveratrol prevents the expression of glaucoma markers induced by chronic oxidative stress in trabecular meshwork cells." Food and Chemical Toxicology **47**(1): 198-204.

Luo, X., et al. (2017). "Tetramethylpyrazine nitrone protects retinal ganglion cells against N-methyl-d-aspartate-induced excitotoxicity." Journal of neurochemistry **141**(3): 373-386.

Luukko, P., et al. (2016). "Introducing libeemd: A program package for performing the ensemble empirical mode decomposition." Computational Statistics **31**(2): 545-557.

Ma, Y., et al. (2018). "Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches." Sleep medicine reviews **37**: 85-93.

Machida, S., et al. (2004). "Predominant loss of the photopic negative response in central retinal artery occlusion." American journal of ophthalmology **137**(5): 938-940.

Machida, S., et al. (2008). "Correlation between photopic negative response and retinal nerve fiber layer thickness and optic disc topography in glaucomatous eyes." Investigative ophthalmology \& visual science **49**(5): 2201-2207.

Machida, S., et al. (2015). "Regional variations in correlation between photopic negative response of focal electoretinograms and ganglion cell complex in glaucoma." Current eye research **40**(4): 439-449.

Mackenzie, W. (1854). A Practical Treatise on the Diseases of the Eye, Longman, Brown, Green & Longmans.

Maetschke, S., et al. (2019). "A feature agnostic approach for glaucoma detection in OCT volumes." PloS one **14**(7): e0219126.

Maklakoff, C. L. (1885). "Ophthalmotonometrie." Arch Ophthalmol (Paris) **5**: 159.

Mallat, S. and Z. Zhang (1993). Matching pursuit with time-frequency dictionaries, Courant Institute of Mathematical Sciences New York United States.

Mallat, S. G. (1989). "A theory for multiresolution signal decomposition: the wavelet representation." IEEE transactions on pattern analysis and machine intelligence **11**(7): 674-693.

Mandelbrot, B. (1967). "How long is the coast of Britain? Statistical self-similarity and fractional dimension." Science **156**(3775): 636-638.

Mandelbrot, B. B. (1983). The fractal geometry of nature. New York, Freeman.

Marmor, M. F., et al. (1989). "Standard for clinical electroretinography." Documenta Ophthalmologica **73**(4): 303-311.

May, J. G., et al. (1982). "Loss in pattern-elicited electroretinograms in optic nerve dysfunction." American journal of ophthalmology **93**(4): 418-422.

McCulloch, D. L., et al. (2015). "ISCEV Standard for full-field clinical electroretinography (2015 update)." Documenta Ophthalmologica **130**(1): 1-12.

McKinnon, S. J. (1997). "Glaucoma, apoptosis, and neuroprotection." Current opinion in ophthalmology **8**(2): 28-37.

Medeiros, F. A., et al. (2019). "From machine to machine: an OCT-trained deep learning algorithm for objective quantification of glaucomatous damage in fundus photographs." Ophthalmology **126**(4): 513-521.

Medeiros, F. A., et al. (2012). "A combined index of structure and function for staging glaucomatous damage." Archives of ophthalmology **130**(9): 1107-1116.

Messenger, H. K. (1964). "Glaukoma and glaucoma." Archives of ophthalmology **71**(2): 264-266.

Meyer, Y. and S. Roques (1993). "Progress in wavelet analysis and applications."

Michalowicz, J. V., et al. (2013). Handbook of differential entropy, Crc Press.

Michalowicz, J. V., et al. (2009). "Calculation of Entropy for a Sinusoid with Beta-Distributed Phase." Entropy **11**(4): 949-958.

Miguel-Jiménez, J., et al. (2010). "Glaucoma detection by wavelet-based analysis of the global flash multifocal electroretinogram." Medical engineering & physics **32**(6): 617-622.

Miguel, A., et al. (2021). "OCT-angiography detects longitudinal microvascular changes in glaucoma: a systematic review." British Journal of Ophthalmology.

Mikelberg, F. S., et al. (1989). "The normal human optic nerve: axon count and axon diameter distribution." Ophthalmology **96**(9): 1325-1328.

Milborrow, S. (2017). earth: Multivariate Adaptive Regression Splines.

Miller, G. (1955). "Note on the bias of information estimates." Information theory in psychology: Problems and methods.

Miyake, Y., et al. (1986). "Congenital stationary night blindness with negative electroretinogram: a new classification." Archives of ophthalmology **104**(7): 1013-1020.

Molina-Martín, A., et al. (2018). Current clinical application of microperimetry: a review. Seminars in Ophthalmology, Taylor & Francis.

Molloy, T. J., et al. (1992). "Electrocardiographic detection of left ventricular hypertrophy by the simple QRS voltage-duration product." Journal of the American College of Cardiology **20**(5): 1180-1186.

Monge, J., et al. (2015). "MEG analysis of neural dynamics in attention-deficit/hyperactivity disorder with fuzzy entropy." Medical engineering & physics **37**(4): 416-423.

Monnik, A. (1870). "Ein neuer Tonometer und sein Gebrauch." Archiv für Ophthalmologie **16**(1): 49-89.

Morgan, S., et al. (1996). "Selective attention to stimulus location modulates the steady-state visual evoked potential." Proceedings of the National Academy of Sciences **93**(10): 4770-4774.

Mortlock, K. E., et al. (2010). "Inter-subject, inter-ocular and inter-session repeatability of the photopic negative response of the electroretinogram recorded using DTL and skin electrodes." Documenta Ophthalmologica **121**(2): 123-134.

Müller, V., et al. (2001). "Investigation of brain dynamics in Parkinson's disease by methods derived from nonlinear dynamics." Experimental brain research **137**(1): 103-110.

Naik, J., et al. (2018). "Short term wind power forecasting using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network." Renewable Energy **118**: 180-212.

Nair, S. S. and K. P. Joseph (2014). "Chaotic analysis of the electroretinographic signal for diagnosis." Biomed Res Int **2014**: 503920.

Nair, S. S. and K. P. Joseph (2014). "Wavelet based electroretinographic signal analysis for diagnosis." Biomedical Signal Processing and Control **9**: 37-44.

Naka, K.-I. and W. A. Rushton (1966). "S-potentials from luminosity units in the retina of fish (Cyprinidae)." The Journal of physiology **185**(3): 587-599.

Nakamura, H., et al. (2011). "Hemispherical focal macular photopic negative response and macular inner retinal thickness in open-angle glaucoma." American journal of ophthalmology **151**(3): 494-506. e491.

Needell, D. and J. A. Tropp (2009). "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples." Applied and computational harmonic analysis **26**(3): 301-321.

Needell, D. and R. Vershynin (2010). "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit." IEEE Journal of selected topics in signal processing **4**(2): 310-316.

Nemenman, I., et al. (2002). Entropy and inference, revisited. Advances in neural information processing systems.

Ng, M., et al. (2012). "Comparison of visual field severity classification systems for glaucoma." Journal of glaucoma **21**(8): 551.

Nguyen, A. T., et al. (2019). "Detecting glaucoma progression using guided progression analysis with OCT and visual field assessment in eyes classified by international classification of disease severity codes." Ophthalmology Glaucoma **2**(1): 36-46.

Nickells, R. W. (1999). "Apoptosis of retinal ganglion cells in glaucoma: an update of the molecular pathways involved in cell death." Survey of ophthalmology **43**: S151-S161.

North, R. V., et al. (2010). "Electrophysiological evidence of early functional damage in glaucoma and ocular hypertension." Investigative ophthalmology & visual science **51**(2): 1216-1222.

Nyquist, H. (1928). "Thermal agitation of electric charge in conductors." Physical review **32**(1): 110.

O'Neill, E. C., et al. (2014). "Glaucomatous optic neuropathy evaluation project: factors associated with underestimation of glaucoma likelihood." JAMA ophthalmology **132**(5): 560-566.

Olsson, J. and H. Rootzén (1994). "An image model for quantal response analysis in perimetry." Scandinavian journal of statistics: 375-387.

Ortiz, G., et al. (2020). "The photopic negative response of the Light-adapted 3.0 ERG in clinical settings." Documenta Ophthalmologica **140**(2): 115-128.

Park, J. W., et al. (2011). "Short-term effects of Ginkgo biloba extract on peripapillary retinal blood flow in normal tension glaucoma." Korean Journal of Ophthalmology **25**(5): 323-328.

Park, K., et al. (2019). "Visual Field Prediction using Recurrent Neural Network." Scientific Reports **9**(1): 1-12.

Peano, G. (1890). "Sur une courbe, qui remplit toute une aire plane." Math. Ann. Mathematische Annalen **36**(1): 157-160.

Pekala, M., et al. (2019). "Deep learning based retinal OCT segmentation." Computers in biology and medicine **114**: 103445.

Percival, W. C. a. D. (2017). wmtsa: Wavelet Methods for Time Series Analysis.

Pham, B. H., et al. (2020). "The rapid N-wave as a potentially useful measure of the photopic negative response." Documenta Ophthalmologica **141**(3): 253-257.

Pincus, S. M., et al. (1991). "A regularity statistic for medical data analysis." Journal of clinical monitoring **7**(4): 335-345.

Pitkow, X. and M. Meister (2012). "Decorrelation and efficient coding by retinal ganglion cells." Nature neuroscience **15**(4): 628-635.

Platner, J. (1745). "Institvtiones Chirvrgiae Rationalis." Leipzig: B Casparis Fritschii: 769.

Porciatti, V. (2015). "Electrophysiological assessment of retinal ganglion cell function." Experimental eye research **141**: 164-170.

Porciatti, V., et al. (2005). "Habituation of retinal ganglion cell activity in response to steady state pattern visual stimuli in normal subjects." Investigative ophthalmology & visual science **46**(4): 1296-1302.

Porciatti, V. and L. M. Ventura (2009). "Adaptive changes of inner retina function in response to sustained pattern stimulation." Vision research **49**(5): 505-513.

Porciatti, V. and L. M. Ventura (2009). "Physiological significance of steady-state PERG losses in glaucoma: clues from simulation of abnormalities in normal subjects." Journal of glaucoma **18**(7): 535.

Preiser, D., et al. (2013). "Photopic negative response versus pattern electroretinogram in early glaucoma." Investigative ophthalmology & visual science **54**(2): 1182-1191.

Prencipe, M., et al. (2020). "The photopic negative response (PhNR): measurement approaches and utility in glaucoma." International ophthalmology: 1-12.

Pukėnas, K., et al. (2012). "Measuring the complexity of a physiological time series: a review." Baltic Journal of Sport and Health Sciences **1**(84).

Quigley, H. A. (2019). "21st century glaucoma care." Eye **33**(2): 254-260.

R Developer Core Team (2017). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Rangaswamy, N. V., et al. (2004). "Photopic ERGs in patients with optic neuropathies: comparison with primate ERGs after pharmacologic blockade of inner retina." Investigative ophthalmology & visual science **45**(10): 3827-3837.

Rangaswamy, N. V., et al. (2004). "Photopic ERGs in patients with optic neuropathies: comparison with primate ERGs after pharmacologic blockade of inner retina." Investigative ophthalmology \& visual science **45**(10): 3827-3837.

Rangaswamy, N. V., et al. (2007). "Effects of spectral characteristics of ganzfeld stimuli on the photopic negative response (PhNR) of the ERG." Investigative ophthalmology & visual science **48**(10): 4818-4828.

Reichstein, D., et al. (2007). "Apoptotic retinal ganglion cell death in the DBA/2 mouse model of glaucoma." Experimental eye research **84**(1): 13-21.

Rényi, A. (1961). On measures of entropy and information, HUNGARIAN ACADEMY OF SCIENCES Budapest Hungary.

Resende, A. F., et al. (2019). "Test–retest repeatability of the pattern electroretinogram and flicker electroretinogram." Documenta Ophthalmologica **139**(3): 185-195.

Richman, J. S. and J. R. Moorman (2000). "Physiological time-series analysis using approximate entropy and sample entropy." American Journal of Physiology-Heart and Circulatory Physiology **278**(6): H2039-H2049.

Riggs, L. A., et al. (1964). "Electrical responses of the human eye to moving stimulus patterns." Science **144**(3618): 567-567.

Risner, M. L., et al. (2018). "Axogenic mechanism enhances retinal ganglion cell excitability during early progression in glaucoma." Proceedings of the National Academy of Sciences **115**(10): E2393-E2402.

Rock, W. J., et al. (1973). "Visual field screening in glaucoma: an evaluation of the Armaly technique for screening glaucomatous visual fields." Archives of ophthalmology **89**(4): 287-290.

Rosolen, S., et al. (2003). "A new model of induced ocular hyperpressure using the minipig." Journal francais d'ophtalmologie **26**(3): 259-267.

Rosolen, S. G., et al. (2004). "Comparing the photopic ERG i-wave in different species." Veterinary ophthalmology **7**(3): 189-192.

Rout, S. and A. Bell (2004). Narrowing the performance gap between orthogonal and biorthogonal wavelets. Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on, IEEE.

Rowe, F. J., et al. (2019). "Accuracy of kinetic perimetry assessment with the Humphrey 850; an exploratory comparative study." Eye **33**(12): 1952-1960.

Rowe, F. J., et al. (2014). "Perimetry Comparisons for Octopus G Top and Dynamic Programmes versus Humphrey 24-2 SITA Fast and SITA Standard Programmes." Ophthalmology Research: An International Journal **2**(1): 24-42.

Sadowsky, J. (1996). "Investigation of signal characteristics using the continuous wavelet transform." johns hopkins apl technical digest **17**(3): 258-269.

Sakata, V. M., et al. (2019). "Full-field electroretinogram behavior in Vogt-Koyanagi-Harada disease: a 24-month longitudinal study in patients from acute onset evaluated with multimodal analysis." Graefe's archive for clinical and experimental ophthalmology **257**(10): 2285-2295.

Sakemi, F., et al. (2002). "Multifocal electroretinograms in early primary open-angle glaucoma." Japanese journal of ophthalmology **46**(4): 443-450.

Salgarello, T., et al. (2008). "Morpho-functional follow-up of the optic nerve in treated ocular hypertension: disc morphometry and steady-state pattern electroretinogram." Current eye research **33**(8): 709-721.

Samojloff, A. (1915). "A SMALL STRING-GALVANOMETER ARRANGED AS A SIGNAL APPARATUS." Quarterly Journal of Experimental Physiology **9**(1): 1-7.

Sarossy, M., et al. (2018). "Discrete wavelet transform of the electroretinogram for glaucoma classification; choice of mother wavelet by variable ranking." Investigative ophthalmology & visual science **59**(9): 5098-5098.

Schillaci, G., et al. (2012). "A review of the role of electrocardiography in the diagnosis of left ventricular hypertrophy in hypertension." Journal of electrocardiology **45**(6): 617-623.

Schiötz, H. (1920). "TONOMETRY." The British journal of ophthalmology **4**(6): 249-266.

Schottdorf, M. and B. B. Lee (2021). "A quantitative description of macaque ganglion cell responses to natural scenes: the interplay of time and space." The Journal of physiology **599**(12): 3169-3193.

Schultz, T., et al. (2020). "Ab externo device for the treatment of glaucoma: direct flow from the anterior chamber to the ocular surface." Journal of Cataract & Refractive Surgery **46**(7): 941-943.

Schuman, J. S., et al. (1995). "Quantification of nerve fiber layer thickness in normal and glaucomatous eyes using optical coherence tomography: a pilot study." Archives of ophthalmology **113**(5): 586-596.

Sedgwick, P. (2013). "Limits of agreement (Bland-Altman method)." Bmj **346**.

Sejdić, E., et al. (2009). "Time--frequency feature representation using energy concentration: An overview of recent advances." Digital Signal Processing **19**(1): 153-183.

Şengür, A., et al. (2016). "Time–frequency texture descriptors of EEG signals for efficient detection of epileptic seizure." Brain informatics **3**(2): 101-108.

Shannon, C. E. (1948). "A mathematical theory of communication." The Bell System Technical Journal **27**: 379-423.

Shapiro, S. S. and M. B. Wilk (1965). "An analysis of variance test for normality (complete samples)." Biometrika **52**(3/4): 591-611.

Sharma, H. (2019). "Heart rate extraction from PPG signals using variational mode decomposition." Biocybernetics and Biomedical Engineering **39**(1): 75-86.

Sheikhani, A., et al. (2007). Analysis of quantitative Electroencephalogram background activity in Autism disease patients with Lempel-Ziv complexity and Short Time Fourier Transform measure. 2007 4th IEEE/EMBS International Summer School and Symposium on Medical Devices and Biosensors, IEEE.

Shrout, P. E. and J. L. Fleiss (1979). "Intraclass correlations: uses in assessing rater reliability." Psychological bulletin **86**(2): 420.

Sihota, R., et al. (2006). "Diagnostic capability of optical coherence tomography in evaluating the degree of glaucomatous retinal nerve fiber damage." Investigative ophthalmology & visual science **47**(5).

Sklar, M. (1959). "Fonctions de repartition an dimensions et leurs marges." Publ. inst. statist. univ. Paris **8**: 229-231.

Smith, P. (1879). Glaucoma: its causes, symptoms, pathology, and treatment, J. & A. Churchill.

Sommer, A., et al. (1984). "Evaluation of nerve fiber layer assessment." Archives of ophthalmology **102**(12): 1766-1771.

Sousa, M. C., et al. (2015). "Suitability of the Visual Field Index according to glaucoma severity." Journal of current glaucoma practice **9**(3): 65.

Spahr, J. (1975). "Optimization of the presentation pattern in automated static perimetry." Vision research **15**(11): 1275-1281.

Srinivasan, V., et al. (2007). "Approximate entropy-based epileptic EEG detection using artificial neural networks." IEEE Transactions on information Technology in Biomedicine **11**(3): 288-295.

Steinhoff, B. J., et al. (1996). "Accuracy and reliability of periodic sharp wave complexes in Creutzfeldt-Jakob disease." Archives of neurology **53**(2): 162-166.

Strela, V. (1996). Multiwavelets--theory and applications. Dept. of Mathematics. Cambridge MA, MIT. **PhD**.

Strömbergsson, D., et al. (2019). "Mother wavelet selection in the discrete wavelet transform for condition monitoring of wind turbine drivetrain bearings." Wind Energy **22**(11): 1581-1592.

Strong, S. P., et al. (1998). "Entropy and information in neural spike trains." Physical review letters **80**(1): 197.

Sturm, B. L. and J. D. Gibson (2006). Matching pursuit decompositions of non-noisy speech signals using several dictionaries. 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, IEEE.

Suda, K., et al. (2018). "Evaluation of structure-function relationships in longitudinal changes of glaucoma using the spectralis OCT follow-up mode." Scientific Reports **8**(1): 1-10.

Sugiyama, K., et al. (1999). "Localized wedge-shaped defects of retinal nerve fiber layer and disc hemorrhage in glaucoma." Ophthalmology **106**(9): 1762-1767.

Sultana, S., et al. (2020). Detection of Stress for Visually Impaired People using EEG Signals based on Time-Frequency Domain Analysis. 2020 International Conference on Machine Learning and Cybernetics (ICMLC), IEEE.

Sustar, M., et al. (2009). "The effect of broadband and monochromatic stimuli on the photopic negative response of the electroretinogram in normal subjects and in open-angle glaucoma patients." Documenta Ophthalmologica **118**(3): 167.

Sutter, E. E. and D. Tran (1992). "The field topography of ERG components in man—I. The photopic luminance response." Vision research **32**(3): 433-446.

Swiderski, B., et al. (2005). Lyapunov exponent of EEG signal for epileptic seizure characterization. Proceedings of the 2005 European Conference on Circuit Theory and Design, 2005., IEEE.

Taglia, D. P., et al. (2002). "Comparison of the Ahmed glaucoma valve, the Krupin eye valve with disk, and the double-plate Molteno implant." Journal of glaucoma **11**(4): 347-353.

Takahashi, T., et al. (2010). "Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: a multiscale entropy analysis." Neuroimage **51**(1): 173-182.

Takahashi, T., et al. (2009). "Age-related variation in EEG complexity to photic stimulation: A multiscale entropy analysis." Clinical Neurophysiology **120**(3): 476-483.

Tallon-Baudry, C., et al. (1996). "Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human." Journal of Neuroscience **16**(13): 4240-4249.

Tang, J., et al. (2014). "The test–retest reliability of the photopic negative response (PhNR)." Translational vision science & technology **3**(6): 1-1.

Tang, J., et al. (2018). "Baseline Detrending for the Photopic Negative Response." Translational vision science & technology **7**(5): 9-9.

Tang, J., et al. (2018). "An optimised protocol for the portable electroretinogram device to probe changes in glaucoma." Investigative ophthalmology & visual science **59**(9): 3028-3028.

Tang, J., et al. (2020). "Short-Term Changes in the Photopic Negative Response Following Intraocular Pressure Lowering in Glaucoma." Investigative ophthalmology & visual science **61**(10): 16-16.

Tao, X., et al. (2020). "Intraocular Pressure Elevation Compromises Retinal Ganglion Cell Light Adaptation." Investigative ophthalmology & visual science **61**(12): 15-15.

Tatham, A. J., et al. (2013). "The relationship between cup-to-disc ratio and estimated number of retinal ganglion cells." Investigative ophthalmology & visual science **54**(5): 3205-3214.

Tham, Y.-C., et al. (2014). "Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis." Ophthalmology **121**(11): 2081-2090.

Thiele, C. and G. Hirschfeld (2021). "cutpointr: Improved Estimation and Validation of Optimal Cutpoints in R." Journal of Statistical Software **98**(11): 1 - 27.

> "Optimal cutpoints" for binary classification tasks are often established by testing which cutpoint yields the best discrimination, for example the Youden index, in a specific sample. This results in "optimal" cutpoints that are highly variable and systematically overestimate the out-of-sample performance. To address these concerns, the cutpointr package offers robust methods for estimating optimal cutpoints and the out-of-sample performance. The robust methods include bootstrapping and smoothing based on kernel estimation, generalized additive models, smoothing splines, and local regression. These methods can be applied to a wide range of binary-classification and cost-based metrics. cutpointr also provides mechanisms to utilize user-defined metrics and estimation methods. The package has capabilities for parallelization of the bootstrapping, including reproducible random number generation. Furthermore, it is pipe-friendly, for example for compatibility with functions from tidyverse. Various functions for plotting receiver operating characteristic curves, precision recall graphs, bootstrap results and other representations of the data are included. The package contains example data from a study on psychological characteristics and suicide attempts suitable for applying binary classification algorithms.

Torres, M. E., et al. (2011). A complete ensemble empirical mode decomposition with adaptive noise. Acoustics, speech and signal processing (ICASSP), 2011 IEEE international conference on, IEEE.

Townsend, K., et al. (2008). "Heidelberg Retina Tomograph 3 machine learning classifiers for glaucoma detection." British Journal of Ophthalmology **92**(6): 814-818.

Tribble, J. R., et al. (2019). "Midget retinal ganglion cell dendritic and mitochondrial degeneration is an early feature of human glaucoma." Brain Communications **1**(1): fcz035.

Tropp, J. A. (2004). "Greed is good: Algorithmic results for sparse approximation." IEEE Transactions on Information theory **50**(10): 2231-2242.

Tropp, J. A. and A. C. Gilbert (2007). "Signal recovery from random measurements via orthogonal matching pursuit." IEEE Transactions on Information theory **53**(12): 4655-4666.

Troy, J. and B. Lee (1994). "Steady discharges of macaque retinal ganglion cells." Visual neuroscience **11**(1): 111-118.

Tun, T. A., et al. (2017). "Evaluation of the anterior segment angle-to-angle scan of cirrus high-definition optical coherence tomography and comparison with gonioscopy and with the visante OCT." Investigative ophthalmology & visual science **58**(1): 59-64.

Vainio-Mattila, B. (1951). "The clinical electroretinogram: II. The difference between the electroretinogram in men and in women." Acta ophthalmologica **29**(1): 25-32.

van Steveninck, R. R. d. R., et al. (1997). "Reproducibility and variability in neural spike trains." Science **275**(5307): 1805-1808.

Vaz, S., et al. (2013). "The case for using the repeatability coefficient when calculating test–retest reliability." PloS one **8**(9): e73990.

Vingrys, A. J. and K. A. Helfrich (1990). "The Opticom M-600™: A new LED automated perimeter." Clinical and Experimental Optometry **73**(1): 3-17.

Viswanathan, S. and L. Frishman (1997). Evidence that negative potentials in the photopic electroretinograms of cats and primates depend upon spiking activity of retinal ganglion cell axons. Soc Neurosci Abstr.

Viswanathan, S., et al. (1999). "The photopic negative response of the macaque electroretinogram: reduction by experimental glaucoma." Investigative ophthalmology & visual science **40**(6): 1124-1136.

Viswanathan, S., et al. (2001). "The photopic negative response of the flash electroretinogram in primary open angle glaucoma." Investigative ophthalmology & visual science **42**(2): 514-522.

von Anrep, B. (1880). "Ueber die physiologische Wirkung des Cocaïn." Archiv für die Gesamte Physiologie des Menschen und der Tiere **21**(1): 38-77.

Von Graefe, A. (1854). "Mittheilungen von Krankheitsfällen und Notizen vermischten Inhalts." Archiv für Ophthalmologie **1**(1): 283-465.

Von Plato, J. (1991). "Boltzmann's ergodic hypothesis." Archive for history of exact sciences **42**(1): 71-89.

Wahlfors, K. (1888). "Ober Druck und Druckmessung im menschlichen Auge. 7. internat." Ophthal. Kongr: 268-274.

Wakili, N., et al. (2008). "The photopic negative response of the blue-on-yellow flash-electroretinogram in glaucomas and normal subjects." Documenta Ophthalmologica **117**(2): 147-154.

Wang, A. Y., et al. (2021). "Retinal ganglion cell dysfunction in mice following acute intraocular pressure is exacerbated by P2X7 receptor knockout." Scientific Reports **11**(1): 1-16.

Wang, W., et al. (2001). "Analysis of the first heart sound using the matching pursuit method." Medical and Biological Engineering and Computing **39**(6): 644-648.

Wang, X., et al. (2011). "Comparative study of retinal nerve fibre layer measurement by RTVue OCT and GDx VCC." British Journal of Ophthalmology **95**(4): 509-513.

