

Sensitivity and Specificity of Multifocal Electroretinography in Detecting Chloroquine and Hydroxychloroquine retinal toxicity

Sina Ahmadi Pirshahid

This thesis is submitted to the Faculty of Graduate and Postdoctoral Studies as a partial fulfillment of the M.Sc. program in Cellular and Molecular Medicine

Submitted November 2014

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

©Sina Ahmadi Pirshahid, Ottawa, Canada, 2015

Contents

Abstract.....	iv
List of Figures.....	v
List of Tables.....	vi
List of abbreviations.....	vii
Acknowledgements.....	viii
1-Introduction.....	1
1.1-Overview (reference citations have been made in the subsequent sections).....	1
1.2- Historical background.....	2
1.3-Pharmacokinetics and mechanisms of anti-inflammatory action.....	2
1.4- CQ and HCQ as rheumatologic and dermatologic drugs.....	3
1.5- Epidemiological concerns.....	5
1.6- Future of CQ and HCQ use in rheumatic disease.....	5
1.7- CQ/HCQ ocular toxicity.....	6
1.7.1- Cellular mechanisms and histopathology findings.....	7
1.7.2- Natural course and clinical picture.....	7
1.7.3- Incidence.....	11
1.7.4- Risk factors.....	12
1.7.5- Prognosis and clinical approach.....	13
1.8- CQ/HCQ retinal toxicity screening.....	13
1.8.1- Automated visual field (AVF) testing.....	16
1.8.2- Multifocal ERG.....	21
1.8.3-Fundus auto fluorescence (FAF).....	27
1.8.4- Spectral Domain Optical Coherence Tomography (sdOCT).....	29
1.9- Design and purpose of this study.....	30
2-Materials and Methods.....	36
2.1- Humphrey Field Analyzer 10-2 automated perimetry.....	38
2.2- Multifocal ERG.....	45
2.3- Spectral domain OCT and FAF.....	46
2.4- Statistical analysis.....	49

2.4.1- Relationship between target and reference test.....	49
2.4.2- Sensitivity and specificity.....	49
3-Results.....	52
3.1- Demographics.....	52
3.2- Risk factors.....	53
3.3- The influence of daily dose.....	55
3.4- Correlation between cumulative dose and test results.....	56
3.5- Sensitivity and specificity of mfERG.....	63
3.6- Relevance of mfERG test results to CQ/HCQ retinal toxicity.....	66
3.7- Analysis of mfERG false positives.....	69
3.8- FAF findings.....	71
4- Discussion.....	72
4.1- The influence of risk factors in the development of toxicity.....	76
4.1.1- Correlation between test results and cumulative dose.....	78
4.2- The model for mfERG sensitivity and specificity calculation.....	81
4.2.1- Inclusion of one or both eyes.....	81
4.2.2- Humphrey 10-2 AVF as the mainstay of screening.....	84
4.3- Sensitivity and specificity of mfERG.....	86
4.3.1- Interpretation of the mfERG “false positive” cases.....	86
4.3.2- Current knowledge of sensitivity and specificity of the screening tests.....	87
4.3.3- Validity of mfERG independent of the reference test.....	89
5- Conclusion.....	91
References.....	92

Abstract

To calculate the sensitivity and specificity of multifocal electroretinography (mfERG) in detection of chloroquine and hydroxychloroquine retinal toxicity, 120 eyes of 63 patients were evaluated using the currently recommended diagnostic tests. The results were compared to those of 54 eyes of 28 control subjects. The sensitivity and specificity of mfERG relative to the combination of automated visual fields and optical coherence tomography (the reference test) were calculated to be 87% and 86.5% respectively. However, analysis of the “false positive” cases proved that mfERG was more sensitive than the reference test and the actual sensitivity and specificity values were higher than the results of this study. Reduction of mfERG amplitude was a strong and reliable sign of early retinal toxicity and was correlated with the cumulative dose of hydroxychloroquine. This correlation was not observed with the reference test quantitative values.

List of Figures

Figure 1 Color fundus photograph of both eyes showing bilateral “bull’s eye maculopathy” due to hydroxychloroquine toxicity	10
Figure 2: Humphrey Field Analyzer 10-2 Single Field Analysis Printout.	20
Figure 3: Multifocal ERG waveform.	23
Figure 4: Multifocal ERG test result of a normal eye.	26
Figure 5: FAF and corresponding sdOCT images.....	32
Figure 6: Normal sdOCT foveal cross section	33
Figure 7: sdOCT retinal thickness map.....	34
Figure 8: sdOCT image of normal retina and 2 different stages of toxicity	35
Figure 9: Humphrey 10-2 AVF Grading of the Pattern Deviation Plot.....	40
Figure 10: Single field analysis of the left eye representing a characteristic early “bull’s eye” scotoma.....	42
Figure 11: Humphrey 10-2 single field analysis printout illustrating advanced Hydroxychloroquine retinal toxicity	43
Figure 12: Humphrey 10-2 AVF progression map showing progressive appearance of new abnormalities in the visual field of a HCQ patient in 3 consecutive years.....	44
Figure 13: Early and late HCQ retinal toxicity detected by FAF.	48
Figure 14: 2X2 table used to calculate the validity of a target test relative to the reference test.	49
Figure 15: Prevalence of Risk Factors in Toxic and Non-toxic Patients based on the reference test results.....	54
Figure 16: Frequency of Positive Test Results in CQ & HCQ Groups	58
Figure 17A-G: Regression and ANOVA analysis between cumulative dose and quantitative data of the tests results.....	62
Figure 18: Interaction line plots demonstrating statistically significant differences in P1 amplitude R2/R5 (Left) and R3/R5 (Right) between the toxic and non-toxic eyes.....	67
Figure 19: mfERG deviation map (Top), sdOCT imaging (Middle) and FAF (Bottom) of both eyes of a patient with HCQ retinal toxicity.....	80
Figure 20: Asymmetry of retinal damage in HCQ retinal toxicity between two eyes using all test modalities.....	83

List of Tables

Table 1: The regression analysis p-values between cumulative dose and test results.....	59
Table 2: Four different categories for mfERG test results relative to the reference test	64
Table 3: Calculated sensitivity and specificity of mfERG with 95% confidence interval	65
Table 4: ANOVA P-value results demonstrating mfERG P1 amplitude R5 ring ratio differences between reportedly toxic, borderline, non-toxic and control groups. Statistically significant differences are limited to the P1 amplitude R5 ring ratios of the parafoveal rings, R2 and R3 between the toxic and non-toxic groups.	68
Table 5: ANOVA P-value results demonstrating the P1 amplitude R5 ring ratio differences between the “False positive”, non-toxic and control groups.....	70
Table 6: Sensitivity and specificity of 10-2 AVF, mfERG, sdOCT as determined by Browning and Lee (Browning and Lee 2014)	88