Weber (1858). "Bijdragc tot de kennis van her glaucoma im zweiten." Archiv für Ophthalmologie **1**: 206.

Weber, A. J. and C. D. Harman (2005). "Structure–function relations of parasol cells in the normal and glaucomatous primate retina." Investigative ophthalmology & visual science **46**(9): 3197-3207.

Wei, Y. and M.-C. Chen (2012). "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks." Transportation Research Part C: Emerging Technologies **21**(1): 148-162.

Weibull, W. (1951). "A statistical distribution function of wide applicability." Journal of applied mechanics.

Weinreb, R. N., et al. (2014). "The pathophysiology and treatment of glaucoma: a review." Jama **311**(18): 1901-1911.

Weleber, R. G. (1981). "The effect of age on human cone and rod ganzfeld electroretinograms." Investigative ophthalmology & visual science **20**(3): 392-399.

Werblin, F. (1971). "Adaptation in a vertebrate retina: intracellular recording in Necturus." Journal of Neurophysiology **34**(2): 228-241.

Werft, M. C. T. a. G. B. a. T. D. a. N. H. a. F. K. a. K. R. a. J. R. a. M. S. a. W. (2017). mutoss: Unified Multiple Testing Procedures.

Westall, C. A., et al. (2002). "The Hospital for Sick Children, Toronto, Longitudinal ERG study of children on vigabatrin." Documenta Ophthalmologica **104**(2): 133-149.

Wild, J. (1988). "Techniques and developments in automated perimetry: a review." Ophthalmic and Physiological Optics **8**(3): 295-308.

Williams, P. A., et al. (2017). "Vitamin B3 modulates mitochondrial vulnerability and prevents glaucoma in aged mice." Science **355**(6326): 756-760.

Williams, P. A., et al. (2013). "Retinal ganglion cell dendritic atrophy in DBA/2J glaucoma." PloS one **8**(8): e72282.

Wise, J. B. and S. L. Witter (1979). "Argon laser therapy for open-angle glaucoma: a pilot study." Archives of ophthalmology **97**(2): 319-322.

Woody, C. D. (1967). "Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals." Medical and biological engineering **5**(6): 539-554.

Wright, T., et al. (2008). "A comparison of signal detection techniques in the multifocal electroretinogram." Documenta Ophthalmologica **117**(2): 163-170.

Wright, T. J., et al. (2011). "Isolating Visual Evoked Responses—Comparing Signal Identification Algorithms." Journal of Clinical Neurophysiology **28**(4): 404-411.

Wu, C., et al. (2011). "Action potential generation at an axon initial segment-like process in the axonless retinal AII amacrine cell." Journal of Neuroscience **31**(41): 14654-14659.

Wu, S., et al. (1995). "Effects of flicker adaptation and temporal gain control on the flicker ERG." Vision research **35**(21): 2943-2953.

Wu, Z., et al. (2016). "Measuring the photopic negative response: viability of skin electrodes and variability across disease severities in glaucoma." Translational vision science & technology **5**(2): 13-13.

Wu, Z., et al. (2016). "Photopic negative response obtained using a handheld electroretinogram device: determining the optimal measure and repeatability." Translational vision science & technology **5**(4): 8-8.

Wu, Z. and N. E. Huang (2009). "Ensemble empirical mode decomposition: a noise-assisted data analysis method." Advances in adaptive data analysis **1**(01): 1-41.

Wu, Z. and F. A. Medeiros (2021). "A simplified combined index of structure and function for detecting and staging glaucomatous damage." Scientific Reports **11**(1): 1-10.

Xia, X.-G., et al. (1996). "Design of prefilters for discrete multiwavelet transforms." IEEE transactions on signal processing **44**(1): 25-35.

Xiao, Q., et al. (2018). "Natural-gas pipeline leak location using variational mode decomposition analysis and cross-time–frequency spectrum." Measurement **124**: 163-172.

Xie, C., et al. (2015). "Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging." Scientific Reports **5**(1): 1-11.

Xin, Y., et al. (2017). "Paroxysmal atrial fibrillation recognition based on multi-scale Rényi entropy of ECG." Technology and Health Care **25**(S1): 189-196.

Yadav, K. S., et al. (2020). "Bio-tactics for neuroprotection of retinal ganglion cells in the treatment of glaucoma." Life sciences **243**: 117303.

Yanti, D. K., et al. (2011). Widely linear based filter using Short Term Fourier Transform for Visual Evoked Potential extraction. 2011 National Postgraduate Conference, IEEE.

Yeh, J.-R., et al. (2010). "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method." Advances in adaptive data analysis **2**(02): 135-156.

Yu, K.-B. and C. D. Mc Gillem (1983). "Optimum filters for estimating evoked potential waveforms." IEEE Transactions on Biomedical Engineering(11): 730-737.

Zadeh, L. A. (1965). "Fuzzy Sets." Information and Control **8**(3): 338-353.

Zeng, X., et al. (2001). "Design and implementation of an estimator of fractal dimension using fuzzy techniques." Pattern Recognition **34**(1): 151-169.

Zhang, D.-x., et al. (2008). The EEG signal preprocessing based on empirical mode decomposition. Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. The 2nd International Conference on, IEEE.

Zhang, L. and V. P. Singh (2007). "Gumbel–Hougaard copula for trivariate rainfall frequency analysis." Journal of Hydrologic engineering **12**(4): 409-419.

Zhang, P.-L., et al. (2012). "Bearing fault detection using multi-scale fractal dimensions based on morphological covers." Shock and Vibration **19**(6): 1373-1383.

Zhang, S., et al. (2019). "Integrating opportunistic glaucoma screening into general health examinations in China: A pilot study." Clinical & experimental ophthalmology **47**(8): 1000-1008.

Zhang, X. and P. Zhou (2012). "Sample entropy analysis of surface EMG for improved muscle activity onset detection against spurious background spikes." Journal of Electromyography and Kinesiology **22**(6): 901-907.

Zhao, X., et al. (2020). "A quadratic penalty item optimal variational mode decomposition method based on single-objective salp swarm algorithm." Mechanical Systems and Signal Processing **138**: 106567.

Zhao, Y., et al. (1990). "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals." IEEE Transactions on Acoustics, Speech, and Signal Processing **38**(7): 1084-1091.

Zhou, W., et al. (2007). "Oscillatory potentials of the slow-sequence multifocal ERG in primates extracted using the Matching Pursuit method." Vision research **47**(15): 2021-2036.

Zvoleff, A. (2020). glcm: Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs). R package.

Appendix

Published papers are shown in the following pages in chronological order

Timing of Changes in the Entropy of the Electroretinogram With Glaucoma

M. Sarossy[1, 2], *Member, IEEE*, B. Aliahmad[1*], *Member, IEEE,* and D. K. Kumar[1], *Senior Member, IEEE*

***Abstract*— The photopic negative response (PhNR) component of the electroretinogram (ERG) is caused by loss of ganglion cells in the retina and can be used for the assessment of glaucoma. The purpose of this study was to use information theory techniques to explore subtle timing changes in the electroretinogram in glaucoma patients. The entropy of ERG signals were calculated within a sliding window, filtered with KZA filter to find the peaks and the timing of entropy changes determined. We found an earlier rise in entropy in glaucoma patients showing that this may have potential as new technique for glaucoma diagnosis.**

I. INTRODUCTION

Glaucoma is a neurodegenerative disease of the optic nerve that is caused by loss of ganglion cells in the retina and is related to increased pressure within the eye (IOP). Vision loss is usually insidious and commences in the peripheral field of vision. Because the loss can be very slow, patients are frequently unaware of the loss at the early stages. This loss is, however, permanent and irreversible. Once detected, the only means of preventing further loss is reduction of the IOP with drops, surgery, laser or a combination of treatments. There is no single safe intraocular pressure and a target or 'safe' pressure must be found or estimated for each patient. As all of the treatments have side effects, the challenge for a clinician is to find the earliest changes in visual function – before any actual symptoms have developed. In assessing the adequacy of treatment, the clinician needs to evaluate the function and structure of the optic nerve to ensure no further loss is occurring at the current IOP. Delayed diagnosis or inadequate treatment of glaucoma will result in permanent structural damage to the optic nerve and vision loss. Tests such as confocal scanning laser ophthalmoscopy and optical coherence tomography assess the structure of the nerve, but it is possible that the changes in the function precede these structural changes. This structure-function 'gap', i.e. the difference between the structural and functional changes, may predict the reversible component of the disease.

Functional changes should manifest at an earlier time point within the disease course compared to structural changes – which are usually caused by the death and apoptosis of cells. Current tests of function in glaucoma are visual field testing: measuring the sensitivity of the visual system at various degrees of eccentricity and electrophysiological testing. The Photopic Negative Response (PhNR) component of electroretinogram (ERG) and the pattern ERG have both been shown to change in glaucoma. They can be effective in detection of the disease prior to permanent loss of ganglion cells [1] and are more objective than visual field testing which requires voluntary responses from the patient.

The pattern ERG (PERG) is a smaller amplitude response dependent upon good patient visual fixation and very sensitive to loss of central acuity. The PhNR is the slow negative potential occurring after the b-wave of ERG with an appropriate ganzfeld stimulus. This component of ERG is known to originate from the action potentials of retinal ganglion cells [2, 3]. The response has been

M. S. is with the school of Engineering, RMIT university, Melbourne, Australia and Center for Eye Research, University of Melbourne, Australia, phone: (03) 99259533; e-mail: marc.sarossy@rmit.edu.au).

B. A*., is with the school of Engineering, RMIT university, Melbourne, Australia, (e-mail: behzad.aliahamd@rmit.edu.au).

D. K. is with the school of Engineering, RMIT university, Melbourne, Australia (e-mail: dinesh@rmit.edu.au).

978-1-5386-1030-5/17/$31.00 ©2017 IEEE

shown to be reduced as a result of glaucoma in clinical studies and animal models and therefore it could be a useful measure for evaluation of inner retinal function of glaucomatous eyes. Niyadurupola *et al.* [4] studied glaucomatous and ocular hypertensive eyes in terms of how lowering IOP would affect the PhNR. They found significant improvement in raw PhNR amplitude and the ratio of PhNR to b-wave in response to lowered IOP. In another study PhNR amplitudes was found to be correlated with thickness of ganglion cells in the center of the macula indicating another test which can be used for glaucoma assessment [5]. ElGohary *et al.* [6] characterized the PhNR in rabbits by inducing glaucoma in the right eyes. The result showed significant decrease in PhNR amplitude within one week of induction (P=0.0001) but no significant association or change in the timing of PhNR was observed (P=0.132). In all of these studies only the amplitude of the PhNR has been reported to be effective for evaluation of glaucoma. However the negative going peak of the response is difficult to define, and the timing/ latency of the PhNR was found to be unrelated to the disease; this makes it unsuitable for diagnosis.

In this work we have hypothesized that the information and hence the time function of the entropy of the ERG signal will vary in post stimulus period with increasing glaucoma. This will be manifested by group differences between glaucoma and non-glaucoma subjects. In this study, we used a sliding window on each trace to compute the entropy of the signal within the window; determining the onset of the light evoked ERG and the offset of the response by the timing of the changes in entropy.

II. DATA COLLECTION

Informed consent was obtained from all participants. The left eyes of 21 glaucoma patients (age 34-86, mean age 66.4) and 18 normal aged matched controls (44-84 mean age 60.7) were tested. ERG signals were recorded with an LKC *RetEval* hand held ERG device. A red flash of 4ms duration was used as the stimulus on a 10cd/m^2 blue background. Flash energy was set to 1.5cd.s/m^2. The sample rate of the device was set to 1.9 K Samples per second at 24 bits resolution, an epoch of 220ms and pre-stimulus acquisition starting at 100ms. 200 sweeps were performed for each session. Proprietary 3 pole skin electrodes were used which incorporated active, indifferent and earth electrodes. B-wave amplitude (the first positive going wave measured relative to the a-wave) and implicit time were measured with the provided device software.

III. METHODOLOGY

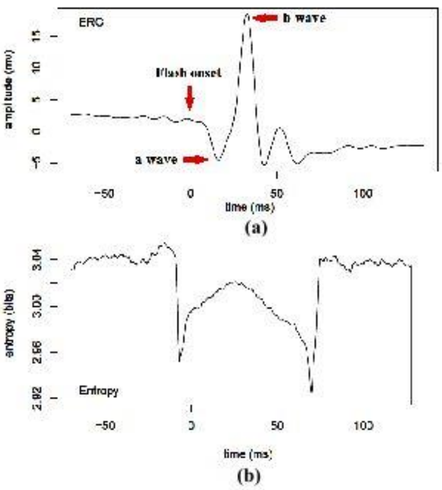
The raw output of the ADC was exported as a file. Entropy analysis was performed in RStudio, ver. 0.99.486 [7] on the raw data using the theory developed by Shannon [8]. Considering p_i as the probability of occurrence of the i^{th} type of element, Shannon entropy H(x) can be calculated as in (1):

$$H(X) = -\sum_{i=1}^{N} p_i \log_2 p_i \qquad (1)$$

where N is the total number of element types. Shannon entropy of a vector of values is invariant to the order of the values in the array.

Entropy was calculated with a sliding window of width 150 samples which was advanced by one sample at a time. For each window, the vector of ADC values was binned into a histogram of 30 bins using the entropy package in R [9]. The entropy of the histogram was estimated using the maximum likelihood plugin estimator. The mean entropy over all sweeps for each window time point was calculated to obtain the entropy signal expressed in the time domain. The mean entropy signal was filtered with the Komolgorov Zurbenko adaptive (kza) filter [10] with a windows size of 30 and 3 iterations. KZA is an iterated adaptive moving average filter and a non-parametric technique for locating breaks in time-series signals when it is buried in noise. The rise in entropy after the conclusion of the signal was detected by finding the maximum of in the first derivative of the response after KZA filtration. The first numerical derivative was estimated using the "sfsmisc" package in R [11] The last positive peak of the derivative corresponds with the time point of the maximum gradient of the terminal entropy rise after the entropy dip as shown in Fig. 2 b.

The b-wave of the ERG signal was also analyzed for its amplitude and latency to find whether there was any significant difference between the normal and glaucomatous eyes.



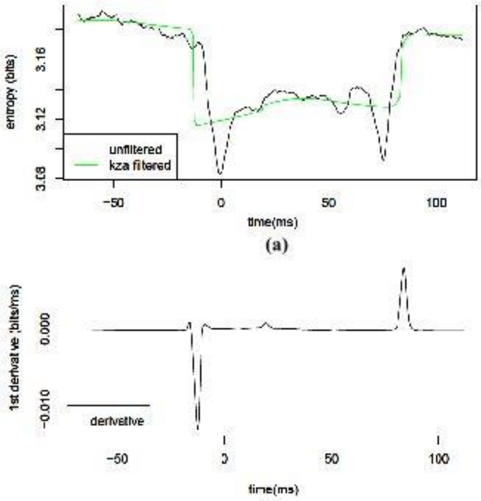
(a)

(b)

Figure 1: Example of A typical (Healthy) ERG and Entropy a) ERG signal b) Shannon entropy

IV. RESULTS

Fig 2.a shows entropy of ERG of a glaucoma patient and effect of KZA filtering on the entropy signal. The upper panel shows the mean entropy of ERG (black curve) as well as the KZA filtered version (green curve). The lower panel shows the first numerical derivative of the KZA filtered entropy signal. As can be seen in this figure, peaks of the first derivative of filtered entropy show the time point of the entropy dip and subsequent rise.



(a)

(b)

Figure 2: KZA filtering and peak detection a) unfiltered (black) and filtered (green) entropy of Glaucomatous ERG b) Numerical first derivative of filtered entropy

Table 1 shows the results of the entropy timing calculations and b-wave parameters for the controls and glaucoma cases. The results show no significant difference between the glaucoma and control eyes in the implicit time of the b-wave (31.1ms vs. 30.6 ms p=0.66) or its amplitude (19.3 vs. 17.0 mv, p=0.18). However, the timing of the rise in entropy after the b-wave was significantly different (74.7ms vs. 79.8ms, p =0.04).

Table 1: Results of the timing calculation

Result	Controls	Glaucoma	*P*
Test subjects	18	21	-
b amplitude (mv)	31.1	30.7	0.66
b latency (ms)	19.3	17.1	0.17
Entropy rise(ms)	79.8	74.7	0.04

V. DISCUSSION & CONCLUSION

The magnitude of the PhNR is well known to decrease in patients with glaucoma and also in animal models of glaucoma. The source of the PhNR itself has been shown to be from the spiking ganglion cells of the retina which will be affected by glaucoma and increases in IOP of the eye. Glaucoma is a disease that usually affects the more peripheral ganglion cells first; manifesting as peripheral loss of visual field and loss of the ganglion and nerve fiber layers of the retina more peripherally. The signals from the more peripheral cells have a longer transit time within the eye. We suggest that the underpinning physiology of our finding: an earlier rise in the entropy or randomness of the ERG signal in Glaucoma patients may be a result from the preferential loss of these more peripheral ganglion cells and is analogous to a shorter "settling time" in the presence of this disease. The algorithm for calculating the entropy via sliding windows is computationally simple and efficient and it lends itself to implementation on an embedded platform like the hand held ERG unit. By studying the timing of the changes of the entropy, the glaucoma physician may be better able to diagnose the earliest changes of glaucoma in a person at risk, or realize that the IOP is too high in a patient already diagnosed and receiving treatment. In this way, electrophysiological testing may help to prevent irreversible vision loss.

REFERENCES

[1] K. T. Tsaousis, S. Plainis, N. R. Parry, I. G. Pallikaris, M. K. Tsilimbaris, and E. T. Detorakis, "Visual electrodiagnosis in glaucoma screening: a clinical study," *J Glaucoma,* vol. 22, pp. 427-31, Jun-Jul 2013.

[2] S. Viswanathan, L. J. Frishman, J. G. Robson, R. S. Harwerth, and r. E. L. Smith, "The photopic negative response of the macaque electroretinogram: reduction by experimental glaucoma," *Investigative Ophthalmology & Visual Science,* vol. 40, pp. 1124-1136, 1999.

[3] A. L. Georgiou, L. Guo, M. Francesca Cordeiro, and T. E. Salt, "Electroretinogram and visual-evoked potential assessment of retinal and central visual function in a rat ocular hypertension model of glaucoma," *Curr Eye Res,* vol. 39, pp. 472-86, May 2014.

[4] N. Niyadurupola, C. D. Luu, D. Q. Nguyen, K. Geddes, G. X. Tan, C. C. Wong, *et al.*, "Intraocular pressure lowering is associated with an increase in the photopic negative response (PhNR) amplitude in glaucoma and ocular hypertensive eyes," *Invest Ophthalmol Vis Sci,* vol. 54, pp. 1913-9, Mar 15 2013.

[5] K. Tamada, S. Machida, T. Oikawa, H. Miyamoto, T. Nishimura, and D. Kurosaka, "Correlation between Photopic Negative Response of Focal Electroretinograms and Local Loss of Retinal Neurons in Glaucoma," *Current Eye Research,* vol. 35, pp. 155-164, 2010/02/01 2010.

[6] A. A. ElGohary and L. H. Elshazly, "Photopic negative response in diagnosis of glaucoma: an experimental study in glaucomatous rabbit model," *Int J Ophthalmol,* vol. 8, pp. 459-64, 2015.

[7] R. C. Team, "R: A Language and Environment for Statistical Computing," ed. Vienna, Austria, 2015.

[8] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal,* vol. 27, pp. 379-423, 1948.

[9] J. Hausser and K. Strimmer, "entropy: Estimation of Entropy, Mutual Information and Related Quantities," ed, 2014, p. R package version 1.2.1.

[10] I. Zurbenko, P. S. Porter, R. Gui, S. T. Rao, J. Y. Ku, and R. E. Eskridge, "Detecting Discontinuities in Time Series of Upper-Air Data: Development and Demonstration of an Adaptive Filter Technique," *Journal of Climate,* vol. 9, pp. 3548-3560, 1996.

[11] M. Maechler, "sfsmisc: Utilities from "Seminar fuer Statistik" ETH Zurich," ed, 2016, pp. R package version 1.1-0.

2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII) | 978-1-6654-4126-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICBSII51839.2021.9445121

Relationship between Glaucoma and Complexity Measures of the Electroretinogram

Marc Sarossy
Dept of Surgery
University of Melbourne
Melbourne, Australia
marc@sarossy.com

Dinesh Kumar
Dept of Electrical Engineering
RMIT University
Melbourne, Australia
dinesh@rmit.edu.au

Zhichao Wu
Macular Research Unit
Centre for Eye Research Australia
East Melbourne, Australia
wu.z@unimelb.edu.au

***Abstract*—Glaucoma is one of the leading causes of vision loss in the world. Early diagnosis is essential to prevent irreversible vision loss. Such early diagnosis requires the use of tests of visual function rather than tests of anatomical structure. Existing tests of visual function can suffer from the use of bulky expensive equipment and poor subject performance. In this work we describe a novel analysis of the electroretinogram to extract features of complexity and detect glaucoma. Cohorts of participants with and without glaucoma had electroretinogram (ERG) testing. The stimulus used to evoke the signal was that used for the Photopic Negative Response (PhNR), a response known to be reduced in glaucoma. There were 18 control and 21 glaucoma participants. Significant differences were found in PhNR amplitude -5.5 mV vs -3.6mV ($p<0.01$), Shannon Entropy 2.99 vs 2.80 ($p<0.01$) and Higuch Fractal Dimension 1.068 vs 1.082 ($p<0.02$). Differences in box count and Katz Fractal Dimension and Sample Entropy were not significant. These complexity measures may have potential for machine learning classifiers for glaucoma.**

Keywords—glaucoma, entropy, fractal dimension, electroretinogram

I. INTRODUCTION

Glaucoma is a disease of the optic nerve of the eye – the structure that electrically connects the eye to the brain. The nerve is comprised of the axons of the ganglion cells of the retina and in a healthy human eye number around one million. Glaucoma is caused by elevated pressure within the eyeball which triggers a cascade of damage within the ganglion cells resulting in death by apoptosis of the ganglion cells [1]. The exact mechanism of the damage remains uncertain however the clinical pattern of evolving glaucoma in an eye is recognizable by a loss of peripheral visual field, stereotypical structural changes at the optic nerve head and changes in the thickness of the ganglion cell and nerve fiber layer on the optical coherence tomography scan. The disease is the second most frequent cause of legal blindess registrations in industrialized countries and the number of people aged 30-80 in the world with glaucoma is estimated to be 64.3 million with 60% of those in Asia [2] Early diagnosis of glaucoma is essential because there are many treatments that reduce the intraocular pressure and consequently reduce the rate and risk of progression.

The electroretinogram is an evoked potential from the eye, measured at the cornea relative to the retina. It is evoked by a brief flash of light. It has been shown [3] that a component of the response – the Photopic Negative Response (PhNR) is reduced by glaucoma, both in experimental monkey models of glaucoma and in human patients with the disease. This response is of low amplitude and may be contaminated by blink artefact. Despite the publication of a standard [4] the test-retest reliability of the response remains low [5]

The PhNR arises from the spiking potentials of ganglion cells. As the disease progresses, the number of ganglion cells generating the response is reduced. Each ganglion cell produces a stochastic spike train where the spike rate follows a Poisson distribution the parameter influenced by the stimulating light level within that cell's receptive field. A reduction in the number of ganglion cells should be manifested by a reduction in the complexity of the part of the ERG generated by the ganglion cells

Various measures have been described to measure the complexity of signals. Shannon entropy [6] is one measure of complexity that can be applied to a time series. It is a function of the probabilities of the levels within the set of measures – that is a function of the probability density function of the amplitude measures. Sample entropy (SampEn) [7] is another measure which relates to the similarity of sequences of different lengths within the signal.

Fractal dimension is a different method that can be used to measure the complexity of a time series [8]. It describes the rate of change of length as scale increases when plotted on a log-log scale. If an object exhibits fractal properties, it will show a linear relationship on this plot. The measure has found utility in analysis of retinal vasculature.[9]

II. METHODOLOGY

A. Participants

This research followed the tenets of the Declaration of Helsinki and was approved by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital. 2 groups of participants were recruited: Glaucoma and non glaucoma controls. Inclusion criterion for the glaucoma cohort was a known diagnosis of glaucoma and a mean deviation on the visual field test of at least -3dB.

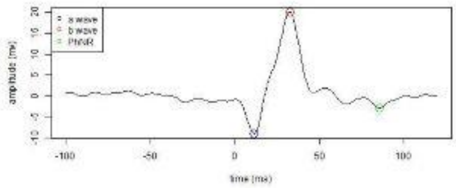


Fig. 1. Sample ERG trace

978-1-6654-4126-1/21/$31.00 ©2021 IEEE

TABLE I. ERG RESULTS

Result	Controls	Glaucoma	p
Test subjects	18	21	-
b amplitude (mv)	31.1	30.7	0.66
b latency (ms)	19.3	17.1	0.17
PhNR (mv)	-5.50 (1.95)	-3.60 (2.53)	0.01

TABLE II. COMPLEXITY MEASURES

Result	Controls	Glaucoma	p
Test subjects	18	21	-
Box count FD	1.148	1.157	0.57
Higuchi FD	1.068	1.082	0.02
Katz FD	1.005	1.007	0.08
SampEn	0.077	0.072	0.07
Shannon entropy	2.99	2.80	0.01

Normal controls were recruited from family and friends of the glaucoma cohort with no reported history of eye disease. Informed consent was obtained for all participants...

B. *ERG Recordings*

In this study, only left eyes were considered. ERG responses were recorded from eyes without pupil dilation. ERG signals were recorded with the RETeval™ hand held device (LKC Technologies, Gaithersburg, MD). A red flash of 4ms duration was used as the stimulus on a 10cd/m2 blue background. Flash energy was set to 1.5cd.s/m2. The sample rate of the device was set to 1.9 K Samples per second at 24 bits resolution, an epoch of 220ms and pre-stimulus acquisition starting at 100ms. 200 sweeps were performed for each session. Proprietary 3 pole skin electrodes were used which incorporated active, indifferent and earth electrodes. B-wave amplitude (the first positive going wave measured relative to the a-wave) and implicit time were measured with the provided device software.