List of abbreviations

AAO	American Academy of Ophthalmology
ACR	American College of Rheumatology
ARMD	Age Related Macular Degeneration
AVF	Automated Visual Field
CQ	Chloroquine Phosphate
DLE	Discoid Lupus Erythematosus
DMARDs	Disease Modifying Anti-Rheumatic Drugs
ETDRS	Early Treatment Diabetic Retinopathy Study
FAF	Fundus Auto Fluorescence
HCQ	Hydroxychloroquine sulphate
LAOs	Lysosome Associated Organelles
LBW	Lean Body Weight
LF	Lipofuscin
MCBs	Membranous Cytoplasmic Bodies
MD	Mean Deviation
mfERG	Multifocal Electroretinography
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
PSD	Pattern Standard Deviation
RA	Rheumatoid Arthritis
RPE	Retinal Pigment Epithelium
SARD	Systemic Autoimmune Rheumatic Disease
	Spectral Domain Optical Coherence
sdOCT	Tomography
SLE	Systemic Lupus Erythematosus
WHO	World Health Organization

Acknowledgements

I would like to express the deepest appreciation to my supervisor Dr. Stuart G. Coupland for the useful advice, remarks and dedication through the process of this thesis. You have been a tremendous mentor to me in encouraging my research and keeping me on the right track. What I learned from you is far beyond the knowledge of retinal electrophysiology; you made me feel much more confident in planning and implementing research projects.

I would also like to express gratitude to my co-supervisor, Dr. Chloe Gottlieb. Your commitment, thoughtfulness, and brilliant comments played an enormous role in helping me to achieve the project's goals.

I wish to thank my committee member, Dr. Catherine Tsilfidis for her support and guidance. It was a great honor to take advantage of your bright ideas and critical approach to scientific methods.

I would like to express my appreciation for John Hamilton's support. Undoubtedly, you are an asset to the Ottawa Hospital Eye Institute, much more than merely a research coordinator. This project could never be accomplished without your affiliation and wise comments.

I have been very fortunate over the past two years to have received an admission scholarship from the University of Ottawa and financial support from members of the Ottawa Hospital Eye Institute, Dr. Stuart G. Coupland, Dr. Chloe Gottlieb, Dr. Bernard Hurley and Dr. Brian Leonard.

Over the period of my Master's program, I received assistance and useful comments from bright medical students, Adrian Tsang and Alison Sifton, and ophthalmology resident, Jennifer Gao. You will soon make very famous scientists and I will always be proud of your collaboration.

Last, but not least, I would like to mention and appreciate superb administrative and technical assistance that I received from Jane Canniff, Debbie Rose-Gallant, Kathya Molino and Rita Balabanian-Garnett.

1-Introduction

1.1-Overview (reference citations have been made in the subsequent sections)

Chloroquine phosphate (CQ) and hydroxychloroquine sulphate (HCQ) are synthetic anti-malarial drugs introduced in the 1940s. Despite their relatively safe systemic profile compared to other rheumatic disease treatment options, their long-term use can become associated with retinal toxicity. The only strategy to preserve vision is to diagnose the toxicity in an early, potentially reversible stage through a screening program. A series of functional and structural tests have been recommended for this purpose. This study is designed to determine the validity of retinal electrophysiologic testing in the early detection of the toxicity.

The purpose of the introduction is to address the following points:

- The significance of CQ and HCQ retinal toxicity and its early detection
- The cellular basis of toxicity
- The histologic and clinically recordable signs of related cellular stress in retinal neurons
- The purpose of the study and the hypotheses to be tested

In this section, the historical background, epidemiology and significance of screening protocols in patients' vision care are outlined. This is followed by a brief review of the diagnostic tests currently recommended for screening of CQ/HCQ retinal toxicity.

1.2- Historical background

CQ and HCQ are two of the well-known antimalarial drugs, a group of chemotherapeutics used in the treatment of malaria. CQ was first introduced during World War II and utilized by the military forces as an effective antimalarial drug. HCQ was added to the malaria treatment protocols in 1945. Shortly thereafter, they were shown to be effective agents in the treatment of inflammatory disease with HCQ being equally effective and less systemically toxic than its predecessor, CQ (Scherbel et al. 1958, Palma Sánchez et al. 2013). HCQ was also shown to be less toxic to ocular structures (Finbloom 1985). A theory to explain this difference was presented by Raines and colleagues. They proposed that, in contrast to HCQ, CQ could destroy the role of the retinal pigment epithelium (RPE) as part of the blood-retinal-barrier, thereby permitting easier access of CQ to the retina (Raines et al. 1989).

1.3-Pharmacokinetics and mechanisms of anti-inflammatory action

CQ and HCQ are closely related and structurally similar chemical compounds. HCQ is the product of beta hydroxylation of CQ (Ochsendorf 2010). They are almost completely absorbed when prescribed orally (McChesney 1983). Due to their pharmacokinetic properties, their achievement of biological steady state is slow; a period of about 6 months might be required for the full clinical response in the treatment of rheumatic disease (Tett et al. 1989). These drugs are also characterized by a relatively long elimination time; half-life of HCQ might be as long as 30 days, HCQ can be detected in blood and urine for 5 months after discontinuation (Tett et al. 1989). The relatively long elimination period has been explained by the sluggish redistribution of the

drug from its large tissue reservoirs like muscle and liver (Fox 1993). When dispersed in bodily tissues, the eye receives a significant amount of these drugs (McChesney 1967) while their distribution in the fat tissue is almost negligible (McChesney 1983). CQ and HCQ show a high tendency for melanin containing tissues like the uvea (Ono et al. 2003). Inside the cells, these drugs tend to accumulate in the acidic intracellular vacuoles, mainly lysosomes (Mahon 2004).

Exact mechanisms of CQ and HCQ anti-inflammatory actions are not yet fully understood. A detailed discussion of this subject has been presented by Fox (Fox 1993). According to this article, some of the more prominent effects are: elevation of pH within the lysosomes and interference with antigen processing and presentation, down-regulation of T-helper cell activities, and reduced production of acute phase reactants. Moreover, the drugs can directly inhibit the stimulated monocytes, resulting in further reduction of interleukins and tumor necrosis factor.