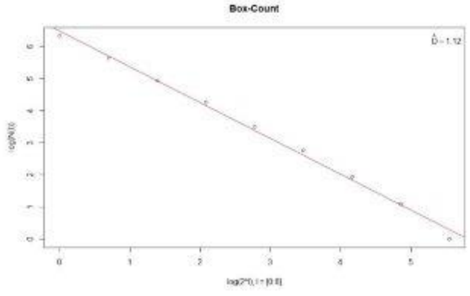


Fig. 3. Box Counting Fractal Dimension Example

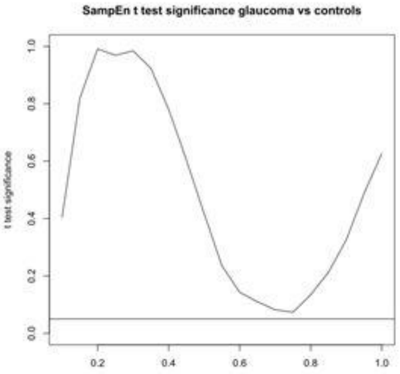


Fig. 2. Effect of radius on Sample Entropy

Raw data were exported from the device at the conclusion of each testing session and analysed with R. The a wave, b wave and PhNR amplitude and implicit times were calculated from the processed waveform produced by the device and corresponded to the first negative going wave and the first positively going wave after the stimulus for the a an b wave respectively. The a wave amplitude was measured relative to the detrended baseline and the b wave was measured relative to the a wave. The PhNR was defined as the first negative going wave after the b wave and was referenced to the prestimulus baseline.

C. *Mathematical Analysis*

All calculations were performed in R. Shannon Entropy was estimated using the Entropy package. Data was discretized into 30 bins for the calculation. Sample entropy was estimated using a custom script in R with a radius of 0.2 times the standard deviation of the signal. Fractal dimensions were estimated using the FractalDim package [10]. Boxcount, Higuchi[11] and Katz [12] entropies were calculated.

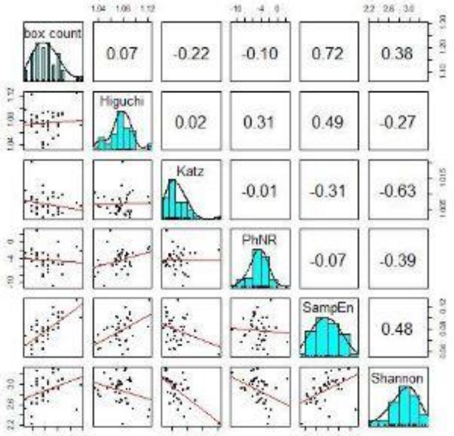


Fig. 4. Pairs plot of complexity measures

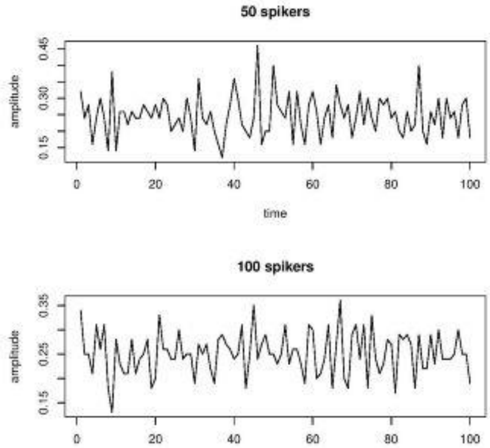


Fig. 5. Simulated Electroretinogram signals

III. RESULTS

A. Electroretinogram Waves

Fig 1 shows a typical ERG trace consisting of the average of the 200 sweeps. Markers have been placed on the a and b waves and the PhNR.

Table I shows the results from the glaucoma and control cohort. There were no significant differences in b wave amplitude or latency between the glaucoma and control groups. Glaucoma is known not to affect this wave. There was a significant difference in the PhNR ampltitudes between control and glaucoma cohorts.

B. Fractal Properties of the Electroretinogram

Fig 2 shows a box count FD calculation for the waveform shown in Fig 1. Good fit was obtained in all signals for the box count log log plot with R^2 measures above 0.95 for all of the ensemble averages

C. Complexity Measures and Glaucoma

Table II shows the performance of the various complexity measures in classifying glaucoma and non glaucoma subjects. Box count FD was far from significance whereas Higuchi FD and Shannon entropy had highly significant differences between the cohorts. Katz entropy and sample entropy approached but did not reach significance

Fig 3 shows how the radius affects the performance of SapmEn in being able to separate glaucoma from non glaucoma subjects. The radius for SampEn is the tolerance within which subsequences are considered to be the same. Fig 2 shows that for this data set, the optimum radius to discriminate between glaucoma and non glaucoma individuals is 0.7 times the standard deviation

D. Relationship between Complexity Measures

Fig 4 shows the relationship between all the various complexity measures and the PhNR for all test subjects (glaucoma and control). There were weak to moderate correlations between the measures. Moderate correlations were found between the PhNR and both Shannon entropy and Higuchi FD – both of which had good classification value when separating glaucoma and non glaucoma. Box count FD was very poorly correlated with PhNR mirroring its inability to differentiate between the cohorts.

IV. DISCUSSION

Making an early diagnosis of glaucoma is critical to the final outcome in the management of the disease. Early disease is often asymptomatic and significant loss of visual function can occur before diagnosis. Screening of at-risk individuals can help: targeting people with a family history of glaucoma or older people with other risk factors. [13]

Current diagnosis relies upon a mixture of objective and subjective measures. The most important subjective measure is the determination of the visual field. This is a quantitative measurement of the brightness of light required to be detected as a function of spatial location within the visual space

Patients do find the performance of the visual field test (VFT) to be stressful [14]. The test can also suffer from unreliability as it does require patient attention and cooperation over a prolonged period. Tests such as optical coherence tomography (OCT) of the optic nerve or analysis of disc appearance via ophthalmolscopy or photographic imaging of the disc are tests of structure rather than function and as such will not be able to detect the presence of sick nerve fibres within the optic nerve.

The RetEVAL is marketed as a low-cost device suitable for screening glaucoma. The device is hand held, and is used with skin electrodes – compared with other ERG devices which are expensive and require the placement of more invasive contact lens or fiber electrodes. The ERG, like the VFT is a test of function and should be sensitive to changes in function in the ganglion cells and their axons, even if the cells have not yet undergone apoptosis [15]

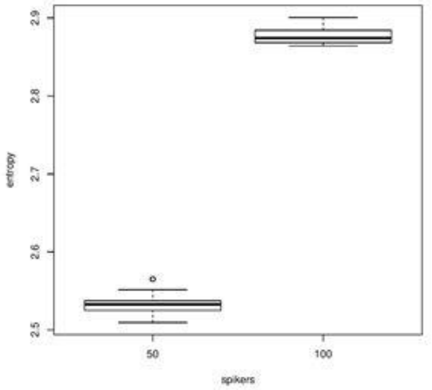


Fig. 6. Shannon entropy from simulated electroretinogram signals

The PhNR has been shown to be affected by glaucoma but as yet lacks the reliability to form part of the diagnostic arsenal for the glaucoma specialist and has not been shown to be of value in glaucoma screening.

In this paper, we have shown that the ERG is affected in ways beyond the simple PhNR in glaucoma and it is possible that these measures of complexity either on their own or combined with the PhNR will aid in either the diagnosis or the management of glaucoma

The response is known to be generated by the summing of action potentials from the ganglion cells. Each of those spiketrains is a Poisson process.

Fig 5 shows a simulated analog signal – formed from the sum of 50 independently discharging Poisson generators in the top panel and 100 in the lower panel. The signals do not show any obvious difference in the time domain

Fig 6 shows boxplots of Shannon Entropies of similar signals for 1000 such simulations. The graph shows very different entropy levels with the signals that are generated with more independent spking generators having the higher entropy and therefore complexity. This is a plausible explanation for the differences seen between the glaucoma and non glaucoma cohort.

This study was limited by the small sample size and did not take into account the severity of glaucoma beyond a threshold level. Even given the limitations, we have shown an interesting and plausible fingerprint of glaucoma within the ERG signal that could suggest a feature to consider for future automated diagnosis statistical model.

ACKNOWLEDGMENT

The assistance of the staff of the glaucoma clinic at the Royal Victorian Eye and Ear Hospital was very valuable in collecting the data for this manuscript.

REFERENCES

[1] McKinnon, S.J.: 'Glaucoma, apoptosis, and neuroprotection', Current opinion in ophthalmology, 1997, 8, (2), pp. 28-37

[2] Tham, Y.-C., Li, X., Wong, T.Y., Quigley, H.A., Aung, T., and Cheng, C.-Y.: 'Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis', Ophthalmology, 2014, 121, (11), pp. 2081-2090.

[3] Viswanathan, S., Frishman, L.J., Robson, J.G., Harwerth, R.S., and Smith, E.r.: 'The photopic negative response of the macaque electroretinogram: reduction by experimental glaucoma', Investigative ophthalmology & visual science, 1999, 40, (6), pp. 1124-1136.

[4] Frishman, L., Sustar, M., Kremers, J., McAnany, J.J., Sarossy, M., Tzekov, R., and Viswanathan, S.: 'ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram', Documenta Ophthalmologica, 2018, pp. 1-5.

[5] Tang, J., Edwards, T., Crowston, J.G., and Sarossy, M.: 'The test–retest reliability of the photopic negative response (PhNR)', Translational vision science & technology, 2014, 3, (6), pp. 1-1.

[6] Shannon, C.E.: 'A mathematical theory of communication', The Bell System Technical Journal, 1948, 27, pp. 379-423.

[7] Richman, J.S., and Moorman, J.R.: 'Physiological time-series analysis using approximate entropy and sample entropy', American Journal of Physiology-Heart and Circulatory Physiology, 2000, 278, (6), pp. H2039-H2049.

[8] Naik, G.R., Arjunan, S., and Kumar, D.: 'Applications of ICA and fractal dimension in sEMG signal processing for subtle movement analysis: a review', Australasian physical & engineering sciences in medicine, 2011, 34, (2), pp. 179-193

[9] awasaki, R., Azemin, M.C., Kumar, D., Tan, A., Liew, G., Wong, T., Mitchell, P., and Wang, J.: 'Fractal dimension of the retinal vasculature and risk of stroke: a nested case-control study', Neurology, 2011, 76, (20), pp. 1766-1767

[10] Gneiting, T., Ševčíková, H., and Percival, D.B.: 'Estimators of fractal dimension: Assessing the roughness of time series and spatial data', Statistical Science, 2012, pp. 247-277

[11] Higuchi, T.: 'Relationship between the fractal dimension and the power law index for a time series: a numerical investigation', Physica D: Nonlinear Phenomena, 1990, 46, (2), pp. 254-264

[12] Katz, M.J.: 'Fractals and the analysis of waveforms', Computers in biology and medicine, 1988, 18, (3), pp. 145-156

[13] Tan, N.Y., Friedman, D.S., Stalmans, I., Ahmed, I.I.K., and Sng, C.C.: 'Glaucoma screening: where are we and where do we need to go?', Current opinion in ophthalmology, 2020, 31, (2), pp. 91-100

[14] Glen, F.C., Baker, H., and Crabb, D.P.: 'A qualitative investigation into patients' views on visual field testing for glaucoma monitoring', BMJ open, 2014, 4, (1)

[15] Crowston, J.G., Kong, Y.X.G., Trounce, I.A., Dang, T.M., Fahy, E.T., Bui, B.V., Morrison, J.C., and Chrysostomou, V.: 'An acute intraocular pressure challenge to assess retinal ganglion cell injury and recovery in the mouse', Experimental eye research, 2015, 141, pp. 3-8

Biomedical Signal Processing and Control 71 (2022) 103164

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Empirical mode decomposition denoising of the electroretinogram to enhance measurement of the photopic negative response

Marc Sarossy [a,*], Jonathan Crowston [b], Dinesh Kumar [c], Zhichao Wu [a,d]

[a] *Ophthalmology, Department of Surgery, University of Melbourne, Melbourne, Victoria, Australia*
[b] *Duke-NUS Medical School, Singapore*
[c] *RMIT University, Melbourne, Australia*
[d] *Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia*

ABSTRACT

Background and objective: The photopic negative response (PhNR) of the electroretinogram (ERG) can be useful in the diagnosis and management of glaucoma. Its utility is limited by its large degree of measurement variability, which can be caused by low frequency oscillations and baseline wander. In this work we show a new denoising strategy using the empirical mode decomposition (EMD) which improves repeatability.

Methods: ERGs were performed on both eyes of 56 patients using red-on-blue stimulus after 2 min of preadaptation, capturing 10 sweeps per recording. Two recordings were obtained to provide intrasession test–retest estimates. Denoising was performed by removing the residue after the decomposition with the EMD, the Ensemble EMD (EEMD) and the Complete ensemble EMD with adaptive noise (CEEMDAN). The PhNR was measured as the amplitude, relative to the prestimulus baseline, of the first negative going trough after the b wave. The coefficient of repeatability (CoR) was calculated to represent 95% of the test–retest differences for the PhNR between sessions.

Results: Denoising with all of the varieties of the EMD improved the test–retest repeatability of the PhNR. The best results were achieved with the CEEMDAN algorithm, which for the PhNR amplitude improved the coefficient of repeatability relative to the effective dynamic range from 20.2% to 8.3%.

Conclusions: The EMD can be used to remove low frequency baseline drift and oscillation from the ERG and improve the test–retest repeatability of PhNR measurements. This technique could be deployed for a range of visual and other electrophysiological tests.

1. Introduction

Glaucoma is an optic neuropathy characterized by the progressive loss of retinal ganglion cells (RGCs), resulting in characteristic changes in the optic nerve appearance clinically; thinning of the neuroretinal tissue as visualised with optical coherence tomography (OCT) and changes in the visual field as measured by standard automated perimetry (SAP) – – see for example Quigley [1]. Globally, 64.3 million people between the ages of 40 to 80 are estimated to have glaucoma in 2013 [2]. In patients with glaucoma, treatment by lowering the intraocular pressure is highly effective at reducing the progressive loss of RGCs and thus irreversible vision loss [3].

One of the key diagnostic challenges for a clinician is to determine whether a given person actually has glaucoma and therefore needs treatment. This is because not all eyes with glaucoma present with abnormal visual fields, since up to 50% of RGCs would need to be lost before visual field loss can be reliably detected [4]. Furthermore, a glaucomatous appearing optic nerve is a poor predictor of future vision loss [5], underscoring the challenges in glaucoma diagnosis.

Electrophysiological tests to provide an objective assessment of RGC function could potentially offer crucial additional information in the clinical diagnosis of glaucoma. The photopic negative response (PhNR) is a late, slow going potential of the electroretinogram (ERG) that originates in the retinal ganglion cells (RGCs) of the eye [6] and is progressively diminished with an increasing severity of experimental and clinical glaucoma [7]. It is usually defined as the lowest trough in the ERG following the b-wave peak, and it is measured relative to the prestimulus baseline. Despite the ability of the PhNR to provide an objective measure of RGC function, two measurements taken within the same visit can vary by more than ± 40% of the dynamic range of this measure [8]. Such degree of measurement variability limits the utility of this technique for characterising severity of glaucoma or the detection of its progression.

There have been few papers exploring the effect of age on the PhNR. Joshi [9] found no significant correlation between the amplitude of the PhNR and age when measured from b wave peak to PhNR trough. Kato [10] similarly showed minimal correlation between age and the PhNR when the PhNR was measured from baseline to post b wave trough.

---

[*] Corresponding author at: Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia.
E-mail address: marc@sarossy.com (M. Sarossy).

https://doi.org/10.1016/j.bspc.2021.103164
Received 21 May 2021; Received in revised form 28 August 2021; Accepted 7 September 2021

One of the major sources of noise contributing to the measurement variability of the PhNR is baseline drift and wander. Baseline drift can be problematic in visual electrophysiology and can contaminate the electroretinogram (ERG) [11], the electrooculogram (EOG) [12], the pattern ERG (PERG) [13] and the Visual Evoked Potential (VEP) [14]. The baseline drift is often not random and is therefore not removed by simple signal averaging. The International Society for the Clinical Electrophysiology of Vision (ISCEV) extended protocol for the PhNR [15] does not specify a detrending algorithm to minimise measurement variability caused by baseline drift. Fig. 1 shows an ERG trace with an obvious linear trend. As the PhNR is a late response measured with respect to the prestimulus baseline, the effect of this trend introduces error into the determined value for the PhNR. This in turn reduces the repeatability and confidence of the measure. Ultimately, this will limit the utility of the measure, either as an input into a diagnostic model, or as a parameter to determine whether or not progression of the disease has occurred. Conversely, removing the component of the response caused by baseline trend could enable the use of the parameter for early diagnosis of vulnerable eyes, or improve the performance of predictive models of the disease.

Various techniques for the removal of baseline drift and wander are in common use. The simplest technique is to use an appropriate high pass filter. Filters used are usually zero-phase, and finite impulse response (FIR) [16] and Butterworth [17] implementations are often used. The ISCEV extended protocol for the PhNR states that the high pass filter should allow frequencies above 0.3 Hz to be included in the analysis [15] although a lower cut frequency is allowed. A high pass filter setting of 0.3 Hz reduces but does not eliminate baseline wander that often includes components far above this frequency [18]. Linear or polynomial curve fitting to the pre and post stimulus data has been used with some success [11] where the coefficient of repeatability was improved without a reduction in the diagnostic power.

Adaptive mode decomposition methods are an alternative approach. The concept is to decompose the signal into a finite number of components. Methods which have been described for denoising of biosignals include the Hilbert Vibrational Decomposition [19], the Variational Mode Decomposition [20] (VMD), the Empirical Wavelet Transform (EWT) [21] and the Empirical Mode Decomposition (EMD) [22]. These algorithms decompose an original signal into a linear sum of component subsignals of the same length by detection of modes. A mode corresponds to a signal which have a compactly supported Fourier spectrum [21].

In the Hilbert Vibrational Decomposition, the signal is recursive and performed iteratively on the residual signal from the previous step [23]. Utilizing, as the name suggests, the Hilbert transform (HT), this algorithm, at each iteration the dominant (most energetic) is extracted with its instantaneous frequency (IF) estimated as the average frequency from the HT. The technique has been used on simulated electrocardiogram (ECG) signals [24] but its use in visual electrophysiology has not been published.

Although the Variational Mode Decomposition also uses the Hilbert Transform and assumes the same definition of modes as the HVD, the algorithm is non recursive and decomposes the signal into a linear sum of AM-FM components known as intrinsic mode functions (IMFs). Each IMF is assumed to be mostly compact around a center frequency [25]. The decomposition is found as a solution to an optimisation problem with a quadratic penalty term and a Lagrangian multiplier. The technique has been used for the detection and removal of electro-oculogram artefacts from electroecephalography (EEG) signals [26]. No use for the electroretinogram has been published.

The Empirical Wavelet Transform (EWT) performs an adaptive discrete Meyer wavelet transform of the input signal. The algorithm proceeds by first computing the Fourier spectrum of the signal [27]. The local maxima of the spectrum are arranged in descending order of magnitude with a pre-specified number of frequencies retained. These values are then used to build the scaling and wavelet functions for the decomposition.

The Empirical Mode decomposition (EMD) proposed by Huang [22] is a decomposition of the input signal into a fixed or variable linear sum of components known as intrinsic mode functions (IMFs). The part of the signal that is left after the removal of all the IMFS is the residue, which has a maximum of one turning point and corresponds with the trend. The IMF is defined [28] as a function with equal numbers of extrema and zero crossings and the envelopes (the curve connecting all local maxima and the curve connecting all local minima) being symmetric with respect to zero.

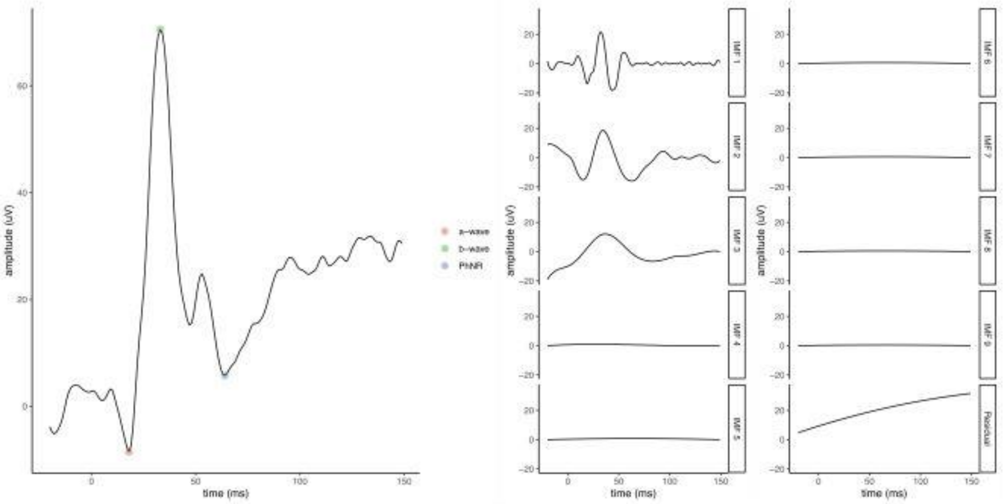


Fig. 1. Left: Sample electroretinogram (ERG) trace based on an average of ten sweeps, with markers for the a-wave, b-wave and photopic negative response (PhNR) included. Right 2 columns: Empirical Mode Decomposition (EMD) of the signal.

The EMD approach is vulnerable to a problem known as mode mixing. This is where the algorithm cannot extract two relatively closely spaced frequency components. To overcome this, the EMD has been extended as the Ensemble Empirical Mode Decomposition (EEMD) [29] and the Complete Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [30]. In both approaches, an ensemble of traces is created from the original signal by the addition of white noise. The EMD is done on the ensemble and the average IMF extracted at each level of the decomposition. The difference between the two techniques is in how the algorithm proceeds between each decomposition level. The original signal can be reproduced from the addition of the components of the CEEMDAN whereas the same does not work with the EEMD. The EMD and its variants are more intuitive than the HVD, VMD and EWT and libraries for R, Python and Matlab are easily available. We have chosen to make the EMD algorithm and its variants the subject of this paper. Denoising is achieved by subtraction of the residue from the original signal which will remove the ultra-low frequency drift and the DC component. The empirical nature of the decomposition makes no assumptions about the number of oscillating components, nor does the oscillation itself need to have any regularity.

EMD denoising has been used to remove both high frequency and low frequency noise from electrocardiogram (ECG) [28] in which the baseline wander was eliminated. Another paper [31] demonstrated that this improved the sensitivity of detecting the R wave. The same principles have been applied to a variety of other fields, such as the forecasting of commuter flows in public transport [32]. The algorithm can be tuned by selection of the stopping criteria and the number of IMFs extracted. A recent paper [33] showed the CEEMDAN to outperform both the EMD and EEMD when used for denoising the ECG. There have been few papers comparing the techniques for baseline wander correction. Singh [34] found EMD and EWT to be similar as denoisers for the ECG. Although Sharma [24] found the HVD to be superior to the EMD for baseline wander in the ECG, this study used simulated data and did not explore the ensemble methods of the EMD. Maji [35] compared the performance of the VMD and EMD in the decomposition of the ECG and found the EMD to perform better for low frequencies than the VMD.

We thus sought to examine in this study whether these EMD approaches, including the EEMD and CEEMDAN, could be used to minimise the impact of baseline drift and improve the measurement repeatability of the PhNR in a real-world unsimulated study. We recruited control subjects without ocular diseases to show the effect of these EMD approaches on the PhNR repeatability. Further work will be needed to explore the correlation of the denoised signal with the severity of the disease and how that compares to the non-denoised response and to compare the EMD variants to the other state-of-the-art methods discussed.

As the intention of the decomposition is to remove the ultra-low frequency trend from the signal and to remove the DC component, the residue is all that is required for denoising. The decomposition itself can either stop after automated detection of residue status, or it can proceed for a fixed number of iterations. The remaining text of this paper is organized under the following sections. Section 2 discusses the EMD, EEMD and CEEMDAN in detail including specific details of the implementation. Section 3 discusses the methodology in detail. The experimental results are presented in Section 4. The discussions on the results are outlined in Section 5. In Section 6, the paper is concluded with some suggested limitations and future works.

2. EMD, EEMD and CEEMDAN

The EMD decomposes the signal into a linear sum of IMFS and a residue. To be an IMF the function must fulfill two conditions [36] : the number of extrema (maxima and minima) and zero crossings must not differ by more than one and the mean of the upper and lower envelopes must be zero. The algorithm as implemented in Libeemd [37] for a signal × length n can be described as follows:

```
Set $k = 0$ and $r_0 = x$
Set first minimum and first maximum to $x_1$ and $S = 0$
Find all extrema as each point where the slope changes sign – in the case of flat regions, use the center point of the flat region as the extremum. For $p = 1, \cdots, n$
    If $x_p > x_{p-1}$ and $x_{p+1} < x_p$ then find $(p, x_p)$ as a maximum
        If $x_{p-1} > x_{p-2}$ and $x_p = x_{p-1}$ and $x_{p+1} < x_p$ then find $(p - 0.5, x_p)$ as a maximum
    If $x_p < x_{p-1}$ and $x_{p+1} > x_p$ then find $(\mathrm{k}, \mathrm{x}_\mathrm{p})$ as a minimum
        If $x_{p-1} < x_{p-2}$ and $x_p = x_{p-1}$ and $x_{p+1} > x_p$ then find $(p - 0.5, x_p)$ as a minimum
Add the other end of the data as extrema as well
If there are at least two interior extrema, test if linear extrapolation provides a more extremal value than the first and last extrema -if so use the more extremal values
For the sets of minima (maxima) of $r_k$ obtain the lower (upper) envelope $e_{min}(e_{max})$ by:
    If 4 or more extrema exist , interpolate with cubic spline with not-a-node end condition using method of Englen-Müllges [38]
    If 3 extrema exist, use polynomial interpolation using the GNU scientific library [39]
    If 2 extrema exist use linear interpolation
Compute the mean envelope $m = \frac{(e_{min} + e_{max})}{2}$
Compute the IMF candidate $d_{k+1} = r_k - m$
Is$d_k = 1$ an IMF ?
    Yes –
        Is $> 3$ ?
            Yes . Save $d_{k+1}$ , compute residue $\mathrm{r}_{k+1} = x - \sum_{i=1}^{k} d_i$ do $k = k+1$ and treat $r_k$ as input data in Step 2
            No. If the number of maxima, minima and zero crossings are the same as previous iteration, Set$S = S + 1$, treat $d_{k+1}$ as input data in Step 2
    No. treat $d_{k+1}$ as input data in step 2
Continue for $num\_imfs - 1$ iterations where $num\_imfs$ is the number of IMFs to be extracted
```

The refinement process (steps 2 to 9) is known as the sifting process. The S number is the stability criterion for ceasing the sifting. In this work it was set to 4 – the number of sifting iterations without change in the number of maxima, minima or zero crossings

The ensemble EMD (EEMD) defines modes as the average of corresponding IMFs obtained from an ensemble of the original signal plus different realizations of finite variance white noise [36]. The Rlibeemd implementation of the algorithm is as follows

```
Generate $x^{(i)} = x + \beta w^{(i)}$ where $w^{(i)} (i = 1, \cdots, I)$ is a zero mean unit variance Gaussian white noise realization, I is the ensemble size, $\beta > 0$ is the noise strength parameter. β is expressed relative to the standard deviation of the noise of the signal. The white noise is generated via the GNU scientific library [39]
Decompose completely each $x^{(i)} (i = 1, \ldots I)$ by EMD as above, obtaining the modes $d_k^{(i)}$ where $k = 1, \ldots, K$ indicates the mode
Assign $\overline{d}_k = \frac{1}{I} \sum_{i=1}^{I} d_k^{(i)}$
```

As each $x^{(i)}$ is decomposed independently, there is no connection between the residues at each stage and thus the EEMD is not complete and suffers from reconstruction error.