1.4- CQ and HCQ as rheumatologic and dermatologic drugs

Long term clinical experience indicates that while steroids and non-steroidal anti-inflammatory drugs (NSAIDs) are useful in reducing signs and symptoms of inflammation, prevention of irreversible tissue damage and deformities mandates chronic prescription of a combination of “disease modifying anti rheumatic drugs (DMARDs)” including HCQ (O’Dell et al. 1996).

Discoid and Systemic Lupus Erythematosus (DLE and SLE) were the first autoimmune diseases widely treated with antimalarial agents (Scherbel 1983). This

indication has not changed over time; a recently published review article recommended prescription of HCQ over the entire course of SLE. In addition, it denotes the systemic safety of HCQ throughout pregnancy in SLE patients (Ruiz-Irastorza et al. 2010). HCQ is currently recommended by the American College of Rheumatology (ACR) for the treatment of rheumatoid arthritis (RA) (Singh et al. 2012, Morris et al. 2011). There is strong evidence in favor of DMARDs use in all cases of RA, whether mild, moderate or severe; HCQ plays a prominent role in every treatment protocol (Davis et al. 1991, Suarez-Almazor et al. 2000).

Therapeutic effects of antimalarial drugs in autoimmune disease therapy are not limited to RA and SLE; Sjögren syndrome, dermatomyositis, sarcoidosis and several other chronic inflammatory conditions are other potential indications for long-term treatment with HCQ mono or combination therapy (Ochsendorf 2010). CQ and HCQ have moderate potency in modulating the immune system activity and must be combined with other anti-inflammatory agents in more severe cases (O'Dell et al. 2002). Nevertheless, in contrast to most other therapies, they do not increase the risk of secondary infections, bone marrow suppression and malignancy (Mavrikakis et al. 1996, Fox 1993). Moreover, their multi-modal action, relative low cost (Singh et al. 2012) and other beneficial effects like improved lipid profile in RA patients (Morris et al. 2011) make them favorable anti-inflammatory drugs in the hands of rheumatologists and dermatologists.

1.5- Epidemiological concerns

A brief look at the prevalence of the major rheumatic diseases indicates that a significant number of the population may benefit from the anti-inflammatory effects of CQ and HCQ. It also gives a rough approximation of the number of people who might develop retinal toxicity. According to the *National Arthritis Data Workgroup* the prevalence of RA in the United States was estimated to be about 1.3 million adults (Helmick et al. 2008). The study results showed that the prevalence of SLE and Sjögren's syndrome were 161,000 to 322,000 and 0.4 million to 3.1 million adults respectively. A recent survey in Canada states that the prevalence of “systemic autoimmune rheumatic disease (SARD)” is about 2 to 5 per 1000 residents in 7 provinces of this country (Brotten et al. 2014).

1.6- Future of CQ and HCQ use in rheumatic disease

It might seem logical that CQ and HCQ, which having been introduced in the early 1950s, should be replaced by newer drugs without their related significant side effects like permanent visual loss; however, present data does not support this idea. Although CQ and HCQ have a relatively low incidence of retinal toxicity associated with their long-term use, they remain very effective medications having a good systemic safety profile. Besides, the newly introduced therapeutic agents carry their own significant, even life-threatening side effects and some options are so expensive that their use is limited to exceptionally drug resistant cases. Some of these life-threatening side effects may require very frequent clinical and paraclinical monitoring, thereby increasing the burden to patients and the health care system (O'Dell et al. 2002).

Additionally, different groups of DMARDs have unique treatment effects, so the use of one type might not obviate the necessity to use others. The current trend in the treatment of rheumatologic disease is the prescription of a combination of different DMARDs (O'Dell et al. 2002). Considering the advantages of CQ and HCQ over other therapies, it seems that at least in the near future these drugs will stay in the list of recommended DMARDs with prevention of induced retinal toxicity remaining the major clinical concern.

1.7- CQ/HCQ ocular toxicity

The ocular structures involved in antimalarial toxicity include the cornea, crystalline lens, ciliary body and retina. In the cornea, lesions termed “cornea verticillata” may develop as whorl-like grey deposits that do not interfere with vision and are reversible upon discontinuation of the offending drug (Desso and Rungger-Brändle 2007). Likewise, lens opacities and ciliary body involvement do not have significant clinical implications; the only major concern is retinal toxicity as it may cause irreversible central vision loss (Yam 2006). Rosental and colleagues (Rosental et al. 1978) found that both shortly after administration, and 3-6 months after discontinuation of CQ treatment to rhesus monkeys, the greatest concentration of the drug and its by-products was detected in those tissues containing the highest amounts of melanin: iris, ciliary body and choroid, followed by sclera and retina. There was a negligible amount of drug found in the lens and cornea. CQ accumulation in melanin-containing eye structures was identified even at low dosage.

1.7.1- Cellular mechanisms and histopathology findings

CQ and HCQ have been known to be weak basic compounds which tend to accumulate in the acidic intracellular vesicles, most importantly lysosomes; this property enables them to increase the pH in the cells and lysosomes, thereby interfering with normal protein degradation and digestion by the malaria parasite (Schlesinger et al. 1988). The same action is also seen in human retinal pigment epithelium and retinal neurons (Mahon et al. 2004). This study indicated that abnormal protein degradation resulting from accumulation of CQ metabolites in the lysosomes is the basic mechanism of the toxicity. The process is initially more prominent in the retinal ganglion cells which are most abundant in the parafoveal area (Mahon et al. 2004, Rosental et al. 1978).

1.7.2- Natural course and clinical picture

CQ/HCQ accumulation in the retinal neurons interferes with normal cellular activities. The cellular stressful condition, if strong and prolonged enough may present as abnormalities in retinal function and finally destruction of the normal retinal architecture. In 1959, Hobbs and colleagues reported 3 cases in different stages of chloroquine toxicity, a combination of visual symptoms, field defects and funduscopic abnormalities involving only the macula or the entire retina depending on the advancement of the toxic effects (Hobbs et al. 1959). In less severe cases, lesions were primarily detected in the macular region, while one patient presented with widespread retinal destruction having signs and symptoms identical to "Retinitis Pigmentosa". With

the current screening programs aimed at earlier detection, these more extensive cases are less often encountered (Weiner et al. 1991).

Presenting symptoms might reflect dysfunctions in the ciliary body and/or retina: difficulty in reading, photophobia, blurred vision, central visual field scotomas and light flashes (Tzekov, Serrato et al. 2004).