The CEEMDAN addresses the problems of the EEMD by recombining the signal and noise after every IMF extraction. It defines $E_k(\cdot)$ as an operator producing the kth mode by EMD. Then set $w^{(i)}$ as a realization of zero mean unit variance Gaussian white noise.

1. For every $I = 1 \cdots, I$ decompose each $x^{(i)} = x + \beta_i w^{(i)}$ by EMD until the first mode and obtain
2. $\overline{d}_1 = \frac{1}{I}\sum_{i=1}^{I} d_1^{(i)}(1)$
3. At the first stage (k = 1) calculate the residue as $r_1 = x - \overline{d_1}$
4. Obtain the second mode as
5. $\overline{d}_2 = \frac{1}{I}\sum_{i=1}^{I} E_1\left(r_1 + \beta_1 E_1\left(w^{(i)}\right)\right)(2)$
6. For$(k = 2, \ldots, K)$calculate the kth residue $r_k = r_{(k-1)} - \overline{d_k}$
7. Obtain the $(k+1)$th mode as
8. $\overline{d}_{k+1} = \frac{1}{I}\sum_{i=1}^{I} E_1\left(r_k + \beta_k E_k\left(w^{(i)}\right)\right) \quad (3)$
9. Go to step 4 for the next k

Iterate steps 4 to 6 K times.

Note that the β used in equations 2 and 3 is set by $\beta_k = \varepsilon_k \beta std(r_k)$ where ϵ_k is the noise strength at each stage. Rlibeemd does not provide for setting a per-stage noise strength so in this implementation it simplifies to $\beta_k = \varepsilon\, std(r_k)$.

The reader is encouraged to compare the 3 algorithms as implemented in Rlibeemd with the method described in Colominas [36]. The differences include the fixed number of iterations, the use of the S number for detecting the end of the sifting operation, the setting of the terminal extrema, the envelope interpolation and the parameterization of the noise level. Other variations have been described including dealing with the boundary conditions by mirrorizing the extrema close to the edges. A more complex method is to add a characteristic wave to each end of the signal. This characteristic wave is a sinusoid extended three periods beyond a determined data span at each end of the original signal. The envelope function itself is usually a cubic spline interpolation but other functions including higher order [40] and non-polynomial [41] splines have been described.

3. Methods

This research followed the tenets of the Declaration of Helsinki and was approved by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital. Healthy control participants without self-reported eye disease were recruited from the staff of the hospital and the staff of the Centre for Eye Research Australia (CERA). Participants without eye disease were chosen for this study in order to characterize the dynamic range of the response and therefore to normalize the repeatability measure. Informed consent was obtained for all participants.

3.1. Electroretinography

Electroretinograms were collected as per the ISCEV extended protocol [8]. Briefly, pupils were dilated to at least 7 mm using 1% tropicamide. Eyes were preadapted for at least 1 min with background room light (of 0.92 cd/m^2). An Espion system (E2/ColorDome; Diagnosys LLC, Lowell, MA, USA) was used for stimulus generation and data acquisition. Brief preadaptation to the blue background of approximately 1 min was performed before the first measurement. Data was collected with a sampling rate of 1kHz from 20 ms prestimulus to 149 ms post stimulus. The equipment has a 16 bit analog to digital converter.

Red (peak wavelength, 635 nm) stimuli of 4 ms duration at 1 cd·s/m^2 were delivered via a Ganzfeld sphere on a blue (peak wavelength 465 nm) background of 10 cd/m^2, at 1 Hz. The response was recorded using a Dawson-Trick-Litzkow (DTL) fiber electrode placed inside the lower lid conjunctival fornix of each eye. The ground electrode was attached to the forehead and the reference electrode was attached to the lateral canthus.

Ten sweeps were collected and the average also calculated. Signals were filtered from 0.15 to 100 Hz. An automatic rejection system removed large artefacts secondary to blink and eye movements so the total sweeps collected was 10 non rejected sweeps. The protocol was repeated once during the same session to enable assessment of intra-session test–retest repeatability. In this study ERG recordings were obtained from both eyes of each participant.

3.2. Signal Analysis

Data was exported from the electrophysiology system and imported into R. EMD, EEMD and CEEMDAN analysis were with the R package Rlibeemd [42]. Denoising consisted of removing the final residual of the EMD decomposition from each trace and then finding the ensemble average. The comparator was the simple arithmetic average of the ensemble of traces. Ten IMFs (including the residue) were specified to minimize the signal in the final residue. Parameters used for the algorithm were based on previously published work [36,43–45]. An S Number of 4 was chosen to ensure stability of each IMF extraction. For the EEMD and CEEMDAN, the ensemble size was set to 250 with a Gaussian White Noise strength of 0.2 relative to the standard deviation of the input signal These parameters were based on previously published works [46,47].

After denoising, the b wave was identified as the largest positive amplitude value. The a wave was identified as the lowest amplitude value between the stimulus onset and the b wave. B wave amplitude is defined as the b wave peak relative to the a wave trough [48]. Prestimulus baseline was determined to be an average of the prestimulus amplitudes. The PhNR was identified as the lowest amplitude after the b wave peak and measured with respect to the prestimulus baseline. The left hand pane of Fig. 1 shows a sample ERG ensemble average with automatic placement of the a-wave, b-wave and PhNR markers. We also calculated the PhNR/B ratio, which we observed to reduce the levels of measurement variability previously [8]. The PhNR/B ratio is derived by dividing (i) b-wave peak to PhNR trough amplitude, by (ii) a-wave trough to b-wave peak amplitude. The 2 right hand panes show the EMD decomposition of the waveform in Fig. 1The residue is what is removed from each trace as the denoising strategy.

The overall algorithm is therefore as follows:

1. Perform EMD (or variant) on the signal with 9 IMFS and residue extracted
2. Obtain the denoised signal by subtracting the residue from the original signal
3. Find the b wave peak as the maximum of the denoised signal
4. Find the a wave trough as the minimum of the interval between the stimulus onset and the b wave peak
5. Find the PhNR trough as the minimum of the interval between the b wave peak and the end of the signal
6. Determine the prestimulus baseline amplitude as the average of the prestimulus signal
7. Obtain the b wave amplitude as the difference between the a wave trough and b wave peak
8. Obtain the a wave amplitude as the difference between the prestimulus baseline and the a wave trough
9. Obtain the PhNR amplitude as the difference between the prestimulus baseline and the PhNR trough

3.3. Statistics

The test–retest repeatability of the different methods were analysed by comparing the coefficients of repeatability (CoR) as per our earlier paper [8]. The CoR is the value below which the absolute differences between the repeated measurements falls with a 95% probability [49]. It is calculated by multiplying the within-subject standard deviation by $1.96\sqrt{2}$. It is an absolute figure in the same units as the measurement. With the PhNR measured in µV, the CoR is the number of microvolts absolute difference that contains 95% of test–retest pairs. To compare the CoR figures between different measures such as the PhNR in µV vs the PhNR/B ratio (which is unitless), the CoR must be normalized. One way to do this is to express the CoR relative to the mean of the measure. However, the mean value may not represent the effective dynamic range (EDR) of the parameter [8]. As in our previous paper, we have normalized the CoR to the EDR.

For each EMD method, the CoR of the PhNR and PhNR/B ratio was compared to the base case (of no extra denoising – only the default detrending included in the manufacturer's software). Significance of the differences in the CoR were determined by bootstrap replication ($n = 1000$ resamples). Replicates of the sample set were made by resampling with replacement, performed at the subject rather than eye level to account for correlations between eyes of the same subject. For each EMD method, the CoR was compared to the base case (of no extra denoising – only the default detrending included in the manufacturer's software).

4. Results

A total of 112 eyes of 56 healthy participants were tested, and these participants were on average 33 years old (range, 20 to 73 years old).

Fig. 2 demonstrates the denoising for an eye of a participant over two recording sessions illustrated over the first two row with each row a session. These graphs show the denoising for each sweep and the average. The third row shows the averages for the two sessions superimposed. The CEEMDAN technique shows close alignment of both baseline and PhNR trough. Fig. 3 shows two sessions superimposed for a further 3 eyes of a different 3 subjects. Again, the alignment of the PhNR region of the signal is improved by all denoising techniques with the best performance by the CEEMDAN

Table 1 shows the marker amplitudes and standard deviations for all eyes, and the CoR between the recordings are shown for the PhNR and the PhNR/B ratio for each denoising method. In the upper part of the table, it can be observed that the denoising technique significantly changed the amplitudes measured for all ERG parameters. This indicates that to compare measurements over time, for example to detect disease progression the same denoising technique must be used for all sessions. The lower part of the table shows the effect of intersession repeatability in this cohort (which has no change in the underlying physiology between sessions). Recordings denoised by the CEEMDAN method showed the lowest CoR (and hence highest repeatability) for both the PhNR and the PhNR/B ratio measures, and for both the absolute and normalized measures ($P < 0.05$ compared to all other methods for all comparisons of the normalized measure).

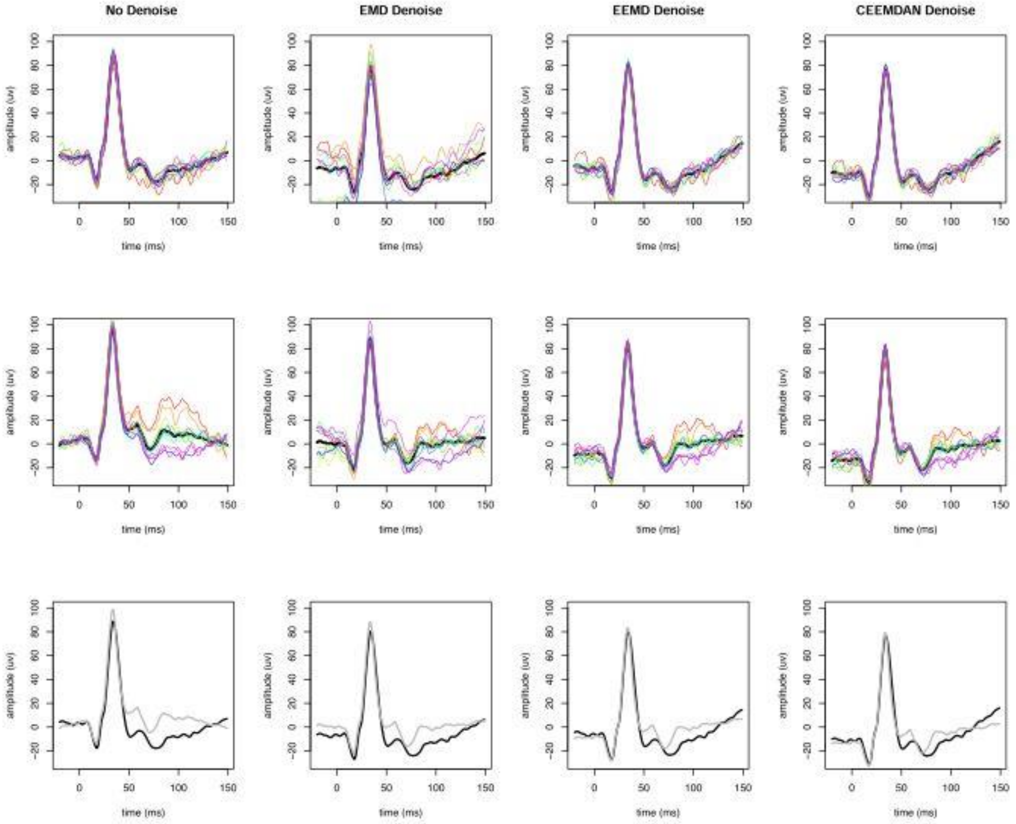


Fig. 2. Comparison of raw ("No Denoise") vs denoised data using empirical mode decomposition (EMD), Ensemble EMD (EEMD) and Complete Ensemble EMD with adaptive noise (CEEMDAN) methods for an example eye over two sessions. Coloured traces are the individual sweeps, and the bold black line is the ensemble average. The first row shows the first recording session, and the second row shows the second session. The third row shows the average from the two recording sessions superimposed. The PhNR and prestimulus baseline are most closely aligned with the CEEMDAN method.

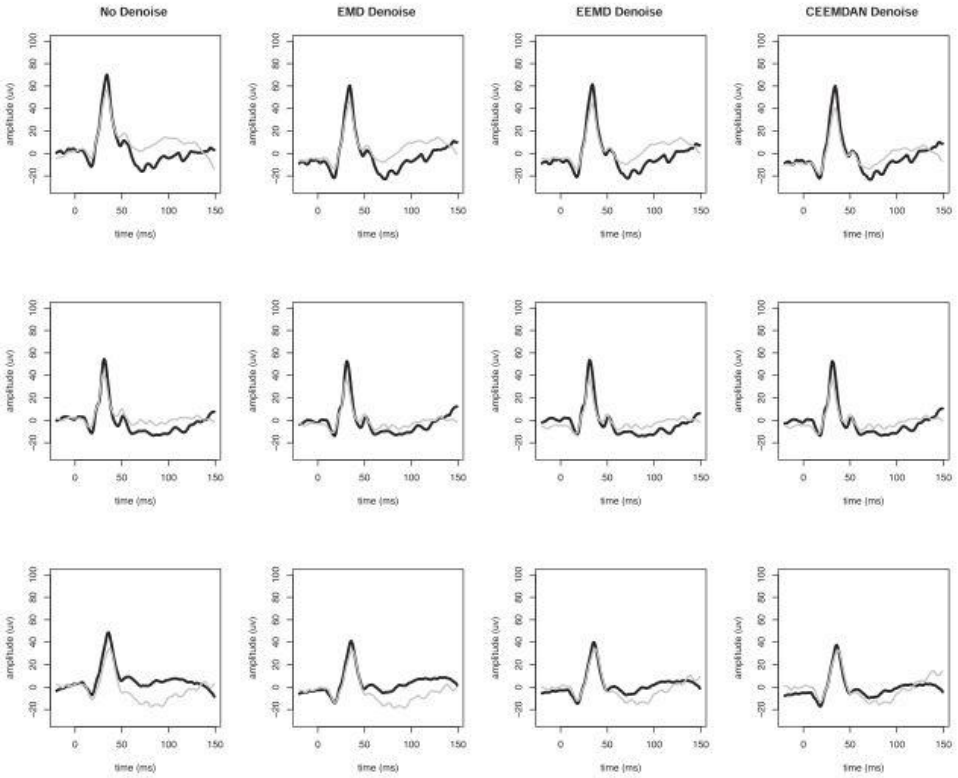


Fig. 3. Comparison of raw ("No Denoise") vs denoised data using empirical mode decomposition (EMD), Ensemble EMD (EEMD) and Complete Ensemble EMD with adaptive noise (CEEMDAN) methods for three example eyes with the net effect of the denoising strategy shown by superimposing the denoised traces for the two sessions. Each row represents a different eye.

5. Discussion

This study demonstrates the effectiveness of a novel and efficient technique for the removal of very low frequency noise from the ERG to improve its measurement repeatability. Low frequency noise, which manifests as baseline drift and wander, is a particular problem for the PhNR response of the ERG, as the response is slow and occurs late after the stimulus.

Although high frequency and stationary noise such as powerline and radiofrequency noise from switchmode power supplies is amenable to conventional filtering with analog or digital filters, very low frequency noise such as DC trend and blink artefact is much more problematic. The PhNR is typically referenced to the prestimulus baseline and a linear trend can have a large influence on the measured magnitude of the response. We have previously shown [8,50] that referencing the PhNR to the b-wave peak or to a ratio relative to the a- and b-wave amplitudes can partially offset this, and improve the repeatability of this measure. However, this adds extra unwanted variance into the signal (even if glaucoma is expected to only affect the PhNR and not the a- and b-waves), resulting in a combination of potentially unrelated physiological processes.

The photomyoclonic response [51] is a response measured via the ERG caused by involuntary blinking and induced by the stimulus flash, which can also contribute to the measurement variability of the PhNR. It is partially time locked to the stimulus, has a latency of around 60 ms and habituates with repeated flashes. There is high inter- and intra-subject variability of this response. These characteristics make the photomyoclonic response difficult to remove with conventional filtering. The fact that the PhNR is within the latency for the photomyoclonic response and blink artefact is highly problematic. An adaptive per-trace filter, such as the EMD and its variants, is an optimal solution for this. The improved repeatability observed with the EMD approaches used in this study likely occurred because they minimized the impact of this photomyoclonic response.

Detrending algorithms exist and are often included in the electrophysiological hardware, but usually use a simple linear algorithm. The elegance of the EMD family of algorithms is that they use the whole response, including the actual ERG, and do not assume a structure to the noise. By visual inspection of the responses, all the EMD techniques were highly effective at improving the alignment of the ERG traces beyond the b-wave. The ensemble techniques of the EMD are known to provide better resilience to mode mixing and it may be that the baseline wander noise and the ERG signal itself are sufficiently close to suffer from this problem, accounting for the better performance of the ensemble techniques. In Fig. 1, the decomposition is essentially complete by the 3rd IMF with the remaining iterations not changing the decomposition or the residual. For the EEMD, this typically took an additional 2 IMFs before stability. The recombination of signal and noise in step 3 of the

Table 1

Results for electroretinography (ERG) measurements for all 112 eyes of 56 patients. The upper part of the table shows the cohort means for both sessions. The lower part of the table shows the effect of denoising on the coefficient of repeatability between the two sessions.

	Raw	EMD	EEMD	CEEMDAN
Cohort Mean (SD)				
a-wave (μV)	−10.6 (6.5)	$-14.5^{\#}$ (7.9)	$-14.9^{\#}$ (7.4)	$-16.2^{\#}$ (7.8)
b-wave (μV)	62.4 (24.5)	$62.6^{\#}$ (24.7)	$62.6^{\#}$ (24.7)	$62.5^{\#}$ (24.6)
PhNR (μV)	−14.9 (13.9)	$-16.7^{\#}$ (8.8)	$-17.3^{\#}$ (8.6)	$-18.2^{\#}$ (7.7)
PhNR/B	0.94 (0.16)	$0.97^{\#}$ (0.14)	$0.97^{\#}$ (0.13)	$0.97^{\#}$ (0.13)
Coefficient of Repeatability				
PhNR absolute (μV)	3.0	$1.9^{\#}$	$1.8^{\#}$	$1.5^{\#}$
PhNR/B absolute	0.04	0.04	0.03	0.03
PhNR normalised (%)*	20.2%	11.3%$^{\#}$	10.3%$^{\#}$	8.3%$^{\#\dagger}$
PhNR/B normalised (%)*	19.5%	13.6%$^{\#}$	13.1%$^{\#}$	11.3%$^{\#\dagger}$

Notes: EMD = empirical mode decomposition; EEMD = ensemble EMD (EEMD); CEEMDAN = complete ensemble EMD with adaptive noise; PhNR = photopic negative response, relative to the prestimulus baseline; PhNR/B = PhNR to b-wave amplitude divided by the a- to b-wave amplitude; * = relative to effective dynamic range; $^{\#}$ = significantly different at $P < 0.01$ compared to the traces without denoising; † = significantly different at $P < 0.05$ compared to the EMD and EEMD methods.

CEEMDAN algorithm leads to a decomposition that does not finish with a series of near identical IMFs and this may explain the better performance of this particular algorithm.

The translational significance of this work is in the demonstration of the technique as an effective and reliable one for detrending and ultra-low frequency denoising for visual electrophysiology. Removing this major artefact improves the measurement repeatability, and ultimately the utility of measurements of the PhNR for glaucoma. An evaluation of the diagnostic performance of the PhNR with and without such denoising is now warranted to understand the clinical relevance of this level of improvement in measurement repeatability. Further, the technique should be of utility in other similar settings such as the PERG or EOG. We envisage that 'detrend by EMD' could become a selectable option on electrophysiology equipment that the operator can use on a case-by-case basis depending on the protocol and the amount of noise present at the time. That stated, extreme care would be needed to compare serial traces where one was filtered, and one was not. For this reason, the technique is much better implemented *post-hoc*, where a decision is made to apply it to all or none of the traces. In the current implementation of the Espion software, the detrending (called drift removal) is an option specified in the protocol for a given step and channel, and it is not reversible. In addition to potentially improving the clinical utility of the test, the use of denoising strategies examined in this study may also allow one to reduce the number of sweeps required to achieve robust responses. In doing so, it may reduce the risk of averaging a non-stationary response.

Limitations of this study include the small number of traces averaged, and the short prestimulus acquisition epoch. An evaluation of the performance of the denoising strategies for recordings obtained with a much larger number of traces would provide valuable insights into the extent that the EMD denoising approaches improve measurement repeatability as the number of sweeps are increased. For future studies, we would recommend a larger number of sweeps and a longer pre- and post-stimulus acquisition epoch to improve the detrending performance. Further work could include a simulation study to quantify the noise reduction, optimization of various tuneable parameters of the empirical mode decomposition and comparison with other similar methods such as the vibrational mode decomposition and the empirical wavelet decomposition.

6. Conclusion

This study performed post-hoc analysis on test–retest ERG recordings and showed a significant improvement in repeatability of PhNR measurements when the individual traces were filtered with an EMD denoising strategy, with the CEEMDAN approach emerging as the optimal method. These EMD denoising strategies will help improve the clinical utility of the PhNR, and it can be deployed to improve the measurement repeatability of a range of other electrophysiological and visual tests.

CRediT authorship contribution statement

Marc Sarossy: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - review & editing. **Jonathan Crowston:** Conceptualization, Resources, Supervision, Project administration, Funding acquisition, Writing - review & editing. **Dinesh Kumar:** Conceptualization, Supervision, Writing - review & editing. **Zhichao Wu:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] H.A. Quigley, 21st century glaucoma care, Eye 33 (2) (2019) 254–260.

[2] Y.-C. Tham, X. Li, T.Y. Wong, H.A. Quigley, T. Aung, C.-Y. Cheng, Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis, Ophthalmology 121 (11) (2014) 2081–2090.

[3] A. Heijl, et al., Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial, Arch. Ophthalmol. 120 (10) (2002) 1268–1279.

[4] R.S. Harwerth, J.L. Wheat, M.J. Fredette, D.R. Anderson, Linking structure and function in glaucoma, Progress in retinal and eye research 29 (4) (2010) 249–271.

[5] F.A. Medeiros, et al., Prediction of functional loss in glaucoma from progressive optic disc damage, Arch. Ophthalmol. 127 (10) (2009) 1250.

[6] S. Viswanathan, et al., The photopic negative response of the macaque electroretinogram: reduction by experimental glaucoma, Invest. Ophthalmol. Vis. Sci. 40 (6) (1999) 1124–1136.

[7] Z. Wu, X. Hadoux, J.C. Fan Gaskin, M.G. Sarossy, J.G. Crowston, Measuring the photopic negative response: viability of skin electrodes and variability across disease severities in glaucoma, Translational vision science & technology 5 (2) (2016) 13, https://doi.org/10.1167/tvst.5.2.13.

[8] Z. Wu, X. Hadoux, F. Hui, M.G. Sarossy, J.G. Crowston, Photopic negative response obtained using a handheld electroretinogram device: determining the optimal measure and repeatability, Translational vision science & technology 5 (4) (2016) 8, https://doi.org/10.1167/tvst.5.4.8.

[9] N.R. Joshi, E. Ly, S. Viswanathan, Intensity response function of the photopic negative response (PhNR): effect of age and test–retest reliability, Doc. Ophthalmol. 135 (1) (2017) 1–16.

[10] K. Kato, A. Sugawara, R. Nagashima, K. Ikesugi, M. Sugimoto, M. Kondo, Factors affecting photopic negative response recorded with reteval system: Study of young healthy subjects, Translational Vision Science & Technology 9 (9) (2020) 19, https://doi.org/10.1167/tvst.9.9.19.

[11] J. Tang, F. Hui, M. Coote, J.G. Crowston, X. Hadoux, Baseline Detrending for the Photopic Negative Response, Translational vision science & technology 7 (5) (2018) 9, https://doi.org/10.1167/tvst.7.5.9.

[12] M.G. Sarossy, M.H.A. Lee, M. Bach, A fast automated method for calculating the EOG Arden ratio, Doc. Ophthalmol. 128 (3) (2014) 169–178.

[13] S.L. Graham, et al., Pattern electroretinograms from hemifields in normal subjects and patients with glaucoma, Invest. Ophthalmol. Vis. Sci. 35 (9) (1994) 3347–3356.

[14] R. Palaniappan, P. Raveendran, Recursive digital filter for fast visual evoked potential estimation and classification, Electron. Lett. 37 (15) (2001) 990–992.

[15] L. Frishman, M. Sustar, J. Kremers, J.J. McAnany, M. Sarossy, R. Tzekov, S. Viswanathan, ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram, Doc. Ophthalmol. 136 (3) (2018) 207–211.

[16] K.R. Alexander, A. Raghuram, A.S. Rajagopalan, Cone phototransduction and growth of the ERG b-wave during light adaptation, Vision Res. 46 (22) (2006) 3941–3948.

[17] J. Meklenburg, E.A. Clancy, R. Tzekov, Signal processing techniques for oscillatory potential extraction in the electroretinogram: automated highpass cutoff frequency estimation, Doc. Ophthalmol. 125 (2) (2012) 101–111.

[18] H. Kundra, J.C. Park, J.J. McAnany, Comparison of photopic negative response measurements in the time and time-frequency domains, Doc. Ophthalmol. 133 (2) (2016) 91–98.

[19] M. Feldman, Time-varying vibration decomposition and analysis based on the Hilbert transform, J. Sound Vib. 295 (3-5) (2006) 518–530.

[20] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2014) 531–544.

[21] J. Gilles, Empirical wavelet transform, IEEE Trans. Signal Process. 61 (16) (2013) 3999–4010.

[22] Huang, N.E., et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. in Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences. 1998. The Royal Society.

[23] M. Civera, C. Surace, A comparative analysis of signal decomposition techniques for structural health monitoring on an experimental benchmark, Sensors 21 (5) (2021) 1825.

[24] H. Sharma, K.K. Sharma, Baseline wander removal of ECG signals using Hilbert vibration decomposition, Electron. Lett. 51 (6) (2015) 447–449.

[25] Z. Feng, D. Zhang, M.J. Zuo, Adaptive mode decomposition methods and their applications in signal analysis for machinery fault diagnosis: a review with examples, IEEE Access 5 (2017) 24301–24331.

[26] C. Dora, P.K. Biswal, An improved algorithm for efficient ocular artifact suppression from frontal EEG electrodes using VMD, Biocybernetics and Biomedical Engineering 40 (1) (2020) 148–161.

[27] S. Boda, M. Mahadevappa, P.K. Dutta, A hybrid method for removal of power line interference and baseline wander in ECG signals using EMD and EWT, Biomed. Signal Process. Control 67 (2021), 102466.

[28] M. Blanco-Velasco, B. Weng, K.E. Barner, ECG signal denoising and baseline wander correction based on the empirical mode decomposition, Comput. Biol. Med. 38 (1) (2008) 1–13.

[29] ZHAOHUA WU, NORDEN.E. HUANG, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Advances in adaptive data analysis 01 (01) (2009) 1–41.

[30] M.E. Torres, et al., A complete ensemble empirical mode decomposition with adaptive noise, IEEE, 2011.

[31] Li, H., et al., Denoising and R-peak detection of electrocardiogram signal based on EMD and improved approximate envelope. 2014. 33(4): p. 1261-1276.

[32] Y.u. Wei, M.-C. Chen, Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks, Transportation Research Part C: Emerging Technologies 21 (1) (2012) 148–162.

[33] Teja, K., R. Tiwari, and S. Mohanty. Adaptive denoising of ECG using EMD, EEMD and CEEMDAN signal processing techniques. in Journal of Physics: Conference Series. 2020. IOP Publishing.

[34] O. Singh, R.K. Sunkaria, ECG signal denoising via empirical wavelet transform, Australas. Phys. Eng. Sci. Med. 40 (1) (2017) 219–229.

[35] U. Maji, S. Pal, Empirical mode decomposition vs. variational mode decomposition on ECG signal processing: A comparative study, IEEE, 2016.

[36] M.A. Colominas, G. Schlotthauer, M.E. Torres, Improved complete ensemble EMD: A suitable tool for biomedical signal processing, Biomed. Signal Process. Control 14 (2014) 19–29.

[37] P.J.J. Luukko, J. Helske, E. Räsänen, Introducing libeemd: A program package for performing the ensemble empirical mode decomposition, Comput. Statistics 31 (2) (2016) 545–557.

[38] Engeln-Müllges, G. and F. Uhlig, Numerical algorithms with C. 2013: Springer Science & Business Media.

[39] Gough, B., GNU scientific library reference manual. 2009: Network Theory Ltd.

[40] Y. Kopsinis, S. McLaughlin, Improved EMD using doubly-iterative sifting and high order spline interpolation, EURASIP Journal on Advances in Signal processing 2008 (1) (2008) 128293, https://doi.org/10.1155/2008/128293.

[41] P. Singh, et al., Nonpolynomial spline based empirical mode decomposition, IEEE, 2013.

[42] J.L. Helske, Perttu, Ensemble empirical mode decomposition (EEMD) and its complete variant (CEEMDAN), Rlibeemd, 2017.

[43] P. Flandrin, P. Goncalves, G. Rilling, Detrending and denoising with empirical mode decompositions, IEEE, 2004.

[44] H. Wang, Z. Liu, Y. Song, X. Lu, Ensemble EMD-based signal denoising using modified interval thresholding, IET Signal Proc. 11 (4) (2017) 452–461.

[45] Marcelo.A. Colominas, Gastón Schlotthauer, María.E. Torres, Patrick Flandrin, Noise-assisted EMD methods in action, Advances in Adaptive Data Analysis 04 (04) (2012) 1250025, https://doi.org/10.1142/S1793536912500252.

[46] J.A. Mucarquer, P. Prado, M.-J. Escobar, W. El-Deredy, M. Zanartu, Improving EEG muscle artifact removal with an EMG array, IEEE Trans. Instrum. Meas. 69 (3) (2020) 815–824.

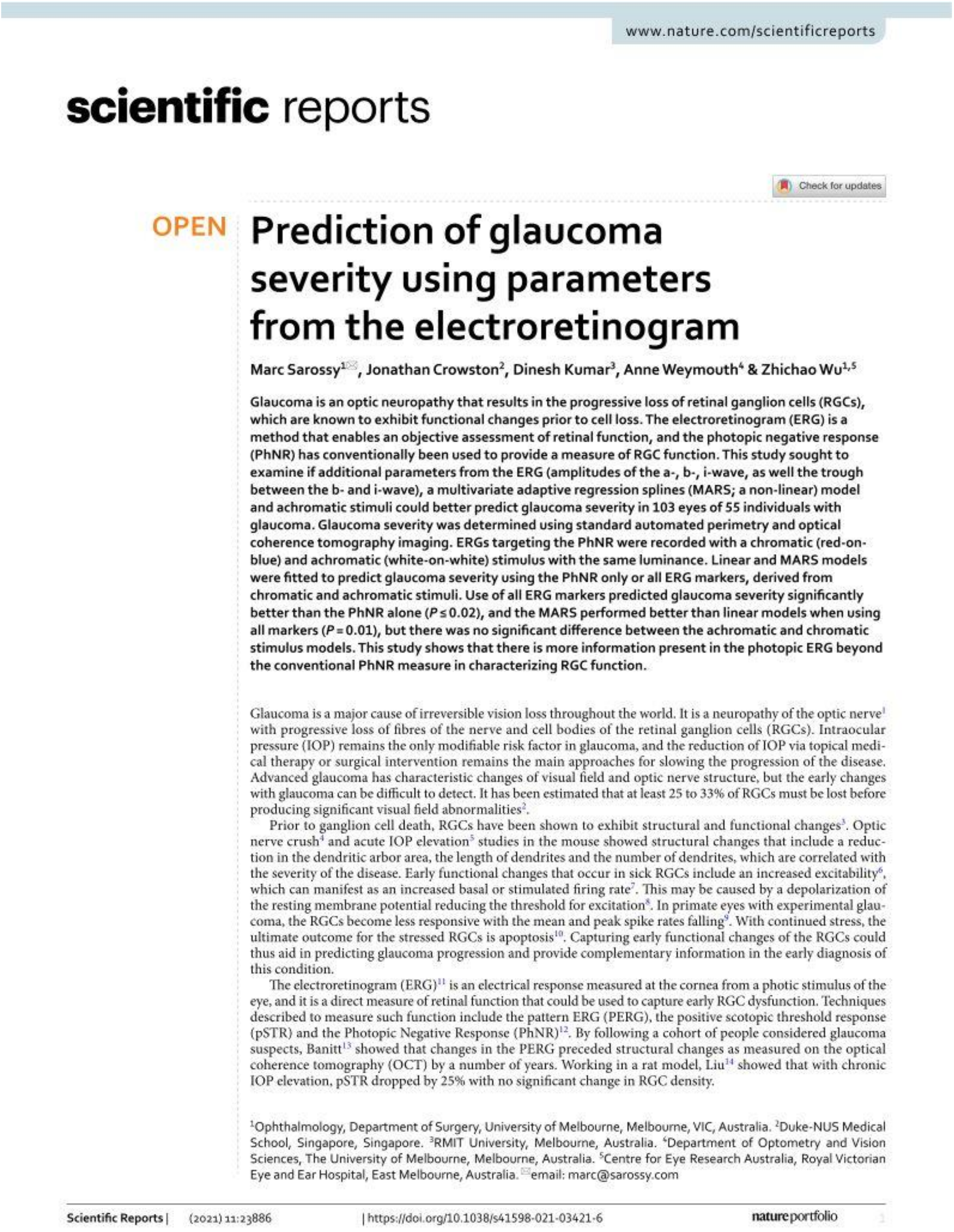
[47] L. El Bouny, M. Khalil, A. Adib, ECG signal filtering based on CEEMDAN with hybrid interval thresholding and higher order statistics to select relevant modes, Multimedia Tools and Applications 78 (10) (2019) 13067–13089.

[48] D.L. McCulloch, M.F. Marmor, M.G. Brigell, R. Hamilton, G.E. Holder, R. Tzekov, M. Bach, ISCEV Standard for full-field clinical electroretinography (2015 update), Doc. Ophthalmol. 130 (1) (2015) 1–12.

[49] S. Vaz, T. Falkmer, A.E. Passmore, R. Parsons, P. Andreou, S. Hempel, The case for using the repeatability coefficient when calculating test–retest reliability. 8 (9) (2013) e73990.

[50] J. Tang, T. Edwards, J.G. Crowston, M. Sarossy, The test–retest reliability of the photopic negative response (PhNR), Translational vision science & technology 3 (6) (2014) 1, https://doi.org/10.1167/tvst.3.6.1.

[51] M.A. Johnson, R.W. Massof, The photomyoclonic reflex: an artefact in the clinical electroretinogram, Br. J. Ophthalmol. 66 (6) (1982) 368–378.



www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Prediction of glaucoma severity using parameters from the electroretinogram

Marc Sarossy[1]✉, Jonathan Crowston[2], Dinesh Kumar[3], Anne Weymouth[4] & Zhichao Wu[1,5]

Glaucoma is an optic neuropathy that results in the progressive loss of retinal ganglion cells (RGCs), which are known to exhibit functional changes prior to cell loss. The electroretinogram (ERG) is a method that enables an objective assessment of retinal function, and the photopic negative response (PhNR) has conventionally been used to provide a measure of RGC function. This study sought to examine if additional parameters from the ERG (amplitudes of the a-, b-, i-wave, as well the trough between the b- and i-wave), a multivariate adaptive regression splines (MARS; a non-linear) model and achromatic stimuli could better predict glaucoma severity in 103 eyes of 55 individuals with glaucoma. Glaucoma severity was determined using standard automated perimetry and optical coherence tomography imaging. ERGs targeting the PhNR were recorded with a chromatic (red-on-blue) and achromatic (white-on-white) stimulus with the same luminance. Linear and MARS models were fitted to predict glaucoma severity using the PhNR only or all ERG markers, derived from chromatic and achromatic stimuli. Use of all ERG markers predicted glaucoma severity significantly better than the PhNR alone ($P \leq 0.02$), and the MARS performed better than linear models when using all markers ($P = 0.01$), but there was no significant difference between the achromatic and chromatic stimulus models. This study shows that there is more information present in the photopic ERG beyond the conventional PhNR measure in characterizing RGC function.

Glaucoma is a major cause of irreversible vision loss throughout the world. It is a neuropathy of the optic nerve[1] with progressive loss of fibres of the nerve and cell bodies of the retinal ganglion cells (RGCs). Intraocular pressure (IOP) remains the only modifiable risk factor in glaucoma, and the reduction of IOP via topical medical therapy or surgical intervention remains the main approaches for slowing the progression of the disease. Advanced glaucoma has characteristic changes of visual field and optic nerve structure, but the early changes with glaucoma can be difficult to detect. It has been estimated that at least 25 to 33% of RGCs must be lost before producing significant visual field abnormalities[2].

Prior to ganglion cell death, RGCs have been shown to exhibit structural and functional changes[3]. Optic nerve crush[4] and acute IOP elevation[5] studies in the mouse showed structural changes that include a reduction in the dendritic arbor area, the length of dendrites and the number of dendrites, which are correlated with the severity of the disease. Early functional changes that occur in sick RGCs include an increased excitability[6], which can manifest as an increased basal or stimulated firing rate[7]. This may be caused by a depolarization of the resting membrane potential reducing the threshold for excitation[8]. In primate eyes with experimental glaucoma, the RGCs become less responsive with the mean and peak spike rates falling[9]. With continued stress, the ultimate outcome for the stressed RGCs is apoptosis[10]. Capturing early functional changes of the RGCs could thus aid in predicting glaucoma progression and provide complementary information in the early diagnosis of this condition.

The electroretinogram (ERG)[11] is an electrical response measured at the cornea from a photic stimulus of the eye, and it is a direct measure of retinal function that could be used to capture early RGC dysfunction. Techniques described to measure such function include the pattern ERG (PERG), the positive scotopic threshold response (pSTR) and the Photopic Negative Response (PhNR)[12]. By following a cohort of people considered glaucoma suspects, Banitt[13] showed that changes in the PERG preceded structural changes as measured on the optical coherence tomography (OCT) by a number of years. Working in a rat model, Liu[14] showed that with chronic IOP elevation, pSTR dropped by 25% with no significant change in RGC density.

[1]Ophthalmology, Department of Surgery, University of Melbourne, Melbourne, VIC, Australia. [2]Duke-NUS Medical School, Singapore, Singapore. [3]RMIT University, Melbourne, Australia. [4]Department of Optometry and Vision Sciences, The University of Melbourne, Melbourne, Australia. [5]Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia. ✉email: marc@sarossy.com

The photopic negative response (PhNR) of the ERG is a slow negative potential following the b-wave. First described by Viswanathan et al.[15] in 1999, it has been shown to be reduced in glaucoma, both in clinical studies and in animal models[16–20] and Ref.[21]. It arises from the spiking potentials of retinal ganglion cells. It is elicited under photopic conditions by a red flash of light on a blue background and recorded from a corneal electrode relative to the lateral canthus of the eye. Previous studies have shown that there are significant correlations between the PhNR amplitude and the mean deviation (MD) of standard automated perimetry (SAP)[22] and peripapillary nerve fiber layer thickness[23]. The PhNR may thus be a useful measure for the objective assessment of RGC function in glaucoma.

Of note, the slow negative potential after the b-wave (the PhNR) is interrupted by a positive peak termed the i-wave of the ERG[24]. This waveform has a variable size and is poorly understood, but it is conserved in primate and non-primate species and thought to arise from cells distal to the RGCs and from the off pathway[25]. The presence of the i-wave can lead to different interpretations of the PhNR wave and therefore can affect the measurement of PhNR amplitude and latency. Recently, the PhNR has been subdivided into: (i) PhNR1, the trough after the b-wave and before the i-wave, and (ii) PhNR2, the first trough after the i-wave[26] (with the latter being typically measured as the PhNR). There has been little published on the significance of the i-wave itself, although authors have noted that it increases in prominence or amplitude in the presence of RGC pathology[25]. Capturing these additional parameters of the ERG beyond the conventional PhNR measure could also help better capture RGC dysfunction.

In addition, the stimulus parameters for optimizing the PhNR for the characterization of RGC dysfunction have been studied. Sustar[27] showed that the PhNR elicited with a chromatic stimulus was better at discriminating between glaucomatous and non-glaucomatous eyes than an achromatic stimulus. Hara[28] reported that the correlations between the PhNR1 and the mean deviation were stronger for the achromatic stimulus but the converse was true for the PhNR2, where the correlations were stronger for the achromatic stimulus. It is likely that these observations relate to the changes in the i-wave between the ERGs elicited by chromatic and achromatic stimuli, which they reported as occurring earlier later in the achromatic compared to chromatic stimulus, and thus had less of an impact on the PhNR I amplitude. Nonetheless, they reported that the PhNR II of the chromatic stimulus performed the best at discriminating between eyes with and without glaucoma. However, neither of these two studies have compared the utility of these three parameters—the PhNR1, i-wave and PhNR2—when used in combination (rather than in isolation) between the chromatic and achromatic stimuli. It is plausible that the utility of the achromatic stimulus may improve when these three parameters are considered in combination, due to its ability to more clearly demarcate the PhNR1 and i-wave.

We thus sought to examine whether the use of a fuller set of parameters (a-wave, b-wave, i-wave, PhNR1 and PhNR2) elicited using both chromatic and achromatic stimuli could better capture RGC functional loss compared with the use of the PhNR alone. This was performed by evaluating its predictive performance for a combined structural and functional measure of disease severity in glaucoma by Meideiros et al.[29], termed the estimated RGC (eRGC) count. Incorporating all markers into a predictive model allows for normalization of overall gain via the a- and b-wave amplitudes and the effect of the i-wave on the estimation of the PhNR via the use of the i-wave itself and the PhNR1 and PhNR2. In this study, we examined both a conventional linear predictive model and a Multivariate Adaptive Regression Splines (MARS) model. The MARS approach models the outcome variable as a linear sum of piecewise linear functions (truncated functions with knots)[30]. As such, it can be used to model the outcome measure as a linear sum of functions that can saturate or change their significance depending on amplitude, and allow intuition about the underlying process without incorporating interaction terms. We thus sought to use MARS as it could better model potentially complex functions for each of the ERG parameters to achieve improved prediction performance and model intuition.

Methods

The study was an approved study by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital, and it was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent prior to any study procedures being undertaken.

Participants. Participants with primary open angle glaucoma were recruited from a private ophthalmology practice, and the diagnosis of glaucoma was based on a comprehensive clinical assessment by an ophthalmologist based on characteristic optic nerve head appearance, the presence of glaucomatous visual field defects and/or neuroretinal tissue loss on OCT imaging. No participants with ocular or systemic diseases that could affect the optic nerve such (as choroidal neovascular membrane, extensive macular atrophy, diabetic retinopathy, multiple sclerosis or epiretinal membrane) were included and only adult patients over the age of 18 were eligible for inclusion in this study. Both eyes were included where glaucoma was bilateral, and only eyes with an acuity of 20/40 (LogMAR 0.3) or better were included in the study.

Standard automated perimetry. All participants performed standard automated perimetry (SAP) testing using the 24-2 Swedish Interactive Threshold Algorithm (SITA) Fast protocol on the Humphrey Field Analzyer 3 (Carl Zeiss Meditec Inc.; Dublin, CA), following correction of the spherical refractive error component from subjective refraction. All visual field results had fixation losses and false negative responses of $\leq 33\%$ and false positives of $\leq 20\%$.

Optical coherence tomography imaging. All participants also underwent optic disc-centered OCT volume scans performed with the Cirrus HD-OCT device (Carl Zeiss Meditec Inc.; Dublin, CA) with dilated pupils. This model has superluminescent diode illumination and obtains 27,000 A-scans per second, with an overall axial resolution of 5 µm. A scab centred on the optic disc was obtained consisting of 200 × 200 A-scans and covering an area of 6 × 6 mm. The circumpapillary retinal nerve fiber layer (RNFL) thickness was derived from this cube from a derived 3.46 mm diameter circle scan, consisting of 256 A scans with segmentation performed by the instrument. All scans had a signal strength score ≥ 7, and only those showing correct centration and accurate segmentation, were analysed in this study.

Estimated retinal ganglion cell counts. For each eye, an estimate was made of the RGC counts (eRGC) as described by Medeiros et al.[29]. This model is based on empirical formulas developed by Harwerth et al.[31]. From non-human primates, Harwerth developed a linear model estimating the RCG soma density (somas/mm^2) loss in dB units at location k (gl_k) at an eccentricity (ec) from the SAP visual sensitivity (s; dB) via slope (m) and intercept (b).

$$m_k = 0.054ec_k + 0.91, \tag{1}$$

$$b_k = -1.5ec_k - 14.8, \tag{2}$$

$$gl_k = (s_k - b_k)/m_k. \tag{3}$$

Translating the non-human primate work to humans required accounting for the different length of the eye in humans vs. non-human primates and the change in visual field threshold strategy from full threshold to SITA. For humans using SITA, the equations became

$$\mathrm{m_k} = 0.07128ec_k + 0.91, \tag{4}$$

$$b_k = -1.98ec_k - 14.8, \tag{5}$$

$$gl_k = \frac{s_k - 1 - b_k}{m_k} + 4.7. \tag{6}$$

The total ganglion cell number SAPrgc is found as

$$\mathrm{SAPrgc} = \sum\nolimits_{k=1}^{K} 10^{0.1gl_k}. \tag{7}$$

Noting that the number of ganglion cell axons entering the optic nerve from an area of the retina must be representative of the number of cell bodies, the authors also developed an estimate of axon numbers from the circumpapillary measurements of RNFL thickness measured by OCT. Recognizing a near linear loss of axons with normal aging and remodelling of the nerve fiber layer, the authors fitted the following formula relating axon density d (axons/µm^2) for a patient of age (ag) in years for a section of the RNFL scan of length px in pixels and average height (mh; µm). A 21.2 pixel length per pixel over a 10.87 mm scan length was used.

$$d = -0.007ag + 1.4. \tag{8}$$

And a total axon count a given simply as

$$a = 21.2mh\,px\,d. \tag{9}$$

Again, to account for the effect of remodelling in glaucoma and to express the final estimate in terms of total ganglion cell axons from OCT imaging (OCTrgc), a mean deviation correction c is calculated from the MD in dB as

$$\mathrm{c} = -0.26\,\mathrm{MD} + 0.12. \tag{10}$$

And using the d from Eq. (8), the OCT model is

$$\mathrm{OCTrgc} = 10^{0.1\left[10log_{10}(10870drnfl) - c\right]}, \tag{11}$$

where $rnfl$ is the mean RNFL thickness in µm from the OCT.

Medeiros et al.[29] finally proposed the use of a weighted mean reflecting the inverse relationship between disease severity of SAP and OCT estimates where the MD ranged from 0 to − 30 dB to obtain a final estimated RGC count in an eye eRGC as

$$\mathrm{eRGC} = \left(1 + \frac{MD}{30}\right)\mathrm{OC}T_{rgc} - \frac{MD}{30}\mathrm{SA}P_{rgc}. \tag{12}$$

Equation (12) was used for the RGC count estimate in this study.

Electroretinography. Both eyes were dilated with 1% tropicamide drops prior to recording. Three strand Dawson, Trick, and Litzkow (DTL)[32] wire electrodes were placed in the lower conjunctival fornices. Local anaesthetic drops were placed prior (1% tetracaine). Gold cup skin electrodes (Grass) were used for ground and indifferent electrodes. The indifferent electrodes were placed at the lateral canthi and the ground electrode was placed

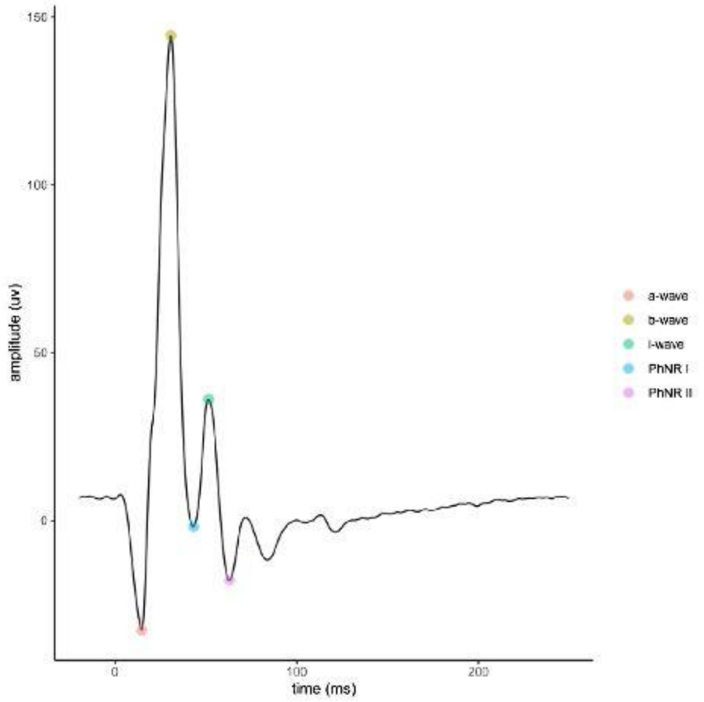


Figure 1. Typical electroretinogram trace (ERG) for the first step of red flash on blue background. Automatic marker placement is shown.

on the forehead. Impedances of all electrodes including the ground were checked using the inbuilt meter and were accepted if the impedance was less than 5 kΩ.

An Espion E3 system with the ColorDome ganzfeld stimulator (Diagnosys LLC; Lowell, MA) was used to collect the ERG data. For the first step, a blue background of 10 cd/m^2 (peak wavelength 465 nm) was used for a preadaptation time of 2 min. A total of 125 flashes were then presented at a rate of 2 Hz, with a luminance of 1 cd s/m^2 using a red flash (peak wavelength 635 nm) of 4 ms duration. This step was compliant with the International Society for the Clinical Electrophysiology (ISCEV) extended protocol for the PhNR[33]. Step 2 used the same photopic intensity of flash and background, but it was delivered as a white-on-white stimulus after a further 2 min preadaptation to the achromatic background. The stimuli were also presented at a rate of 2 Hz, with a 4 ms duration. The signals were collected with a sample rate of 4 kHz (which is higher than the minimum 1 kHz sample rate specified by the ISCEV protocol for recording the ERG) and an epoch length of 150 ms, and with 30 ms of pre-stimulus recording. Both raw data and filtered (band pass 0.3–100 Hz) were recorded and signals with an amplitude of over 150 μV were rejected.