Early fundus changes can typically be detected as granularity of the RPE and retinal pigmentary changes in the para-foveal region. More advanced toxicity typically presents as “bull’s eye maculopathy” (Tzekov, Serrato et al. 2004). It consists of an oval band of depigmentation surrounded by another band of pigmentation (Figure-1). Extension of the lesion into the peripheral retina may be associated with narrowing of the retinal vessels, optic disc pallor, granularity of the retina, and increased clinical visualization of the choroidal vasculature in the periphery (Tzekov, Serrato et al. 2005, Hobbs et al. 1959). Retinopathy may further be complicated by cystoid macular edema and epiretinal membrane formation (Kellner et al. 2014).

A definitive diagnosis of CQ/HCQ retinopathy is generally made when the classic bull’s eye maculopathy is associated with a history of long-term drug consumption. A bull’s eye pattern may present in some macular dystrophies, however, such lesions are not characteristic of age-related retinal degeneration or vasculopathies associated with underlying auto immune diseases (Wolfe and Marmor 2010). Diagnosis has also been defined based on the documented functional abnormalities like reproducible visual field defects (Easterbrook 1999). However, precise diagnostic modalities are capable of

detecting subtle changes which at times may be confounded by other pathology or result from technical recording error. Therefore, new diagnostic terms: “definite or probable” and “possible” (discussed later) have been suggested by the American Academy of Ophthalmology (AAO) (Marmor et al. AAO 2011).

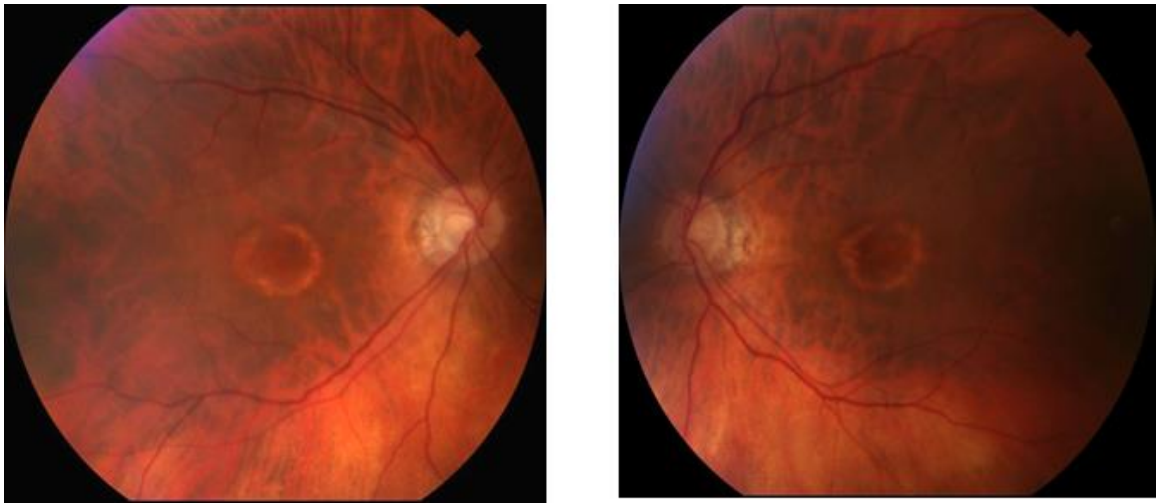


Figure 1 Color fundus photograph of both eyes showing bilateral “bull’s eye maculopathy” due to hydroxychloroquine toxicity

Retinal toxicity is generally irreversible once it progresses to the stage of bull's eye maculopathy. It also may continue to progress despite discontinuation of therapy, resulting in significant central and peripheral vision loss (Marmor et al. AAO 2002, Michaelides et al. 2011). Nevertheless, it has been proposed that the disease goes through an early stage, characterized only by subtle functional abnormalities that might be reversible upon cessation of therapy (Marmor et al. AAO 2011).

1.7.3- Incidence

There is currently a lack of consensus regarding the definitive diagnosis of CQ and HCQ retinal toxicity in its early stages, and therefore data regarding the incidence of this condition is controversial. Easterbrook reported the CQ retinal toxicity in 7.4% of patients receiving less than 3mg/kg/day, and this value rose to 40% with doses between 3-4 mg/kg/day (Easterbrook 1999). The risk of HCQ retinal toxicity is quite low, especially during the first 5 years of continuous consumption. It rises to more than 1% after 5-7 years and even higher following 15 to 20 years of exposure (Marmor et al. AAO 2011). All of the reports about the incidence of the disease have been affected by the definition of toxicity and the modes of diagnosis (Tzekov, Serrato et al. 2004). In a recent survey of more than 2000 patients who had received a cumulative dose of more than 1000 g HCQ, 150 cases (7.5%) had clear signs of toxicity (Marmor and Melles 2014).

1.7.4- Risk factors

Treatment of autoimmune disease using CQ/HCQ can continue for a relatively long period of a patient's life, sometimes 2 decades or more. Considering the long half-life of these drugs, a large amount of their metabolites may build up in ocular tissues over the period of treatment. Current literature indicates that a cumulative dose of HCQ in excess of 1000g (7 years of standard dose of 400mg/day) and CQ of more than 460g (5 years of standard dose of 250mg/day) significantly increases the rate of retinal toxicity (Marmor et al. AAO 2011). However, this complication has also been reported with lower doses (Michaelidis et al. 2011, Palma Sánchez et al. 2013). Retinal toxicity has been reported to start several years after cessation of CQ therapy (Ehrenfeld et al. 1986, Kazi et al. 2014).

Since these drugs are not distributed significantly in the fat tissue the safe limit of daily dose is defined based on the fat free, lean body weight (LBW) (Hume 1966) rather than total body weight. Daily dose of more than 3mg/kg of CQ and 6.5 mg/kg based on LBW is known as a risk factor (Marmor et al. AAO 2011). In fact, there is strong evidence suggesting that both daily and cumulative doses are important risk factors (Easterbrook 1999, Wolf and Marmor 2012, Marmor et al. AAO 2011).

Other possibly less important factors include age, underlying retinal disease and genetic predisposition (Marmor et al. AAO 2011). One group of investigators have recently reported statistically significant relationship between arterial hypertension and development of toxicity (Palma Sánchez et al. 2013).

1.7.5- Prognosis and clinical approach

In advanced toxic conditions visual field loss reaches an irreversible stage, with central and peripheral expansion of the paracentral visual sensitivity loss (*scotoma*, described later). Eventually vision of the involved eye will be irreversibly reduced. There is no known treatment for CQ/HCQ retinal toxicity. Discontinuation of the drug upon detection of the earliest toxic signs is potentially the sole method for preserving vision (Michaelides et al. 2011). Clinical experience indicates that diagnosis and cessation of the offending drug generally does not affect the final visual outcome when toxicity becomes evident on clinical exam or when the patient complains of symptoms (Marmor and Hu 2014). While screening for early diagnosis does not necessarily prevent retinal damage altogether, there is strong evidence that earlier detection of toxicity and discontinuation of the drug can significantly improve the visual prognosis (Marmor et al. AAO 2011, Nika et al. 2014). When the disease is diagnosed before RPE involvement foveal architecture is often saved (Marmor and Hu 2014). Therefore, early diagnosis and visual preservation can only be achieved when all patients at risk are closely monitored. The following section describes the past and currently recommended screening protocols, and the challenges faced by clinical practitioners.

1.8- CQ/HCQ retinal toxicity screening

A detailed list of the screening program requirements can be found in the World Health Organization (WHO) publication on this topic (Wilson and Jungner 1968). According to The Royal College of Ophthalmologists, the following requirements should be met to justify the HCQ ocular toxicity screening program (Fielder et al. 1998):

1. Presence of cause and effect relationship between the drug and the side effect.
2. Presence of a test or group of tests sensitive and specific enough to detect the toxicity when it still can be reversed.
3. Stopping the drug should be able to prevent irreversible damage.
4. Screening should not be associated with significant concerns for the patients and should not cause significant risks including financial burden.

The relationship between CQ/HCQ and their characteristic retinal toxic effects has already been described. The last 3 requirements have been the area of concern in the screening protocols. Over the past 60 years and especially prior to the year 2000, most of the diagnostic tests for macular function and structure were used for this purpose, but never offered any significant advantage over clinical exam alone in the early detection of disease (see below). The diagnostic approach has suffered from a wide range of opinions regarding the most sensitive test or combination of tests (Fielder et al. 1998). In 2002, the American Academy of Ophthalmology attempted to address this issue by arranging a task force of experts in this field. They reviewed the most updated knowledge on the available tests, the socioeconomic impact of screening, and the liabilities of the attending physicians. The authors agreed that considering all of the pertinent factors, screening was advisable, and they emphasized that early diagnosis could be helpful in preventing severe toxicity but would not assure that vision loss could be avoided altogether (Marmor et al. AAO, 2002).

The main concerns of the 2002 AAO document on CQ/HCQ retinal toxicity screening were the patients' vision, the economic burden of the screening program, and treating physician's liability. The following summarizes the 2002 AAO CQ/HCQ recommendations:

1. After obtaining a normal baseline clinical exam and ancillary diagnostic test(s), further testing is not necessary during the first 5 years if there is no apparent risk factor and the dose of the drugs is kept below 6.5 mg/kg of HCQ and 3 mg/kg of CQ.
2. Upon the first ophthalmic exam, recording the vision, overall eye condition, presence of retinal lesions that may simulate HCQ toxicity should be recorded and photographically documented. Fluorescein angiography may help in diagnosis by enhancing early parafoveal pigmentary changes.
3. A set of *psychophysical tests*, automated visual field testing, Amsler grid and color vision can be helpful in making early diagnosis.
4. Mass response obtained from the entire retina by full field electroretinography (ffERG) and electro-oculogram (EOG) is not a reliable test of early toxicity. Multifocal ERG could play a role in early detection of the toxicity but this remained to be established.

Despite the incorporation of these guidelines into clinical practice, the screening outcome was unsatisfactory and the number of patients suffering from vision loss due to CQ/HCQ retinal toxicity did not diminish significantly (Marmor et al. AAO, 2011). In addition, it was found that the incidence of toxicity was actually higher than previously expected (Marmor et al. AAO, 2002).

Between 2000, when the first article about mfERG findings in HCQ retinal toxicity was published (Kellner et al. 2000), and 2011, several studies were conducted to determine the reliability of mfERG in CQ/HCQ toxicity screening. Moreover, 2 more sophisticated testing modalities for evaluating retinal structure, fundus auto fluorescence (FAF) and spectral domain optical coherence tomography (sdOCT) were introduced to detect early signs of toxicity with promising results (Kellner et al. 2006, Kellner et al. 2009). These studies will be briefly reviewed along with a description of each test later. A review of the updated literature regarding these newest diagnostic testing modalities prompted revision of the AAO screening guidelines in 2011. The

highlights of the AAO revised recommendations for HCQ retinal toxicity screening (Marmor et al. AAO, 2011) included:

1. Maximum 5 year period between baseline test recording and starting regular, annual follow up is still recommended unless there are significant risk factors or clinical suspicion of toxicity.
2. The safe daily dose should be calculated based on ideal body weight rather than total body weight.
3. The following tests are no longer recommended for screening due to their low sensitivity: fluorescein angiography, time domain OCT, Amsler grid and color vision.
4. 10-2 automated visual field (AVF) should be included in baseline testing along with one of the following objective tests: sdOCT, FAF or mfERG, as available.
5. The sensitivity and specificity of the objective screening tests have not yet been determined.

In addition to the currently recommended diagnostic modalities, newer methods like microperimetry have been tested with promising results for early detection of this toxicity (Jivrajka et al. 2013). But, these are limited to few reports and have not affected the clinical care extensively. A study published 2 years later discovered that the recommendations have increased the cost of screening without detecting more toxic cases (Browning 2013).

The following section will briefly review the basics and historical background of the currently recommended diagnostic screening tests for CQ/HCQ retinal toxicity.

1.8.1- Automated visual field (AVF) testing

Visual field testing, or *perimetry*, is a subjective measure of retinal function which also depends on the integrity of the optic nerves, intracranial visual pathways and cortex, the subject's concentration, ability to fixate on the central target, level of intelligence and physical ability to respond to the stimuli. The retinal photoreceptors differ in sensitivity to light stimulation based on their cell type, topographical location

and also on the background light intensity. Generally speaking, in the light adapted state, the cone photoreceptors located in the fovea are the most sensitive and can detect the dimmest light on a bright background. This capability gradually diminishes towards the peripheral retina, along with a decrease in retinal contrast sensitivity (Humphrey field analyser manual, 2012). Conversely, in the dark adapted state, the central retina becomes less sensitive to light stimulation. The minimum light intensity which can be perceived 50% of the time at a given test point is called the *threshold light sensitivity* of that area (Humphrey field analyser manual, 2012). Based on this definition, all of the lights brighter than threshold can be detected while all lights less intense than threshold cannot.