Signal processing. The location of the a-wave and b-wave from the signal were determined by locating the b-wave as the global maximum and then the a-wave as the minimum in the segment between the start of the vector and the b-wave. The i-wave was located by determining the first peak after the b-wave. The PhNR1 and PhNR2 were determined to be the negative going troughs either side of the i-wave. A typical ERG averaged trace elicited from chromatic stimuli is shown in Fig. 1.

Statistical analysis. The Caret[34] package in R was used for predictive model fitting and tuning. Linear and multivariate adaptive regression splines (MARS) were fitted. Overfitting was avoided by using tenfold cross-validation. For the MARS models, the degree was set to 1 and the maximum number of terms was set to 40. The degree is the level of interaction of terms allowed—in this case none allowed. An example second degree interaction term would be the a-wave amplitude multiplied by the b-wave amplitude. The number of terms in the model is a function of the number of input features and the number of breaks in the hinge function. Without a breakpoint, a single feature will have 2 terms: intercept and gradient. With a single breakpoint, this becomes 4 terms: intercept, the 2 gradients and the breakpoint location.

Characteristics	(103 eyes from 55 individuals)
Individual level	
Age (years)	75 (66 to 80)
Gender (female)	21 (38%)
Diabetes (present)	7 (13%)
Hypertension (present)	35 (64%)
Eye level	
Refraction sphere (D)	0.00 (− 1.00 to 0.50)
Visual acuity (logMAR)	0.0 (− 0.1 to 0.1)
Intraocular pressure (mmHg)	15.0 (12.0 to 16.0)
Mean deviation (dB)	− 2.5 (− 5.9 to − 0.5)
Retinal nerve fiber layer thickness (μm)	73.0 (62.8 to 82.5)
eRGC ('000s)	601 (470 to 753)

Table 1. Characteristics of the individuals and eyes with glaucoma in this study. Continuous statistics presented as median and interquartile range, categorical statistics as number and percentage. *logMAR* logarithm of the minimum angle of resolution, *eRGC* estimated retinal ganglion cell count.

	R^2		*P*-value for comparison between				
			PhNR vs. markers		Linear vs. MARS	Chromatic vs. achromatic	
	Linear	MARS	Linear	MARS		Linear	MARS
Chromatic							
PhNR	0.09	0.11	0.02	0.02	0.14	0.30	0.40
Markers	0.22	0.33			0.01	0.19	0.20
Achromatic							
PhNR	0.16	0.18	0.02	0.01	0.15	–	–
Markers	0.31	0.45			0.01	–	–

Table 2. Model performance as proportion of variance explained (R^2) for the prediction of ganglion cell count from models based on photopic negative response (PhNR) alone or the full set of amplitude features ("Markers", including the a-, b- and i-waves and PhNR1 and PhNR2). Both simple linear regression and multivariate adaptive regression spline (MARS) models are shown. Testing was with red flashes on blue background (chromatic) or white on white (achromatic). P-values were calculated by bootstrap resampling.

For each of the tuned 'final models', the prediction performance of the ERG for the eRGC counts was examined based on the proportion of variance explained (R^2) by the model. Each model was compared to the base case of the PhNR only model. The significance of difference was determined using a bootstrap technique ($n = 1000$ resamples) which dealt with the within-subject correlations through resampling at the individual level.

Results

A total of 103 eyes with glaucoma from 55 participants were included in this study, and their characteristics are shown in Table 1.

Prediction of estimated retinal ganglion cell counts.

The PhNR/B ratio gave poor predictive performance with R^2 values ranging from 0.028 for the linear chromatic model through to 0.073 for the MARS chromatic model. Table 2 summarizes the predictive ability of the different models using the ERG for the eRGC measure for the PhNR alone and the full set of markers. The predictions of eRGC were significantly better with the models utilising the full feature set from the ERG, compared with those using the PhNR alone ($P \geq 0.02$ for all). The predictive performance of the MARS models was also significantly better than the linear models using the full set of markers ($P = 0.01$ for both), but not when using the PhNR alone ($P \geq 0.14$ for both). Finally, the models derived from ERG recordings with achromatic stimuli performed better at predicting eRGC counts than ones using chromatic stimuli, but it this difference did not reach statistical significance ($P \geq 0.19$ for all).

For the MARS models using both the chromatic and achromatic stimulus, automated tuning removed the b-wave amplitude and it was not used as a feature, leaving 4 terms.

The achromatic MARS marker model is illustrated in Fig. 2 with an example prediction. This model has an intercept of 940,162 RGCs and the additional effect of each term is read from the y axis. In this example, the PhNR2 and i-wave have no effect. The a-wave and PhNR2 terms contribute to give an overall estimate of 287,902 RGCs, being similar to the estimate of 322,678 RGCs for this example determined from the visual field and OCT data.

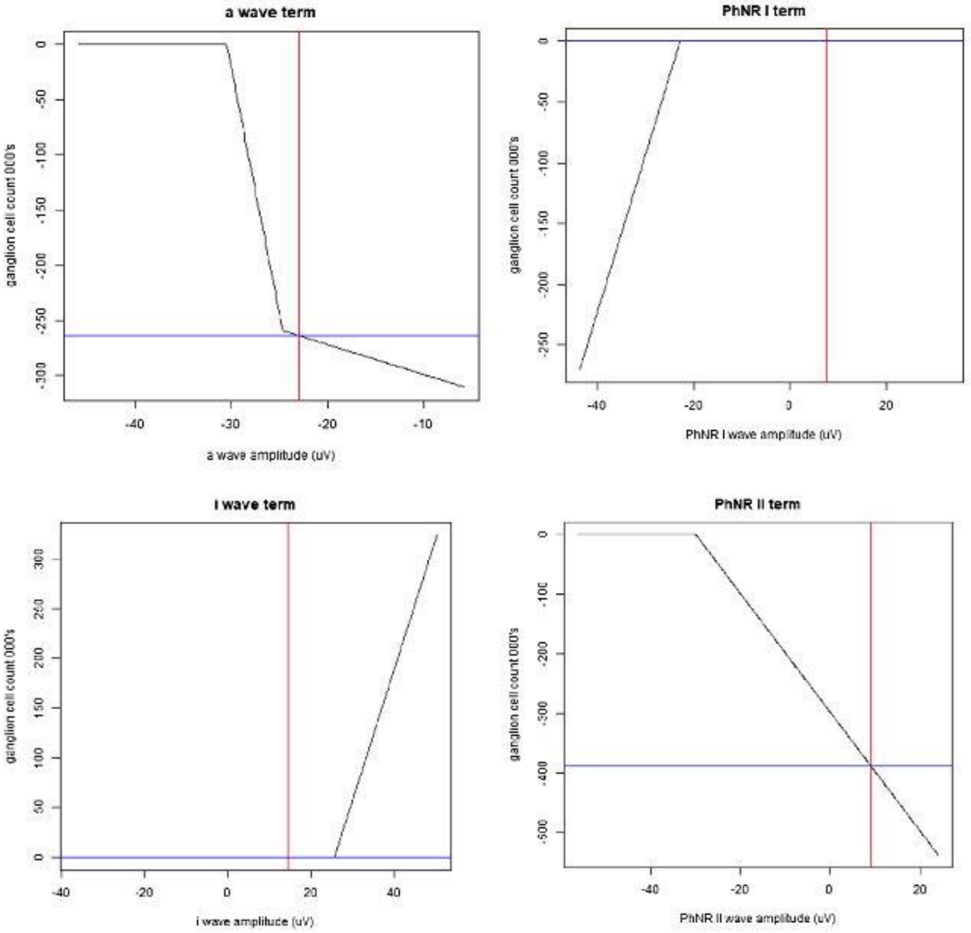


Figure 2. MARS model—achromatic stimulus. The panels represent the hinge functions and the influence of each term on the predicted estimate of retinal ganglion cells (eRGC). The red line shows amplitude of the various waves from an example eye. The effect on the eGRC for each term is shown on the vertical axis in each panel. The final eRGC is the intercept (940,162 retinal ganglion cells) plus the effect of each term.

Discussion

This study showed that using additional information from the ERG—the a-, b-, and i-wave, as well as the troughs on either side of the i-wave (PhNR1 and PhNR2)—improved the prediction of the eRGC measure compared to using the PhNR alone. Further improvements were achieved by replacing a linear model with the quasi-linear MARS model, whilst the use of achromatic stimuli resulted in non-significant improvements compared to the use of chromatic stimuli. These findings demonstrate how there is a wealth of information present in the ERG for characterizing RGC function beyond the PhNR measure alone.

The PhNR amplitude has previously been shown to correlate with glaucoma severity[17,21,23], but this is the first time to our knowledge that multiple parameters—namely, the a-, b-, and i-waves, and the PhNR1 and PhNR2 measures—from the ERG have been used in combination to predict glaucoma severity. This study showed that the additional parameters led to significantly better performance than using the conventional PhNR measure alone, which may be attributed to the ability of such a model to allow for overall eye gain and to also to allow for the effect of the i-wave contamination of the measurement of the process driving the PhNR itself. It also allows for the inclusion of the process driving the i-wave which might itself be affected in glaucoma. We also found that the MARS models performed better than the linear models, while still allowing for a high level of intuition. As can be seen from Fig. 2 for the responses from an achromatic stimulus, the a-wave amplitude has an impact at smaller amplitudes, perhaps correcting for overall eye gain. The influence of the PhNR2 term saturates at larger (negative) amplitudes whereas the PhNR1 term only starts having an effect at large (negative) amplitudes. The i-wave only had an influence on the model when its amplitude was high. Piecewise linear functions like these are discontinuous in the first derivative and although this is unlikely to be the optimum representation of the contribution of the input features, it can be an improvement on a simple linear model in that it allows saturation. These findings highlight how non-linear models could make better use of the multiple parameters from the ERG in characterizing RGC function than described by linear models.

In this study, we did not find any significant difference between chromatic and achromatic stimuli in their prediction performance of glaucoma severity. We tested the hypothesis that the response with a white-on-white stimulus although smaller, might have better discriminatory ability for the assessment of glaucoma severity—especially with the ability to combine the other parameters in a linear and non-linear way. Although the achromatic stimulus was not worse, it did not provide any significant advantage.

Whilst this study sought to examine whether alternative approaches with the ERG could improve the prediction of glaucoma disease severity (represented by the extent of RGC loss), we did not expect a perfect relationship since the ERG may capture RGC dysfunction that precedes its loss. However, this study provides us with an opportunity to examine whether we can better fully utilize the information available from the ERG response (using more parameters and more complex modelling) and whether different methods for eliciting the response (achromatic stimuli) would be useful for capturing RGC function. Indeed, we observed that there was benefit in considering more information derived from the ERG than the PhNR alone, and that maximum benefit from the additional parameters requires more complex modelling using the MARS approach. These findings highlight how future studies seeking to explore the utility of the ERG for tackling diagnostic dilemmas in the clinical management of glaucoma, such as of whether an eye needs treatment, what the target intraocular pressure should be, and the likelihood of disease progression, would benefit from considering additional parameters present in the ERG that are exploited through the MARS approach.

A potential limitation of this study was the use of the eRGC parameter, which previous studies have cautioned may not provide estimates of the actual number of RGCs present[35,36]. Nonetheless, these studies acknowledged that this issue of estimating the number of RGCs present differed from the clinical value of combining measures of structure and function in characterizing glaucoma severity. Indeed, we used the eRGC counts parameter in this study as a means of characterizing glaucoma severity, which we consider a strength of this study as previous studies have shown that this parameter outperformed clinical measures of structure and function used in isolation for this purpose[37–40]. We showed in this study that extracting additional information from the ERG could allow better understanding of RGC function. This study was limited to simple amplitude measures, and it is possible that better performance could be obtained by adding features such as latency, frequency domain characteristics, such as median frequency or even features from time–frequency analysis such as wavelet coefficients.

Finally, another potential limitation of this study was its sample size and the inclusion of both eyes from each participant. The Intra-cluster correlation coefficient (ICC) calculated by the method of Lohr[41] using the psych package in R[42] yielded an ICC for the eRGC using the side as the clustering parameter of only 0.011 indicating near independence. Our study had 79% power to detect a difference of 0.10 in the R^2 if only one eye per participant was included, based on a one-sided alpha of 0.05. Based on the effective sample size after accounting for the between-eye correlations for participants where both eyes were included, this study would have 96% power to detect this same effect size, when using the appropriate statistical methodology to account for between eye correlations (performed using the bootstrap method in this study, resampling at the individual level). The study was thus adequately powered to detect the differences in predictive performance across the different ERG models, as evident from the significant improvements detected when using the MARS model and full set of ERG features.

In conclusion, this study shows that there is significant benefit in considering more information derived from the ERG than the PhNR alone. Maximum benefit from the additional parameters requires more complex modelling, and we showed a high predictive performance for glaucoma severity using the MARS approach with the full set of ERG amplitude features. These findings highlight key analytical approaches to incorporate for future studies evaluating the utility of the photopic ERG in the clinical management of glaucoma.

Received: 4 September 2021; Accepted: 29 November 2021

References

1. Weinreb, R. N., Aung, T. & Medeiros, F. A. The pathophysiology and treatment of glaucoma: A review. *JAMA* **311**, 1901–1911 (2014).
2. Kerrigan-Baumrind, L. A., Quigley, H. A., Pease, M. E., Kerrigan, D. F. & Mitchell, R. S. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Investig. Ophthalmol. Vis. Sci.* **41**, 741–748 (2000).
3. Fry, L. E. *et al.* The coma in glaucoma: Retinal ganglion cell dysfunction and recovery. *Prog. Retin. Eye Res.* **65**, 77–92 (2018).
4. Kalesnykas, G. *et al.* Retinal ganglion cell morphology after optic nerve crush and experimental glaucoma. *Investig. Ophthalmol. Vis. Sci.* **53**, 3847–3857 (2012).
5. Della Santina, L., Inman, D. M., Lupien, C. B., Horner, P. J. & Wong, R. O. Differential progression of structural and functional alterations in distinct retinal ganglion cell types in a mouse model of glaucoma. *J. Neurosci.* **33**, 17444–17457 (2013).
6. Tao, X., Sabharwal, J., Wu, S. M. & Frankfort, B. J. Intraocular pressure elevation compromises retinal ganglion cell light adaptation. *Investig. Ophthalmol. Vis. Sci.* **61**, 15–15 (2020).
7. Risner, M. L., Pasini, S., Cooper, M. L., Lambert, W. S. & Calkins, D. J. Axogenic mechanism enhances retinal ganglion cell excitability during early progression in glaucoma. *Proc. Natl. Acad. Sci.* **115**, E2393–E2402 (2018).
8. Calkins, D. J. Adaptive responses to neurodegenerative stress in glaucoma. *Prog. Retin. Eye Res.* **84**, 100953 (2021).
9. Weber, A. J. & Harman, C. D. Structure–function relations of parasol cells in the normal and glaucomatous primate retina. *Investig. Ophthalmol. Vis. Sci.* **46**, 3197–3207 (2005).

10. Nickells, R. W. Apoptosis of retinal ganglion cells in glaucoma: An update of the molecular pathways involved in cell death. *Surv. Ophthalmol.* **43**, S151–S161 (1999).
11. McCulloch, D. L. *et al.* ISCEV Standard for full-field clinical electroretinography (2015 update). *Doc. Ophthalmol.* **130**, 1–12 (2015).
12. Porciatti, V. Electrophysiological assessment of retinal ganglion cell function. *Exp. Eye Res.* **141**, 164–170 (2015).
13. Banitt, M. R. *et al.* Progressive loss of retinal ganglion cell function precedes structural loss by several years in glaucoma suspects. *Investig. Ophthalmol. Vis. Sci.* **54**, 2346–2352 (2013).
14. Liu, H. H., He, Z., Nguyen, C. T., Vingrys, A. J. & Bui, B. V. Reversal of functional loss in a rat model of chronic intraocular pressure elevation. *Ophthalmic Physiol. Opt.* **37**, 71–81 (2017).
15. Viswanathan, S., Frishman, L. J., Robson, J. G., Harwerth, R. S. & Smith, E. R. The photopic negative response of the macaque electroretinogram: Reduction by experimental glaucoma. *Investig. Ophthalmol. Vis. Sci.* **40**, 1124–1136 (1999).
16. Viswanathan, S., Frishman, L. J., Robson, J. G. & Walters, J. W. The photopic negative response of the flash electroretinogram in primary open angle glaucoma. *Investig. Ophthalmol. Vis. Sci.* **42**, 514–522 (2001).
17. Machida, S. *et al.* Correlation between photopic negative response and retinal nerve fiber layer thickness and optic disc topography in glaucomatous eyes. *Investig. Ophthalmol. Vis. Sci.* **49**, 2201–2207 (2008).
18. North, R. V., Jones, A. L., Drasdo, N., Wild, J. M. & Morgan, J. E. Electrophysiological evidence of early functional damage in glaucoma and ocular hypertension. *Investig. Ophthalmol. Vis. Sci.* **51**, 1216–1222 (2010).
19. Nakamura, H., Hangai, M., Mori, S., Hirose, F. & Yoshimura, N. Hemispherical focal macular photopic negative response and macular inner retinal thickness in open-angle glaucoma. *Am. J. Ophthalmol.* **151**, 494-506.e491 (2011).
20. Machida, S., Kaneko, M. & Kurosaka, D. Regional variations in correlation between photopic negative response of focal electoretinograms and ganglion cell complex in glaucoma. *Curr. Eye Res.* **40**, 439–449 (2015).
21. Preiser, D., Lagreze, W. A., Bach, M. & Poloschek, C. M. Photopic negative response versus pattern electroretinogram in early glaucoma. *Investig. Ophthalmol. Vis. Sci.* **54**, 1182–1191 (2013).
22. Kita, Y. *et al.* RETeval portable electroretinogram parameters in different severity stages of glaucoma. *J. Glaucoma* **29**, 572–580 (2020).
23. Cvenkel, B., Sustar, M. & Perovšek, D. Ganglion cell loss in early glaucoma, as assessed by photopic negative response, pattern electroretinogram, and spectral-domain optical coherence tomography. *Doc. Ophthalmol.* **135**, 17–28 (2017).
24. Rosolen, S. G. *et al.* Comparing the photopic ERG i-wave in different species. *Vet. Ophthalmol.* **7**, 189–192 (2004).
25. Rangaswamy, N. V. *et al.* Photopic ERGs in patients with optic neuropathies: comparison with primate ERGs after pharmacologic blockade of inner retina. *Investig. Ophthalmol. Vis. Sci.* **45**, 3827–3837 (2004).
26. Ortiz, G. *et al.* The photopic negative response of the Light-adapted 3.0 ERG in clinical settings. *Doc. Ophthalmol.* **140**, 115–128 (2020).
27. Sustar, M., Cvenkel, B. & Brecelj, J. The effect of broadband and monochromatic stimuli on the photopic negative response of the electroretinogram in normal subjects and in open-angle glaucoma patients. *Doc. Ophthalmol.* **118**, 167 (2009).
28. Hara, Y. *et al.* Comparisons of photopic negative responses elicited by different conditions from glaucomatous eyes. *Jpn. J. Ophthalmol.* **64**, 114–126 (2020).
29. Medeiros, F. A. *et al.* A combined index of structure and function for staging glaucomatous damage. *Arch. Ophthalmol.* **130**, 1107–1116 (2012).
30. Friedman, J. H. Multivariate adaptive regression splines. *Ann. Stat.* **19**(1), 1–67 (1991).
31. Harwerth, R., Wheat, J., Fredette, M. & Anderson, D. Linking structure and function in glaucoma. *Prog. Retin. Eye Res.* **29**, 249–271 (2010).
32. Dawson, W. W., Trick, G. L., Litzkow, C. A. J. I. O. & Science, V. Improved electrode for electroretinography. *Investig. Ophthalmol. Vis. Sci.* **18**, 988–991 (1979).
33. Frishman, L. *et al.* ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram. *Doc. Ophthalmol.* **136**(3), 1–5 (2018).
34. Kuhn, M. The caret package. J. Stat. Softw. **28**(5) (2009).
35. Raza, A. S. & Hood, D. C. Evaluation of a method for estimating retinal ganglion cell counts using visual fields and optical coherence tomography. *Investig. Ophthalmol. Vis. Sci.* **56**, 2254–2268 (2015).
36. Swanson, W. H. & Horner, D. G. Assessing assumptions of a combined structure-function index. *Ophthalmic Physiol. Opt.* **35**, 186–193 (2015).
37. Zhang, C. *et al.* Relationship between ganglion cell layer thickness and estimated retinal ganglion cell counts in the glaucomatous macula. *Ophthalmology* **121**, 2371–2379 (2014).
38. Medeiros, F. A. *et al.* Retinal ganglion cell count estimates associated with early development of visual field defects in glaucoma. *Ophthalmology* **120**, 736–744 (2013).
39. Rolle, T., Bonetti, B., Mazzucco, A. & Dallorto, L. Diagnostic ability of OCT parameters and retinal ganglion cells count in identification of glaucoma in myopic preperimetric eyes. *BMC Ophthalmol.* **20**, 1–10 (2020).
40. Rolle, T., Dallorto, L. & Bonetti, B. Retinal and macular ganglion cell count estimated with optical coherence tomography RTVUE-100 as a candidate biomarker for glaucoma. *Investig. Ophthalmol. Vis. Sci.* **57**, 5772–5779 (2016).
41. Lohr, S. L. *Sampling: Design and Analysis*. 177 (CRC Press, 2019).
42. Revelle, W., & Revelle, M. W. Package 'psych'. The comprehensive R archive network 337, 338 (2018).

Author contributions

All authors conceived the study. M.S. recruited the participants and collected the data. M.S. and Z.W. performed the analysis. M.S. wrote the main manuscript text. All authors reviewed the manuscript. J.C., D.K., A.W. and Z.W. provided supervision.

Funding

Supported in part by a Hugh Noel Puckle Scholarship (MS) and National Health & Medical Research Council Early Career Fellowship (#1104985, ZW).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.S.

Reprints and permissions information is available at www.nature.com/reprints.

tvst

Glaucoma

Time–Frequency Analysis of ERG With Discrete Wavelet Transform and Matching Pursuits for Glaucoma

Marc Sarossy[1], Jonathan Crowston[2], Dinesh Kumar[3], Anne Weymouth[4], and Zhichao Wu[1,5]

[1] Ophthalmology, Department of Surgery, University of Melbourne, Melbourne, Victoria, Australia
[2] Duke-NUS Medical School, Singapore
[3] RMIT University, Melbourne, Victoria, Australia
[4] Department of Optometry and Vision Sciences, University of Melbourne, Melbourne, Victoria, Australia
[5] Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Victoria, Australia

Correspondence: Marc Sarossy, Ophthalmology, Department of Surgery, University of Melbourne, 7/32 Gisborne St East Melbourne, Victoria 3002, Australia. e-mail: marc@sarossy.com

Received: March 19, 2022
Accepted: September 13, 2022
Published: October 13, 2022

Keywords: electroretinogram; glaucoma; photopic negative response; predictive modeling

Citation: Sarossy M, Crowston J, Kumar D, Weymouth A, Wu Z. Time–frequency analysis of ERG with discrete wavelet transform and matching pursuits for glaucoma. Transl Vis Sci Technol. 2022;11(10):19, https://doi.org/10.1167/tvst.11.10.19

Purpose: To examine the performance of two time–frequency feature extraction techniques applied to electroretinograms (ERGs) for the prediction of glaucoma severity.

Methods: ERGs targeting the photopic negative response were obtained in 103 eyes of 55 patients with glaucoma. Features from the ERG recordings were extracted using two time–frequency extraction techniques based on the discrete wavelet transform (DWT) and the matching pursuit (MP) decomposition. Amplitude markers of the time-domain signal were also extracted. Linear and multivariate adaptive regression spline (MARS) models were fitted using combinations of these features to predict estimated retinal ganglion cell counts, a measure of glaucoma disease severity derived from standard automated perimetry and optical coherence tomography imaging.

Results: Predictive models using features from the time–frequency analyses—using both DWT and MP—combined with amplitude markers outperformed predictive models using the markers alone with linear ($P = 0.001$) and MARS ($P \leq 0.011$) models. For example, the proportions of variance (R^2) explained by the MARS model using the DWT and MP features with amplitude markers were 0.53 and 0.63, respectively, compared to 0.34 for the model using the markers alone ($P = 0.011$ and $P = 0.001$, respectively).

Conclusions: Novel time–frequency features extracted from the photopic ERG substantially added to the prediction of glaucoma severity compared to using the time-domain amplitude markers alone.

Translational Relevance: Substantial information about retinal ganglion cell dysfunction exists in the time–frequency domain of ERGs that could be useful in the management of glaucoma.

Introduction

Glaucoma is an optic neuropathy characterized by the progressive loss of retinal ganglion cells (RGCs),[1] which exhibit structural and functional changes prior to their death by apoptosis.[2] Early RGC structural changes include reduction in the length and number of dendrites and the area of the dendritic arbor, as seen from mouse models of optic nerve injury.[3,4] Functional changes in the RGCs prior to apoptosis can include an increased firing rate from increased excitability[5,6] and a fall in the mean and peak spike rates.[7] Detecting these early functional changes might potentially aid the prediction of future RGC loss in glaucoma and thus enhance the clinical management of progression of this condition, which was estimated to affect 64.3 million people in the world between the ages of 30 and 80 years in 2013.[8]

Electrophysiological recordings from the eye may be able to detect the earliest changes in RGC function. In a study on glaucoma suspects, Banitt et al.[9] showed that a 10% change in the pattern electroretinogram (PERG) amplitude preceded the same change in peripapillary retinal nerve fiber layer (RNFL) thickness by 8 years. Liu et al.[10] worked in a rat model and showed that, with chronic intraocular pressure (IOP) elevation, the positive scotopic threshold response was reduced by 25% in animals with elevated pressure without any changes in the optical coherence tomography (OCT) parameters and that these changes reversed with normalization of the IOP. It is therefore plausible that the electroretinogram (ERG) is a tool that could be used to capture such early functional abnormalities of the RGCs. The ERG is a time-domain signal of the electrical activity of the retina in response to light stimuli. When recorded under light-adapted conditions with a brief presentation of a red-on-blue stimulus, a slow negative going wave occurring after an initial a-wave trough and b-wave peak—termed the photopic negative response (PhNR)—is typically observable.[11] The PhNR is a response that arises from the RGCs and has been demonstrated to be reduced in glaucoma, both in experimental models and in clinical studies.[12–17] It can show improvement in patients with glaucoma when the IOP is lowered[18] and has been used as a marker of inner retinal function in patients with glaucoma when treated with nicotinamide.[19] The signal between the b-wave peak and the PhNR trough is frequently interrupted by the i-wave, which is thought to arise from the cells distal to the RGCs.[20] If the i-wave is present, the PhNR can be recorded as either the negative wave between the b-wave and i-wave (the PhNR1) or the negative wave following the i-wave (the PhNR2). We have previously shown that a combination of key amplitude markers of the ERG—namely, the a-wave, b-wave, i-wave, and PhNR (PhNR1 and PhNR2) amplitudes—better predicts glaucoma severity than the PhNR amplitude alone.[21] These findings underscore how a photopic ERG contains more information about RGC function than is captured by conventional PhNR measures alone.