The aim of visual field testing is to estimate the threshold light sensitivities of the entire field of the vision. As this is practically difficult and time consuming, a standard number of spots are tested. The threshold results are then compared to those of an age- and sex-matched normative database. A *scotoma* is said to be present when the threshold value of an area is lower than normal. A scotoma is *absolute* when the maximum light stimulus cannot be distinguished or *relative* when the light stimulus can be distinguished at a lesser sensitivity (brighter stimulus) than normal. The area occupied by the normal optic nerve head, where there are no photoreceptors, is a good example of an absolute scotoma.

Figure 2 illustrates a 10-2 single field analysis printout. Automated *static perimetry* uses brief, stationary flashes of light in predefined locations throughout the visual field. During the test, the background lighting is set to a standard intensity less

than that of the target stimuli to be presented. At each position in the visual field, flashes of light with increasing intensity are presented to the retina and the minimum light intensity perceived and reported by the subject, the “threshold”, is recorded. Traditionally, a higher threshold number represents the perception of a dimmer light stimulus. In light adapted state, the retinal light sensitivity is highest in the foveal center and gradually diminishes towards periphery. Threshold numbers obtained from testing each individual eye are compared to the normal database of the instrument to create deviation maps. The term 10-2 AVF refers to the test protocol evaluating the central 10 degrees surrounding central fixation. This protocol determines the visual threshold of a high concentration of test points in the area of highest concern in CQ/HCQ retinal toxicity. The area tested by this protocol does not include the optic nerve head *blind spot*. Generally, the software utilizes a series of light stimuli in a bracketing approach to determine thresholds. Faster strategies have been developed and widely used in order to enhance the subjects’ comfort and reduce testing time. The results are displayed both on a *Gray Scale* plot and in a *Threshold* numerical map. A *Total Deviation* map presents the difference between measured threshold for each retinal point tested and the age-matched normal.

The Pattern Deviation map attempts to compensate for overall depression (e.g. due to cataract) by taking the total deviation map and adjusting each point by an amount equal to the average of the 17 worse test points, which allows for better recognition of localized defects.

The statistical significance of the threshold deviations is demonstrated by a symbol representing its p-value. Mean deviation (MD) reveals the average of all points in the total deviation (Humphrey field analyser manual PDF 2012). Pattern standard deviation (PSD) is the quantitative representation to which the subject's field departs from that of the age-matched normal reference. High PSD indicates field irregularities, i.e. presence of localized scotomas with or without generalized depression of retinal sensitivity (Humphrey field analyser manual PDF 2012). The test reliability is indicated by three indices: Fixation Losses, False Positive Errors, and False Negative Errors. A fixation loss occurs when the subject responds to a stimulus presented in their blind spot. A false positive error occurs when the subject responds when a light was not even presented. A false negative error occurs when the subject fails to respond to a light much brighter than the previously established threshold at that point. Fixation loss rates exceeding 20% and false positive or negative rates exceeding 15% may indicate low test reliability (Humphrey field analyser manual PDF 2012).

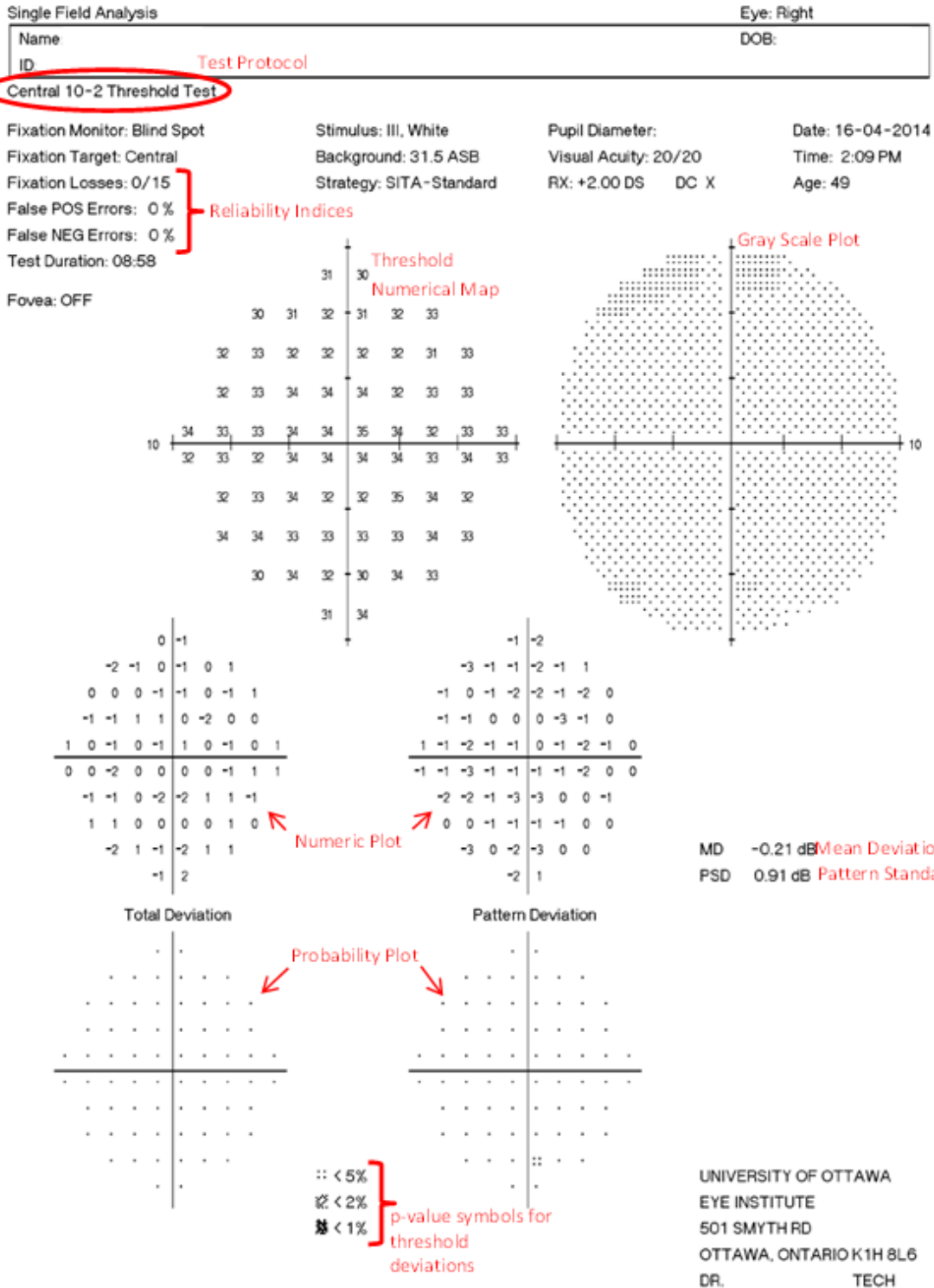


Figure 2: Humphrey Field Analyzer 10-2 Single Field Analysis Printout

The different plots and analysis parameters have been labeled in red.

1.8.2- Multifocal ERG

Electroretinography (ERG) testing is a reliable, objective means of evaluating retinal function, which records the electrical signals in the eye in response to a light stimulus. In contrast to visual fields, ERG records retinal activity independent of the visual pathways and cortex, but it has its own limitations. For instance, full field ERG records a mass response from the entire retina, which is not capable of detecting localized electrophysiological disturbances such as those associated with early HCQ retinopathy (Marmor et al. AAO, 2002).

In general, full field ERG is not affected until at least 20% of the retina becomes dysfunctional (Creel 2011). In response to the demand for recording localized electroretinal function, and along with technological advances, multifocal ERG (mfERG) was introduced by Eric Sutter in 1995. Multifocal ERG testing is performed when the functional defect is suspected to be localized to the central region of the retina, as in CQ/HCQ toxicity (Figure-3A). It uses an algorithm which calculates the electrophysiological responses from much smaller sectors of the retina less than a square millimetre (Sutter 2010). Like full field ERG, the electrical responses are recorded using electrodes on or around the eye, but the methods of light stimulation, recording, and analysis differ. The multifocal ERG stimulus is a pattern presented on an LCD monitor, which the patient views from a distance of about 33 centimetres. The pattern consists of a geometric arrangement of 61 or 103 hexagonal elements (arranged in 5 or 6 concentric rings), half of which are black and half white in an apparent random arrangement of colour assignment. The hexagonal tiles change colour continually from

black to white at a rate of 75 Hz in a predetermined but apparently random manner. The mfERG signals are derived by correlating the continuous ERG signal with the “on and off” responses from each element (Hood 2011).

The resulting waveform is a biphasic response called first-order kernel (Figure-3B). It consists of 3 consecutive deflections, a positive wave called P1 intervening 2 negative waves N1 and N2. The mfERG waveforms resemble that containing “a- and b-waves” of the full field ERG however, due to the differing pattern and duration of retinal stimulation, the recorded responses are different.

mfERG responses mainly represent cone and bipolar cell activity (Hood et al. 1997). mfERG analysis software assesses each component (practically N1 and P1) of the waves in terms of their amplitude and implicit timing (latency).

The results are displayed in both quantitative and qualitative formats in the print out. The *Trace Array* enables detecting focal abnormalities in each hexagon or group of hexagons (Figure 4A). The *Ring Average* analysis presents the average waveform of all hexagons in each concentric ring (Figure 4B). The peak of each wave is indicated by a dot which can be further adjusted by the operator to achieve higher accuracy. Horizontal and vertical lines of the rectangular boxes designate the amplitude and implicit time normal limits respectively. The P1 amplitude and implicit timing of each ring average are quantitatively presented in a table. The central hexagon is called ring 1. Rings 2 and 3 roughly correspond to the para foveal area classically involved in CQ/HCQ retinal toxicity. Normal deviation map (Figure 4C) is a topographical illustration of the data difference from the normal database.

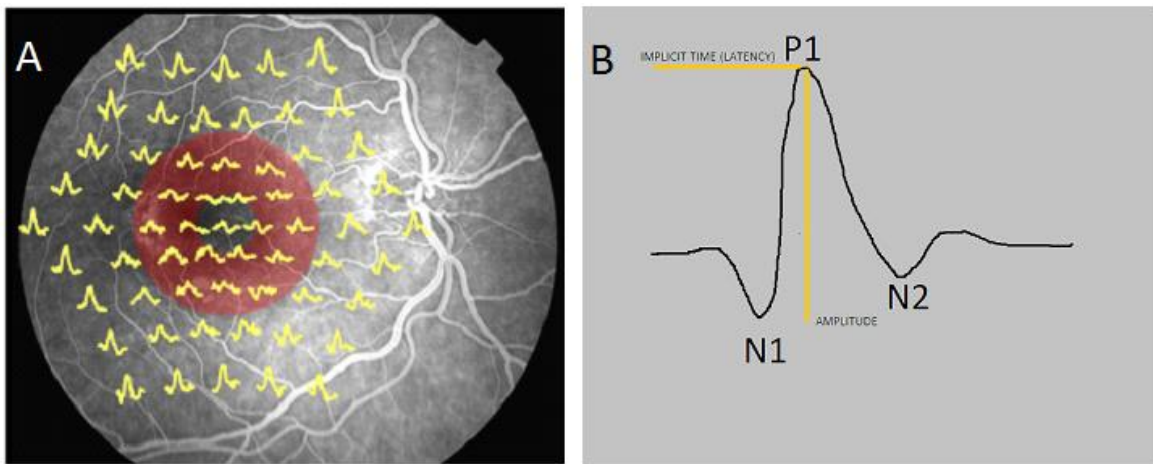


Figure 3: Multifocal ERG waveform

A: topographic map of retinal electrical activity as recorded by mfERG. The area covered by the red ring roughly depicts the parafoveal region of the macula and rings 2 and 3 of the mfERG trace array.

B: First order kernel, N1, P1 and N2 deflections. Two principal quantitative measures include: P1 implicit time (horizontal line - the duration between the light stimulation and P1 peak) and P1 amplitude (vertical line – voltage difference between the trough of N1 and peak of P1).

Some studies have reported improved mfERG test results following discontinuation of therapy (Lai et al. 2005, Maturi et al. 2004), which suggests that mfERG might be sensitive enough to detect disease in a reversible stage; however, they were limited to case reports and their statistical significance could not be determined.

“Relevance of mfERG findings to true toxicity” has been a matter of debate. Studies on mfERG have detected abnormalities in a relatively large number of patients, which may be as high as 60% after several years of drug use (Lai et al. 2005, Maturi et al. 2004, Moschos et al. 2004). Since this number of patients never develop clinical evidence of toxicity even after 2 decades of treatment with HCQ, Marmor argued that mfERG abnormalities may in fact, reflect generalized electrophysiological alterations of the retinal neurons which are not relevant to true CQ/HCQ toxicity (Marmor 2005). This idea was later challenged by the fact that neither the findings of full field ERG(ffERG), which detects global electrical activity of the retina, nor electro oculogram (EOG), which represents functional integrity of the RPE layer, detected consistent abnormalities in early stages of toxicity (Nebbioso et al. 2009, Marmor et al. AAO 2011).