The ERG measured from a single differential pair electrode is a mixture of the underlying processes, which occur at different times after the stimulus.[22] The underlying mechanisms are complex[23] with feedback[24] and feedforward pathways,[25,26] and there are slow and fast frequency components from the different sources of the signal. Time–frequency analysis has the potential to extract these components from the photopic ERG to provide further insights into RGC function, above and beyond what is captured by the amplitude markers of the ERG. With time–frequency analysis, a vector of amplitude measurements is transformed into a matrix of coefficients with axes of time and frequency so that the magnitude and timing of frequency components within the signal can be determined. The wavelet transform, a form of time–frequency analysis, has been described as an alternative to short-time Fourier transform.[27,28] A wavelet is a small function (little wave) that acts as a filter and can localize energy within the signal in time and frequency. This is achieved through multiplication of the signal and wavelet after translation and dilation of the wavelet relative to the signal for time and frequency resolution, respectively. Wavelets have been applied to the PERG,[29] multifocal ERG,[30] and photopic ERG[31] in the assessment of glaucoma.

There are two types of wavelet transforms: continuous wavelet transform (CWT)[28] and discrete wavelet transform (DWT).[32]

The CWT is obtained by convoluting the wavelet with the signal for all values of scale (e.g., frequency) and time lag, creating a continuous scalogram showing the energy of the signal at each frequency and time point. However, this creates a highly redundant output matrix, as both its width and length are now equal to the number of samples in the original time series vector. Forte and colleagues[33] showed that the Morlet wavelet could be used to isolate oscillatory potentials in rat ERGs. Behbahani and colleagues[34] used a Mexican hat wavelet to determine the dominant frequencies associated with the PhNR in patients with central retinal vein occlusion and found that the dominant frequency decreased. However, the CWT has not thus far found clinical utility for glaucoma.

On the other hand, the DWT[32,35,36] uses set scales and time lags at discrete values, where the output is in the form of a binary tree and the total number of coefficients is equal to the number of samples in the input, which necessarily must be of length equal to a power of two. Specifically, it uses a low-pass filter (scaling function) and a high-pass filter (wavelet function), followed by downsampling by two. There are many mother wavelets available for the DWT. Selection of the mother wavelet optimizes the resolution in time and frequency for the temporal and spectral content of the signal. Various techniques for mother wavelet selection have been described,[37] including minimizing informational cost. The DWT has been applied to glaucoma, but previous studies have generally used a single feature derived from the transform.[30,38] A recent paper explored the use of the DWT in the photopic ERG and found differences between individuals diagnosed with autism spectrum disorder compared with control subjects.[39]

A different approach to time–frequency analysis is the matching pursuit (MP) algorithm.[40] The MP algorithm decomposes a time-domain signal into a linear combination of subsignals of the same length termed atoms. The full set of atoms is referred to as the dictionary. The decomposition output is a set consisting of a coefficient for each atom in the dictionary. The process of generating the dictionary usually begins with a discrete wavelet family: small vectors representing digital filters.[32] In contrast to the DWT, the atoms in the MP decomposition are padded to the length of the signal rather than a much shorter vector that slides along the signal. The dictionary thus includes all of the time shifts of a given wavelet as separate atoms. The algorithm process is known as "greedy." It starts by finding the best match to the signal from the entire dictionary, removes that part of the signal, and then finds the next best match and so on. The number of iterations is usually equal to the length of the dictionary, although a smaller fixed number can be used, or the algorithm can reach a stopping criterion. The result of the transformation is the coefficient for each atom and the index into the dictionary used. The dimensionality of the transformation may be larger or smaller than the original signal.

We therefore posed this optimization problem: Given a continuous outcome measure (the estimated RGC), can additional features extracted by the DWT or the MP informing linear or MARS models yield better performance than time-domain amplitude features alone? Although the DWT is more straightforward for others to replicate with widely available software, the MP technique may offer better time localization for low frequencies. Some studies have compared the two techniques in hyperspectral imaging[41] and electroencephalogram analysis,[42] although these studies addressed classification rather than regression problems and the differences were modest. Both techniques in theory allow the extraction of multiple features from the underlying processes and could plausibly better characterize the extent of retinal dysfunction than time-domain features alone.

Given the ability of the DWT and MP algorithms to extract novel time–frequency features from the photopic ERG, we examined whether these approaches, when used together with an extended set of amplitude markers, could be used to better predict glaucoma severity (or the extent of RGC loss and dysfunction). We compared the incremental benefit of each method and the combination of both. Our aim was not to develop a new clinical tool for the diagnosis of glaucoma or the classification of its severity, as clinicians currently have OCT, standard automated perimetry (SAP), and clinical examination for that. Rather, the aim of this study was to elucidate additional information within the ERG that can be extracted by time–frequency techniques and which might ultimately be useful in, for example, building a predictive model of progression.

Methods

This study was an approved study by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital, and it was conducted in accordance with the tenets of the Declaration of Helsinki. All participants provided written informed consent prior to any study procedures being undertaken.

Participants

Participants with primary open-angle glaucoma were recruited from a private ophthalmology practice, and the diagnosis of glaucoma was based on a comprehensive clinical assessment by an ophthalmologist based on characteristic optic nerve head appearance, the presence of glaucomatous visual field defects, and/or neuroretinal tissue loss on OCT imaging. Participants with ocular or systemic diseases that could affect the optic nerve (such as choroidal neovascular membrane, extensive macular atrophy, diabetic retinopathy, multiple sclerosis, or epiretinal membrane) were excluded. Only individuals over the age of 18 were eligible. Both eyes were included where glaucoma was bilateral, and only eyes with an acuity of 20/40 (or 0.30 logMAR) or better were included in this study.

Automated Perimetry

All participants performed SAP testing using the 24-2 Swedish interactive thresholding algorithm (SITA) Fast protocol on the Humphrey Field Analyzer 3 (Carl Zeiss Meditec, Jena, Germany), following correction of the spherical refractive error component from subjective refraction. Only visual field results where fixation losses or false-negative responses were $\leq 33\%$ and false positive responses were $\leq 20\%$ were considered reliable and analyzed in this study.

Optical Coherence Tomography

All participants underwent optic disc-centered OCT volume scans performed with the CIRRUS HD-OCT device (Carl Zeiss Meditec) with dilated pupils. Each scan consisted of 200 × 200 A-scans and covered an area of 6 × 6 mm. The global circumpapillary RNFL thickness scan was calculated from a derived 3.46-mm-diameter circle scan consisting of 256 A-scans. All scans were checked for centration and segmentation accuracy.

Estimated RGC Counts

For each eye, the measure of estimated RGC (eRGC) counts, as described by Medeiros and colleagues,[43] was derived to provide an index of disease severity. This measure was based on empirical formulas developed by Harwerth and colleagues[44] in non-human primates, which related RGC counts to visual field measurements. In brief, the eRGC is based on an estimate of the number of RGC somas in each retinal region sampled on SAP ($\mathrm{SAP_{rgc}}$) and also an estimate of the number of RGC axons at the circumpapillary circle on OCT imaging (with corrections for the effects of disease-related remodeling of the RNFL axonal and non-axonal composition, or $\mathrm{OCT_{rgc}}$). Medeiros and colleagues[43] proposed the use of a weighted mean of the structural and functional estimates to reflect the inverse relationship between the accuracies of clinical SAP and OCT imaging estimates with disease severity. Using the mean deviation (MD) values from SAP as a measure of disease severity, the eRGC count in an eye was thus calculated as

$$\mathrm{eRGC} = \left(1 + \frac{\mathrm{MD}}{30}\right)\mathrm{OCT_{rgc}} - \frac{\mathrm{MD}}{30}\mathrm{SAP_{rgc}} \quad (2)$$

A recent paper found the range of the eRGC in healthy eyes without glaucoma to be 1,102,108 ± 72,669.[45]

ERG Recordings

An Espion E3 system (Diagnosys, Lowell, MA) was used with a Diagnosys ColorDome light-emitting diode Ganzfeld stimulator to collect the ERG data. The recordings were in compliance with the International Society for Clinical Electrophysiology of Vision (ISCEV) 2018 extended protocol for the photopic negative response of the electroretinogram.[11] Participants were not dark adapted at any stage of the testing. They were adapted to ambient light in the testing suite for at least 10 minutes during pupil dilation. Both eyes were dilated with 1% tropicamide drops prior to recording, and active Dawson–Trick–Litzkow electrodes were used with gold-cup skin electrodes as the reference and ground electrodes. The ground electrode was placed at Fz and the reference electrode at the lateral canthus. Impedances of all electrodes, including the ground, were checked to ensure that they were less than 5 kΩ. The stimulus parameters were compliant with the protocol and used a blue background of 10 cd/m^2 (peak wavelength = 465 nm) with a preadaptation time within the Ganzfeld dome of 2 minutes and presenting red flashes (peak wavelength = 635 nm) with a 4-ms duration at 1 cd·s/m^2. A sample rate of 4000 Hz was used with 20 ms of pre-stimulus and 250 ms of post-stimulus recording, yielding a vector length of 1079 samples. Flashes were presented at 2 Hz. Voltage-based automatic rejection of traces with blink or movement artifacts was used[17,46] with signals with an absolute amplitude of 200 µV within 100% of the sweep range rejected. This value replicated the level used in our laboratory for routine photopic full-field electroretinogram recordings with this equipment and electrode type. A total of 125 unrejected sweeps per stimulus were collected. Bandpass filtering from 0.3 to 100 Hz was performed. This upper bound, which is lower than the high-frequency cutoff of the ISCEV protocol, was chosen as the pilot for this project used a achromatic stimulus and the standard photopic ERG protocol of the laboratory.

Preprocessing

After signal averaging of the unrejected traces, the ensemble averages were trimmed so that the b-wave peak occurred at the 200th sample and the epoch was 1024 samples long. The alignment process is illustrated in Figure 1. The alignment was performed because the region of interest within the time–frequency spectrum is much later than for analysis of outer retinal function via traditional electroretinogram markers. Aligning the averages in this way ensures that the largest coefficient of either the DWT or the MP corresponds in all cases to the b-wave.

Amplitude Markers

For each ensemble average of ERG traces, amplitude markers were determined as described previously.[21] Briefly, the a-wave trough was determined as the first negative trough after the stimulus onset, and the b-wave peak was found as the first peak after the a-wave. The i-wave was found as the first peak or inflexion point after the b-wave. The PhNR1 was the first trough between the b-wave and the i-wave, or equal to the i-wave in the case of that being an inflexion point. The PhNR2 was found as the first trough after the i-wave.

Discrete Wavelet Transform

For each trace, the DWT was performed using Wavelet Methods for Time-Series Analysis (WMTSA) version 2.0-3 in R (R Foundation for Statistical Computing, Vienna, Austria).[47] Figure 2 shows a typical ERG and the DWT performed in this case with a Daubechies wavelet with 8 vanishing moments (d8 wavelet), with the heatmap showing the absolute magnitude of the wavelet coefficients by decomposition level and position in time.

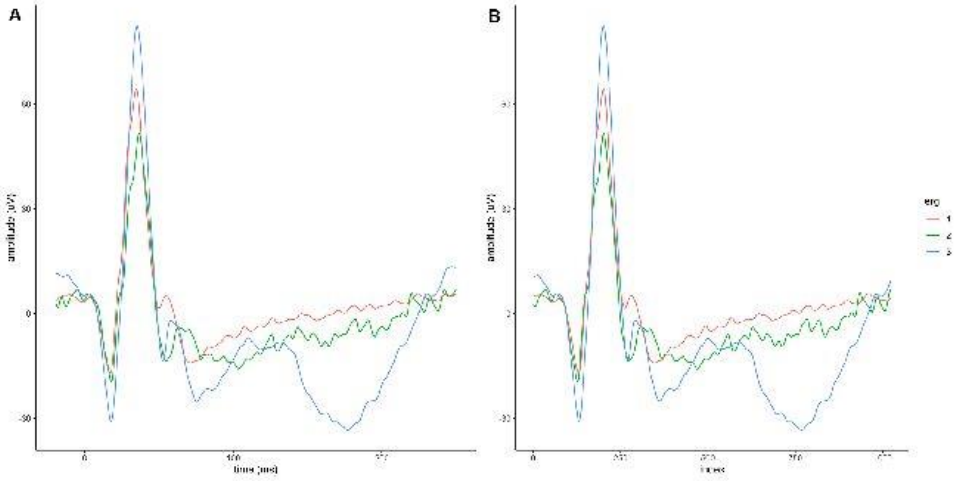


Figure 1. Alignment of the traces. Three ERG ensemble averages are illustrated. (A) Original traces plotted against post-stimulus time. (B) Alignment of the traces so that the b-wave peak occurs at sample 200 of 1024.

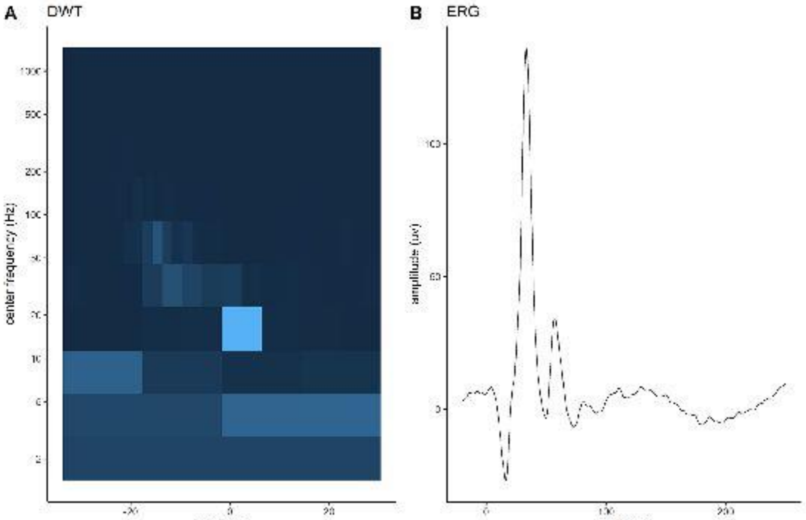


Figure 2. DWT of the ERG. (A) Magnitude of the detail coefficients of the DWT (scalogram). (B) ERG as a function of amplitude relative to time.

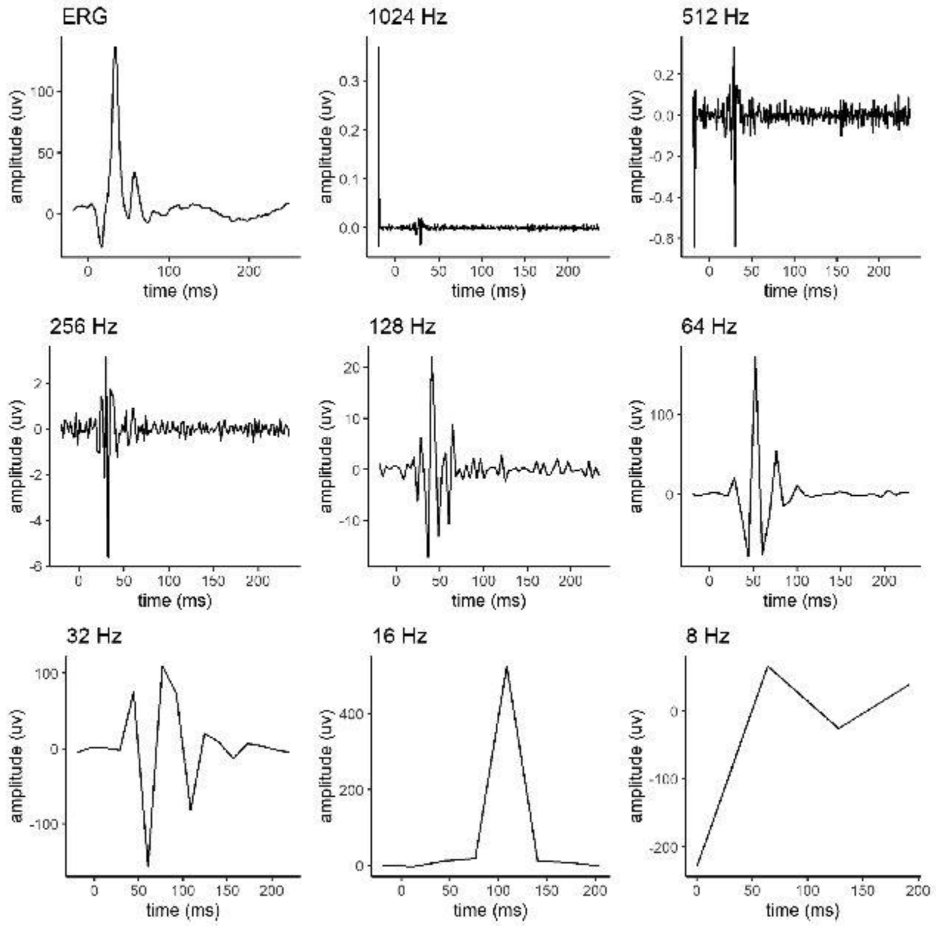


Figure 3. DWT decomposition of the ERG by level. The first panel represents the time-domain signal. Each subsequent panel shows the detail coefficients plotted on a horizontal axis scaled to the same length as the original time-domain signal. The panel label shows the center frequency of the wavelet. At the first decomposition level, the length of the signal is reduced by half and so on for each subsequent decomposition level.

Figure 3 shows each decomposition of the waveform presented in Figure 1 by the DWT, with the magnitude retaining the sign information. The horizontal axis of each subplot has been scaled to retain the same overall width for comparison.

The list of wavelets examined in this study includes the Coiflet, Symlet, Daubechies, and Best Localized wavelets over all available vanishing moments. These wavelets differ in terms of orthogonality (the extent to which the coefficients are correlated), symmetry (which provides linear phase), compact support, and filter order (an increased filter order of the mother wavelet increases smoothness).[48] Optimal wavelet selection was performed by minimizing the ratio of energy to Shannon entropy.[49] The energy for each wavelet was calculated by the sum of all the wavelet coefficients squared. The Shannon entropy was calculated by the plug-in estimator[50] for all of the coefficients over 20 bins. The plug-in estimator is a technique for estimating the entropy by using the empirically measured frequencies of the bins to directly estimate the probabilities used in the calculation. That is, $\hat{H} = -\sum_k \widehat{p_k} ln\widehat{p_k}$,

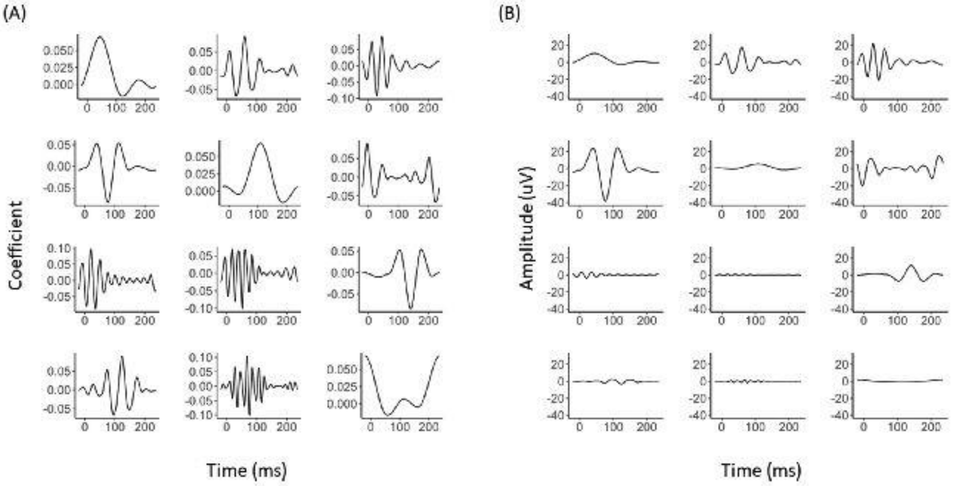


Figure 4. Sample matching pursuit decomposition. (A) The dictionary of atoms used for the decomposition and the coefficients for each, and (B) the atoms multiplied by the coefficient. Note that some of the coefficients are negative; the upper right trace of (A) has been inverted in (B). The decomposition consists of the coefficients for each atom in the dictionary.

where $\hat{H}$ is the estimated entropy for the k bins, and $\widehat{p_k}$, the estimated probability for the kth bin, is given simply by the count for that bin divided by the total count of all bins.

The ratio of energy to Shannon entropy for each wavelet over the ensemble was summed to determine the optimum wavelet family. Bootstrap resampling at the individual level (to account for between-eye correlations) was performed 1000 times initialized by a fixed seed to determine the 95% confidence intervals. The wavelet selection process used the input feature matrix only and not the outcome measure (in a similar fashion to principal components analysis), and thus did not require cross-validation. The optimum wavelet was used to extract features from the traces for subsequent predictive modeling. In this study, the 16 wavelet coefficients of the sixth decomposition level (center frequency of 32 Hz) were used for both the wavelet selection and the subsequent predictive modeling.

Matching Pursuit

The matching pursuit calculations were performed in MATLAB 2019b (MathWorks, Natick, MA) using the wavelet toolbox. Figure 4 illustrates a sample MP decomposition of the waveform presented in Figure 2 with the 12 atoms of highest importance shown. Figure 4A shows the atoms themselves and Figure 4B shows each atom multiplied by the coefficient. Summing all of the traces in Figure 4B would return the original signal.

Dictionaries were created in MATLAB for the matching pursuit algorithm. A length of 1024 was specified via the wmpdictionary function, and, for each wavelet packet family, dictionaries were created across a range of orders and vanishing moments. Each dictionary had a size of 1024 × 1024, representing the length of the signal and the number of atoms created. Dictionaries were created from the following wavelet packets: Coiflet, Daubechies, Meyer, Fejer–Korovkin, and Symlet. Like the DWT, these resulting dictionaries had varying properties. For example, the Feyer–Korovkin wavelet packet generates more concentrated wavelet packets with less leakage of high-frequency noise.[51] Optimization of the dictionary to the ensemble was performed over the choice of generating wavelet, the order, and the number of vanishing moments. For each dictionary, a smaller subdictionary of 25 atoms was selected as the top 25 atoms used within each full dictionary. The subdictionaries were compared to one another by decomposing and reconstructing each signal and finding the total root mean square error (RMSE) of the residuals over the whole ensemble. Reconstruction was performed with the wmpalg function using the orthogonal matching pursuits option and setting the number of iterations to 25 (i.e., the size of the smaller dictionary). The function returns the fit and the residual and the latter was used to calculate the RMSE for each fit. The best dictionary was considered that with the smallest RMSE. Bootstrap resampling was performed again in the same manner as for DWT to determine the 95% confidence intervals for the RMSE. As for the DWT, the outcome measure was not used for this step, so cross-validation was not required. The inputs to the MP model therefore consisted of 25 features corresponding to the coefficients for the atoms in the dictionary.

Predictive Modeling

To examine the value of time–frequency features from the ERG for predicting glaucoma severity based on the eRGC counts, predictive models utilizing the amplitude markers (a-wave, b-wave, i-wave, PhNR1, and PhNR2) were used as a base-case scenario for comparisons, given that we previously demonstrated that the use of this set of amplitude markers improved the prediction of glaucoma severity compared with the PhNR measure alone.[21] The performance of the base-case amplitude marker models was compared with that of models using these features plus additional DWT and/or MP features.

Two types of predictive models were fitted for each of the sets of features extracted from the ERGs. The models tested were a simple linear model and a multivariate adaptive regression spline (MARS) model,[52] an interpretable type of machine learning model that, as we have previously shown, improves the prediction of glaucoma severity with the ERG.[21] The outcome measure of a MARS model is a linear sum of piecewise linear functions. The Caret[53] package (version 6.0-86) in R was used for model fitting and tuning. For both linear and MARS models, overfitting was avoided by using 10-fold cross-validation. For the MARS models, the degree was set to 1 (that is, with no interaction terms), and the maximum number of terms was set to 40. MARS models were fitted using the Earth 5.3.0 package in R.[54]

For each of the tuned "final models," the performance in predicting the eRGC counts was evaluated by the proportion of variance of the eRGC counts explained (R^2) value. Each model was compared to the base case of the amplitude-based marker models. Comparisons were also made between linear and MARS models. The significance of difference in the prediction performance was determined by bootstrapped resampling ($n = 1000$ resamples with a fixed seed for the session) at the participant level.

Results

A total of 103 eyes with glaucoma from 55 participants were included in this study, and their characteristics are shown in Table 1. On the basis of the MD, 12 eyes (11.6%) had severe glaucoma ($\text{MD} < -12$ dB) and 16 eyes (15.5%) had moderate glaucoma (-12 dB $< \text{MD} < -6$ dB).

Discrete Wavelet Transform

Figure 5 shows the energy to entropy ratio for all the discrete wavelets used in the study. A higher ratio indicates a more optimum wavelet. With this metric for this set of traces, the Coiflet 8 was the best wavelet for the decomposition with an energy-to-entropy ratio of 9534 (SD = 2371) Using the Dunnett test for post hoc pairwise comparisons,[55] this wavelet was significantly better than all others ($P < 0.01$).

Matching Pursuit

The results for the matching pursuit dictionary selection are shown in Figure 6. Optimization was undertaken across mother wavelet, order of wavelet, and number of vanishing moments. The best overall dictionary was created by the symlet of order 6 (wpsym6) with six vanishing moments (VMs) having an RMSE of 33.5 nV (SD = 16.6 nV) Using the Dunnett test for post hoc pairwise comparisons, this dictionary was found to be significantly better ($P <$ 0.01) than all other dictionaries except symlets of order 6 with VM 7 ($P = 0.24$) and order 8 with VMs 7 and 8 ($P = 0.29$ and $P = 1.0$, respectively), as well as Coiflet

Table 1. Characteristics of the Individuals and Eyes With Glaucoma in the Study

Characteristic	
Individuals ($n = 55$)	
Age (y), median (IQR)	75 (66–80)
Gender (female), *n* (%)	21 (38)
Diabetes (present), *n* (%)	7 (13)
Hypertension (present), *n* (%)	35 (64)
Eyes ($n = 103$), median (IQR)	
Refraction sphere (D)	0.00 (−1.00 to 0.50)
Visual acuity (logMAR)	0.0 (−0.1 to 0.1)
IOP (mmHg)	15.0 (12.0 to 16.0)
Mean deviation (dB)	−2.5 (−5.9 to −0.5)
RNFL thickness (μm)	73 (63 to 83)
eRGC (1000s)	601 (470 to 753)

IQR, interquartile range.

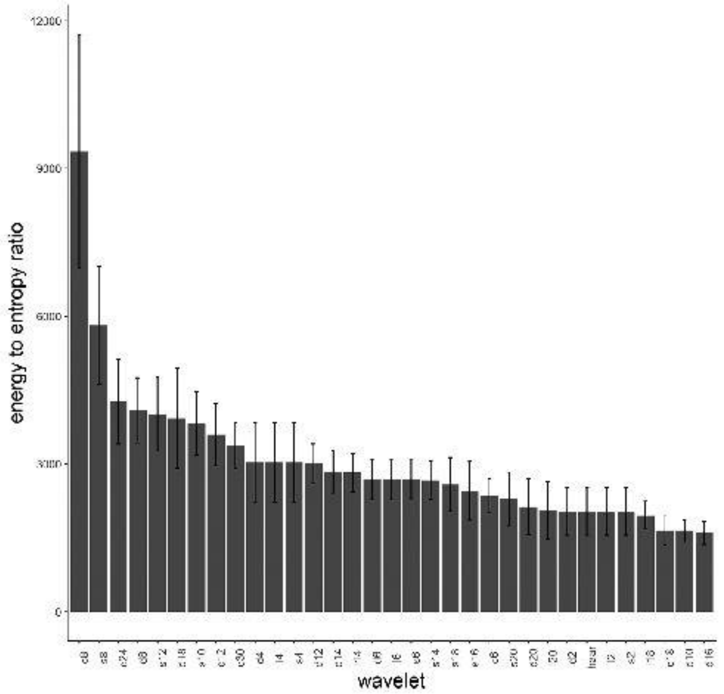


Figure 5. Energy-to-entropy ratio for all of the wavelets in descending order. Wavelets are identified by family and length. C, Coiflet; s, symlet; d, Daubechies; l, best localized. Higher values indicate more optimum discrete wavelets for the decomposition. The *error bars* represent the 95% confidence intervals determined by bootstrap resampling.

dictionaries with order 4 VM 6 ($P = 0.87$) and order 3 VM 7 ($P = 1.0$).

Predictive Modeling

Table 2 shows the summary of model performance for the base-case amplitude marker model compared to those with additional time–frequency features for both the linear and MARS models. Compared to a model using the ERG amplitude markers alone ($R^2 = 0.21$ and $R^2 = 0.34$ for the linear and MARS models, respectively), the models that additionally included the time–frequency features to the amplitude markers significantly improved the predictive performance of the eRGC counts ($R^2 \geq 0.41$ and $R^2 \geq 0.53$ for the linear and MARS models, respectively; $P \leq 0.011$) (Table 2). Pairwise differences among models 2, 3, and 4 did not reach significance within the model type (linear or MARS $P > 0.05$ for all comparisons).

Discussion

The use of electrophysiological tests to assess glaucoma is appealing in that it can be an objective measure of retinal function. The two established techniques are the PERG and the PhNR. The PERG is elicited with a patterned stimulus in which the pattern changes but the overall luminance remains constant over time.[56] Studies have shown changes in the amplitude markers of the transient[57] and steady-state[58] PERG. The PhNR elicited with a Ganzfeld dome has the advantage over the PERG of not requiring clear optics or refractive correction and may be technically easier to collect.[17]

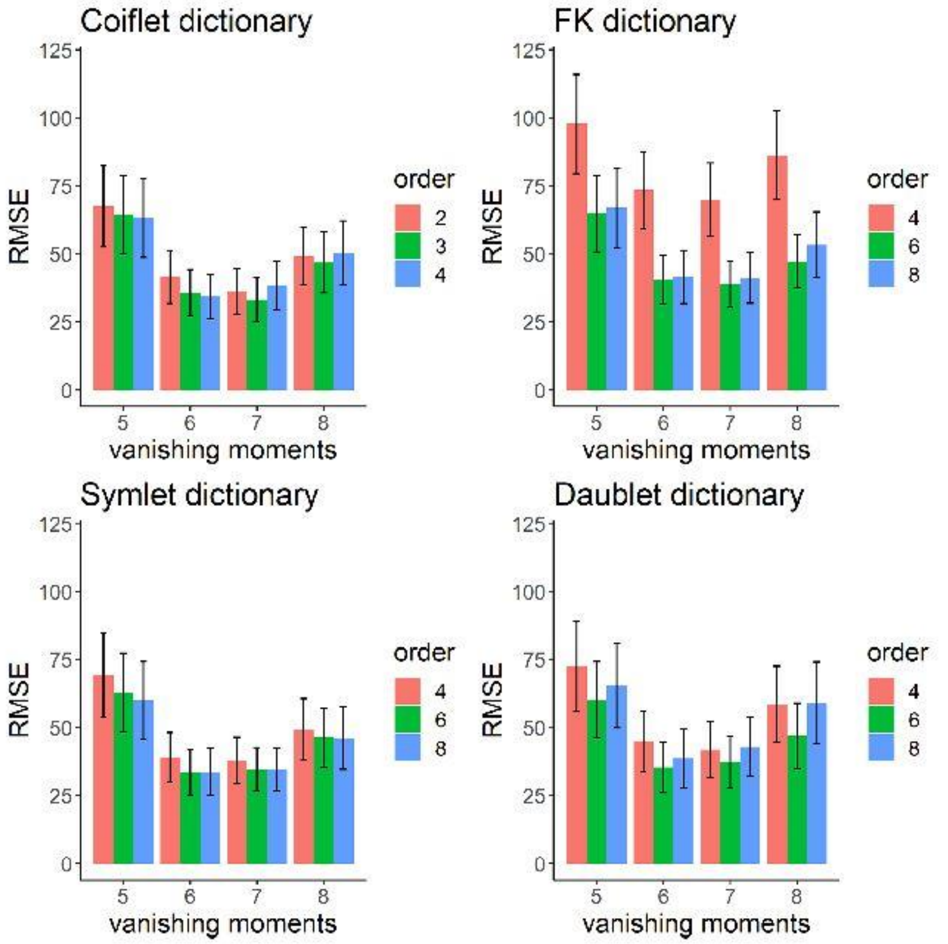


Figure 6. Matching pursuit dictionary optimization. FK, Fejer–Korovkin wavelet. The bar graphs show the residual sum of squares from reconstruction of the ensemble traces for different mother wavelet packets (by panel), vanishing moments (abscissa), and order of wavelet packet (by *color*). *Error bars*: 95% confidence intervals.

This study, which collected electroretinograms with a PhNR protocol, demonstrated that features from time–frequency analyses of the photopic ERG—when using either the DWT or MP—significantly improved the prediction of glaucoma severity when added to amplitude markers evaluated in the time-domain signal. These findings underscore how there is more information in the ERG about RGC dysfunction that could be uncovered using time–frequency analyses, which could potentially aid in the prediction of future RGC loss in glaucoma.

The ERG is a mass signal response, the sum of many different generating components from the inner and outer retina. Feedback loops are well described within the retina,[24,59,60] and RGC discharge spike trains have been shown to spontaneously oscillate in a variety of species.[61] These oscillations and the characteristic waveforms could have a distinctive time–frequency pattern; thus, it is possible that RGC dysfunction occurring in glaucoma could be identifiable with time–frequency analyses of the ERG. Our findings that the combination of the ERG time-domain

Table 2. Performance of Linear and MARS Models in Predicting eRGC From Time and Time–Frequency Features Derived From ERGs

	Linear		MARS		
	R^2	P^a	R^2	P^a	P, Linear vs. MARS
Model 1: Markers	0.21 (0.07)	—	0.34 (0.06)	—	0.004
Model 2: DWT + markers	0.41 (0.08)[b]	0.001	0.53 (0.07)[b]	0.011	0.055
Model 3: MP + markers	0.43 (0.08)[b]	0.001	0.63 (0.08)[b]	<0.001	0.002
Model 4: DWT + MP + markers	0.50 (0.09)[b]	0.001	0.63 (0.07)[b]	0.001	<0.001

Model 1 was informed by the amplitudes of the a-wave, b-wave, i-wave, PhNR1, and PhNR2. Numbers in parentheses indicate standard deviation from the bootstrap resampling.

[a] When compared to model 1.

[b] Not significantly different at $P < 0.05$ for pairwise comparisons among models 2, 3, and 4 for the linear and MARS models separately.

amplitude markers and time–frequency features improved the prediction of glaucoma severity is consistent with observations seen in other fields. One previous study showed that the combination of time and time–frequency features from the electroencephalogram improved the prediction of the error rate for a behavioral task.[62] Another study demonstrated that the addition of time–frequency features to the time-domain features also improved the prediction of sudden cardiac death from heart rate variability data from the electrocardiogram.[63]

The findings of this study build upon our previous work, further reinforcing how there is substantially more information present in the photopic ERG beyond the PhNR amplitude or ratio of PhNR to b-wave alone,[21] which currently remain the primary recommended measures for assessing RGC function on the ERG.[11] However, the improved prediction of glaucoma severity at cross-section when incorporating features from time–frequency analyses does not guarantee its improved utility for the clinically important task of predicting RGC loss in glaucoma, which remains a significant challenge.[64,65] The findings in this study provide further evidence to support the conduct of future longitudinal studies to explore whether the ERG could provide added clinical utility for this purpose. Furthermore, recovery of the PhNR in some eyes with glaucoma following IOP lowering and nicotinamide has been previously described,[18,19] and it may be that time–frequency analysis may also show such functional recovery better than amplitude measurements of the PhNR alone. The ERG as a measure of function rather than structure may not be able to distinguish between reversible and irreversible loss of RGC function. However reversible RGC dysfunction could potentially be identified through evaluating the discordance between the observed and expected RGC function based on estimates of RGC loss (such as through the eRGC counts parameter in this study), which could also be examined in future longitudinal studies. Further exploration of the nature of feedback and feedforward pathways could be done with hierarchical decomposition analysis.[66] This technique uses multiple electrodes to decompose a set of signals into their constituent sources. The technique assumes that there is a dominant generator with subsequent generators dependent upon the dominant one—hence, the hierarchy. This technique may become possible with multi-electrode contact lens electrodes such as those that have been used in the rat.[67]

Limitations of this study include the relatively small sample size and the limited range of glaucoma severity, but the cohort nonetheless enabled the demonstration of the significant added value of features from time–frequency analyses. We used the analysis of ensemble averages with timing relative to the peak of the b-wave rather than relative to the flash onset to account for latency and phase variations among individuals. This is an alternative technique to the local wavelet maximum technique used by Gauvin and colleagues[68] and does have the advantage that the decomposition is invertible. This technique results in the loss of timing information, although timing features are generally not used for analysis of the PhNR. Both the MP and the DWT optimization were performed on glaucoma participants alone, and it may be that inclusion of eyes without glaucoma could yield a different result. Construction of a mixed cohort for test optimization would require knowledge of the prior probability of the existence and extent of glaucoma in an eye undergoing the test. Such an approach was beyond the scope of this study. Although in our study better performance was obtained with the MP technique, the DWT is more intuitive, and our DWT methods are easier to replicate by others compared to the MP decomposition, which requires access to our custom dictionary.

In this work, we did not attempt to establish that the electroretinogram can better estimate the glaucoma severity than existing techniques such as the eRGC. To do so would require a ground truth such as histological ganglion cell counts, and even this would not necessarily reveal the functional state of the RGCs. We have, however, shown that conventional time-domain amplitude analysis techniques might not be extracting all of the useful information available from the ERG in glaucoma.

Conclusions

This study demonstrated that features extracted from the ERG using time–frequency analyses yield additional predictive information about the severity of glaucoma. This is biologically plausible given that the ERG is made up of a sequence of different processes with different frequency responses that can become impaired in glaucoma. These findings encourage future longitudinal studies to understand whether this technique could be clinically useful for predicting glaucoma progression.

Acknowledgments

Supported in part by a Hugh Noel Puckle Scholarship (MS) and by grants from the National Health and Medical Research Council (1104985 and 2008382 to ZW).

Disclosure: **M. Sarossy**, None; **J. Crowston**, None; **D. Kumar**, None; **A. Weymouth**, None; **Z. Wu**, None

References

1. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA*. 2014;311:1901–1911.
2. Fry LE, Fahy E, Chrysostomou V, et al. The coma in glaucoma: retinal ganglion cell dysfunction and recovery. *Prog Retin Eye Res*. 2018;65:77–92.
3. Kalesnykas G, Oglesby EN, Zack DJ, et al. Retinal ganglion cell morphology after optic nerve crush and experimental glaucoma. *Invest Ophthalmol Vis Sci*. 2012;53:3847–3857.
4. Della Santina L, Inman DM, Lupien CB, Horner PJ, Wong RO. Differential progression of structural and functional alterations in distinct retinal ganglion cell types in a mouse model of glaucoma. *J Neurosci*. 2013;33:17444–17457.
5. Risner ML, Pasini S, Cooper ML, Lambert WS, Calkins DJ. Axogenic mechanism enhances retinal ganglion cell excitability during early progression in glaucoma. *Proc Natl Acad Sci USA*. 2018;115:E2393–E2402.
6. Calkins DJ. Adaptive responses to neurodegenerative stress in glaucoma. *Prog Retin Eye Res*. 2021;84:100953.
7. Weber AJ, Harman CD. Structure–function relations of parasol cells in the normal and glaucomatous primate retina. *Invest Ophthalmol Vis Sci*. 2005;46:3197–3207.
8. Tham Y-C, Li X, Wong TY, Quigley HA, Aung T, Cheng C-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121:2081–2090.
9. Banitt MR, Ventura LM, Feuer WJ, et al. Progressive loss of retinal ganglion cell function precedes structural loss by several years in glaucoma suspects. *Invest Ophthalmol Vis Sci*. 2013;54:2346–2352.
10. Liu HH, He Z, Nguyen CT, Vingrys AJ, Bui BV. Reversal of functional loss in a rat model of chronic intraocular pressure elevation. *Ophthalmic Physiol Opt*. 2017;37:71–81.
11. Frishman L, Sustar M, Kremers J, et al. ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram. *Doc Ophthalmol*. 2018;136:207–211.
12. Viswanathan S, Frishman LJ, Robson JG, Walters JW. The photopic negative response of the flash electroretinogram in primary open angle glaucoma. *Invest Ophthalmol Vis Sci*. 2001;42:514–522.
13. Machida S, Gotoh Y, Toba Y, Ohtaki A, Kaneko M, Kurosaka D. Correlation between photopic negative response and retinal nerve fiber layer thickness and optic disc topography in glaucomatous eyes. *Invest Ophthalmol Vis Sci*. 2008;49:2201–2207.
14. North RV, Jones AL, Drasdo N, Wild JM, Morgan JE. Electrophysiological evidence of early functional damage in glaucoma and ocular hypertension. *Invest Ophthalmol Vis Sci*. 2010;51:1216–1222.
15. Nakamura H, Hangai M, Mori S, Hirose F, Yoshimura N. Hemispherical focal macular photopic negative response and macular inner retinal thickness in open-angle glaucoma. *Am J Ophthalmol*. 2011;151:494–506.e1.
16. Machida S, Kaneko M, Kurosaka D. Regional variations in correlation between photopic

negative response of focal electoretinograms and ganglion cell complex in glaucoma. *Curr Eye Res*. 2015;40:439–449.

17. Preiser D, Lagreze WA, Bach M, Poloschek CM. Photopic negative response versus pattern electroretinogram in early glaucoma. *Invest Ophthalmol Vis Sci*. 2013;54:1182–1191.
18. Tang J, Hui F, Hadoux X, et al. Short-term changes in the photopic negative response following intraocular pressure lowering in glaucoma. *Invest Ophthalmol Vis Sci*. 2020;61:16.
19. Hui F, Tang J, Williams PA, et al. Improvement in inner retinal function in glaucoma with nicotinamide (vitamin B3) supplementation: a crossover randomized clinical trial. *Clin Exp Ophthalmol*. 2020;48:903–914.
20. Rosolen SG, Rigaudière F, LeGargasson JF, et al. Comparing the photopic ERG i-wave in different species. *Vet Ophthalmol*. 2004;7:189–192.
21. Sarossy M, Crowston J, Kumar D, Weymouth A, Wu Z. Prediction of glaucoma severity using parameters from the electroretinogram. *Sci Rep*. 2021;11:23886.
22. Robson JG, Frishman LJ. Dissecting the dark-adapted electroretinogram. *Doc Ophthalmol*. 1998;95:187–215.
23. Sieving PA, Murayama K, Naarendorp F. Push–pull model of the primate photopic electroretinogram: a role for hyperpolarizing neurons in shaping the b-wave. *Vis Neurosci*. 1994;11:519–532.
24. Crook JD, Manookin MB, Packer OS, Dacey DM. Horizontal cell feedback without cone type-selective inhibition mediates "red–green" color opponency in midget ganglion cells of the primate retina. *J Neurosci*. 2011;31:1762–1772.
25. Dong C-J, Agey P, Hare WA. Origins of the electroretinogram oscillatory potentials in the rabbit retina. *Vis Neurosci*. 2004;21:533.
26. Puller C, Haverkamp S, Neitz M, Neitz J. Synaptic elements for GABAergic feed-forward signaling between HII horizontal cells and blue cone bipolar cells are enriched beneath primate S-cones. *PLoS One*. 2014;9:e88963.
27. Morlet J. Sampling theory and wave propagation. In: Chen CH, ed. *Issues in Acoustic Signal—Image Processing and Recognition*. Berlin: Springer; 1983:233–261.
28. Goupillaud P, Grossmann A, Morlet J. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*. 1984;23:85–102.
29. Hassankarimi H, Noori SMR, Jafarzadehpour E, Yazdani S, Radinmehr F. Analysis of pattern electroretinogram signals of early primary open-angle glaucoma in discrete wavelet transform coefficients domain. *Int Ophthalmol*. 2019;39:2373–2383.
30. Miguel-Jiménez J, Boquete L, Ortega S, Rodriguez-Ascariz J, Blanco R. Glaucoma detection by wavelet-based analysis of the global flash multifocal electroretinogram. *Med Eng Phys*. 2010;32:617–622.
31. Kundra H, Park JC, McAnany JJ. Comparison of photopic negative response measurements in the time and time–frequency domains. *Doc Ophthalmol*. 2016;133:91–98.
32. Daubechies I. *Ten Lectures on Wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1992.
33. Forte JD, Bui BV, Vingrys AJ. Wavelet analysis reveals dynamics of rat oscillatory potentials. *J Neurosci Methods*. 2008;169:191–200.
34. Behbahani S, Ramezani A, Karimi Moridani M, Sabbaghi H. Time–frequency analysis of photopic negative response in CRVO patients. *Semin Ophthalmol*. 2020;35:187–193.
35. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell*. 1989;11:674–693.
36. Daubechies I. Orthonormal bases of compactly supported wavelets. *Commun Pure Appl Math*. 1988;41:909–996.
37. Kumar P, Foufoula-Georgiou E. Wavelet analysis for geophysical applications. *Rev Geophys*. 1997;35:385–412.
38. Gauvin M, Sustar M, Little JM, Brecelj J, Lina J-M, Lachapelle P. Quantifying the ON and OFF contributions to the flash ERG with the discrete wavelet transform. *Transl Vis Sci Technol*. 2017;6:3.
39. Constable PA, Marmolejo-Ramos F, Gauthier M, Lee IO, Skuse DH, Thompson DA. Discrete wavelet transform analysis of the electroretinogram in autism spectrum disorder and attention deficit hyperactivity disorder. *Front Neurosci*. 2022;16:890461.
40. Mallat S, Zhang Z. *Matching Pursuit with Time–Frequency Dictionaries*. New York: Courant Institute of Mathematical Sciences; 1993.
41. Hsu P-H. Feature extraction of hyperspectral images using wavelet and matching pursuit. *ISPRS J Photogramm Remote Sens*. 2007;62:78–92.
42. Zarei A, Asl BM. Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals. *Comput Biol Med*. 2021;131:104250.
43. Medeiros FA, Lisboa R, Weinreb RN, Girkin CA, Liebmann JM, Zangwill LM. A combined index of structure and function for staging glaucomatous damage. *Arch Ophthalmol*. 2012;130:1107–1116.

44. Harwerth R, Wheat J, Fredette M, Anderson D. Linking structure and function in glaucoma. *Prog Retin Eye Res*. 2010;29:249–271.
45. Wu Y, Cun Q, Tao Y, et al. Evaluation of macular and retinal ganglion cell count estimates for detecting and staging glaucoma. *Front Med*. 2021;8:740761.
46. Wakili N, Horn FK, Jünemann AG, et al. The photopic negative response of the blue-on-yellow flash-electroretinogram in glaucomas and normal subjects. *Doc Ophthalmol*. 2008;117:147–154.
47. DB Percival, Walden AT. *Wavelet Methods for Time Series Analysis*. Cambridge, UK: Cambridge University Press; 2008.
48. Adamo F, Andria G, Attivissimo F, Lanzolla AML, Spadavecchia M. A comparative study on mother wavelet selection in ultrasound image denoising. *Measurement*. 2013;46:2447–2456.
49. Shigueoka LS, de Vasconcellos JPC, Schimiti RB, et al. Automated algorithms combining structure and function outperform general ophthalmologists in diagnosing glaucoma. *PLoS One*. 2018;13:e0207784.
50. Zhang Z, Zhang X. A normal law for the plug-in estimator of entropy. *IEEE Trans Inform Theory*. 2011;58:2745–2747.
51. Kadir R, Saha R, Akhter MM, Awal MA, Kadir MI. LDPC coded hybrid discrete cosine transform and Fejér–Korovkin wavelet transform-based SC-FDMA for image communication. *Array*. 2021;12:100107.
52. Friedman JH. Multivariate adaptive regression splines. *Ann Statist*. 1991;19:1–67.
53. Kuhn M. caret: classification and regression training. Available at: https://cran.r-project.org/web/packages/caret/caret.pdf. Accessed September 26, 2022.
54. Milborrow S. earth: multivariate adaptive regression splines. Available at: https://cran.r-project.org/web/packages/earth/earth.pdf. Accessed September 26, 2022.
55. Vizzeri G, Tafreshi A, Weinreb RN, Bowd C. Effect of operator and optical defocus on the variability of pattern electroretinogram optimized for glaucoma detection (PERGLA). *J Glaucoma*. 2010;19:77.
56. Bach M, Brigell MG, Hawlina M, et al. ISCEV standard for clinical pattern electroretinography (PERG): 2012 update. *Doc Ophthalmol*. 2013;126:1–7.
57. Parisi V, Manni G, Centofanti M, Gandolfi SA, Olzi D, Bucci MG. Correlation between optical coherence tomography, pattern electroretinogram, and visual evoked potentials in open-angle glaucoma patients. *Ophthalmology*. 2001;108:905–912.
58. Bach M, Unsoeld AS, Philippin H, et al. Pattern ERG as an early glaucoma indicator in ocular hypertension: a long-term, prospective study. *Invest Ophthalmol Vis Sci*. 2006;47:4881–4887.
59. Purpura K, Tranchina D, Kaplan E, Shapley RM. Light adaptation in the primate retina: analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Vis Neurosci*. 1990;4:75–93.
60. Davenport CM, Detwiler PB, Dacey DM. Effects of pH buffering on horizontal and ganglion cell light responses in primate retina: evidence for the proton hypothesis of surround formation. *J Neurosci*. 2008;28:456–464.
61. Neuenschwander S, Castelo-Branco M, Singer W. Synchronous oscillations in the cat retina. *Vision Res*. 1999;39:2485–2497.
62. Munneke G-J, Nap TS, Schippers EE, Cohen MX. A statistical comparison of EEG time- and time–frequency domain representations of error processing. *Brain Res*. 2015;1618:222–230.
63. Ebrahimzadeh E, Pooyan M, Bijar A. A novel approach to predict sudden cardiac death (SCD) using nonlinear and time–frequency analyses from HRV signals. *PLoS One*. 2014;9:e81896.
64. Ernest PJ, Schouten JS, Beckers HJ, Hendrikse F, Prins MH, Webers CA. Prediction of glaucomatous visual field progression using baseline clinical data. *J Glaucoma*. 2016;25:228–235.
65. De Moraes CG, Sehi M, Greenfield DS, Chung YS, Ritch R, Liebmann JM. A validated risk calculator to assess risk and rate of visual field progression in treated glaucoma patients. *Invest Ophthalmol Vis Sci*. 2012;53:2702–2707.
66. Repucci M, Schiff N, Victor J. General strategy for hierarchical decomposition of multivariate time series: implications for temporal lobe seizures. *Ann Biomed Eng*. 2001;29:1135–1149.
67. Derafshi Z, Kunzer BE, Mugler EM, et al. Corneal potential maps measured with multi-electrode electroretinography in rat eyes with experimental lesions. *Invest Ophthalmol Vis Sci*. 2017;58:2863–2873.
68. Gauvin M, Little JM, Lina J-M, Lachapelle P. Functional decomposition of the human ERG based on the discrete wavelet transform. *J Vis*. 2015;15:14.