In the early years of mfERG clinical use, this test suffered from inconsistency of the results and significant inter-individual variability (Marmor 2005). Age, intraocular opacities and electrical noise were among the most prominent factors affecting the mfERG. Along with improvements in instrumentation, the protocols for data acquisition and processing were standardized by the International Society for Clinical

Electrophysiology of Vision (ISCEV) (Marmor et al. ISCEV 2003, Hood 2011). Ring average calculation was an effective innovation to mitigate variation due to age (Tzekov, Serrato et al. 2004). The introduction of ring ratio analysis significantly furthered this approach. Lyons and Severns found that normalization of the amplitude ring averages by calculating R1 ring ratios significantly reduces the inter-individual variations and thereby increase the confidence in diagnosis of CQ/HCQ retinal toxicity (Lyons and Severns 2007). Later it was found that the sensitivity of the measures could be enhanced by normalization to R5 (R1/R5, R2/R5...) (Adam et al. 2012).

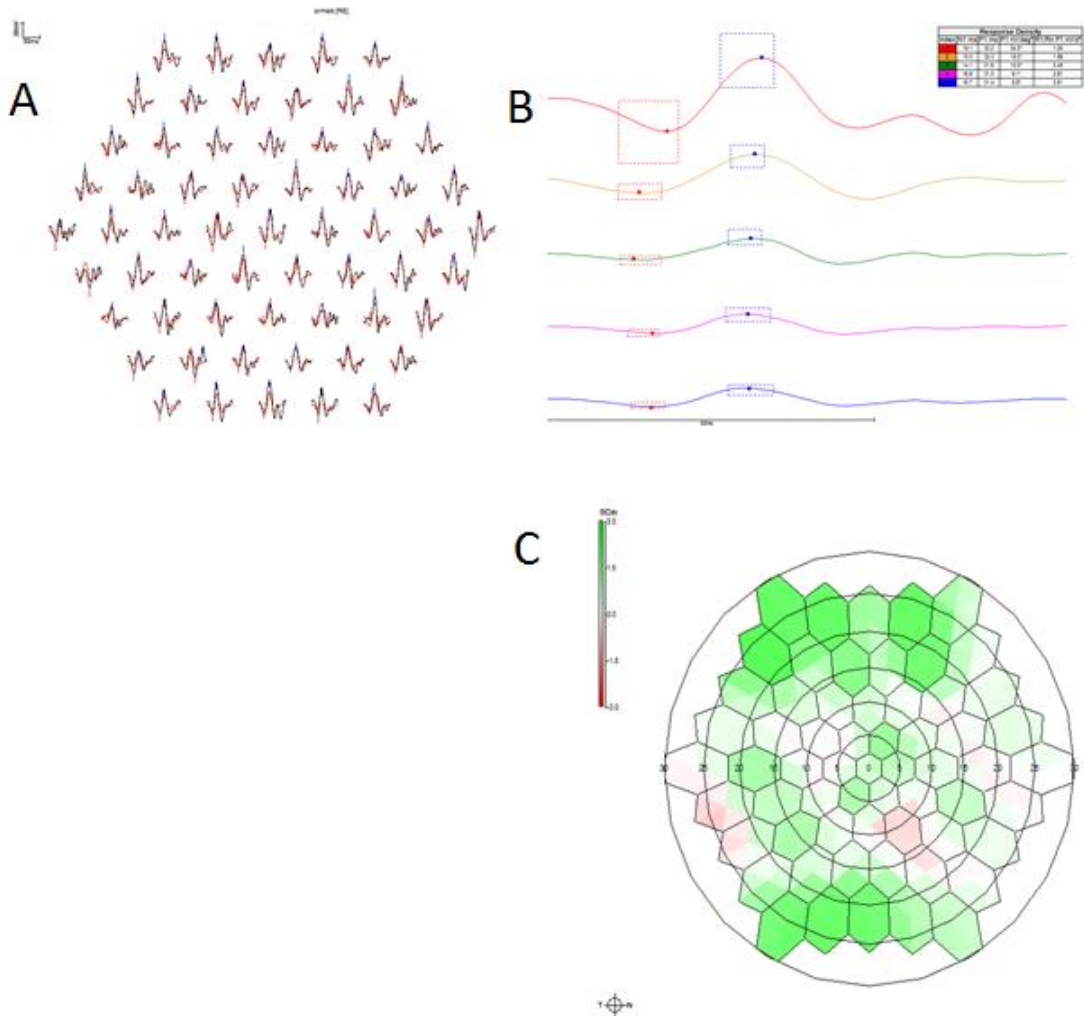


Figure 4: Multifocal ERG test result of a normal eye.

- A: Trace Array
- B: Ring Averages, the rectangles illustrate the normal amplitude and implicit time ranges (mean±2SD), quantitative measures are displayed in the table.
- C: Normal Deviation map, green and red colors indicate standard deviation measurements more and less than the average normal values respectively.

Multifocal ERG changes in HCQ retinal toxicity involve but are not exclusive to the reduction of amplitudes; implicit time of the waves may also become abnormally prolonged, but this finding is less common at least in earlier stages of disease (Maturi et al. 2004).

1.8.3-Fundus auto fluorescence (FAF)

FAF is one of the objective diagnostic tests recommended by the AAO for HCQ retinal toxicity screening. This test modality is primarily a means of detecting RPE abnormalities (Holz et al. 2001).

Fluorescence is the capability of some materials (fluorophores) to absorb an electromagnetic radiation and emit it with a longer wavelength (Sauer et al. 2011). The imaging device exposes the retina to a monochromatic light of a wavelength that excites fundus fluorophores. All of the back reflected rays are then filtered out except those emitted from the fundus fluorophores. Intravenous fluorescein and indocyanine green angiography (IFA & ICGA) are examples of clinical applications of this phenomenon. IFA and ICGA use injected dyes to highlight retinal and choroidal vasculature. Auto fluorescence, on the other hand, is a fluorescing property of the ocular fundus structures which is not dependent on the presence of any injected material (Figure 5A). Lipofuscin (LF) and its related fluorophores are the most important sources of ocular fundus auto

fluorescence (Kellner et al. 2006). LF, which accumulates in the RPE, is the major waste material produced by the outer segments of photoreceptors.

Hyper-auto-fluorescence takes place when photoreceptor cells shed abnormally high amounts of outer segment membranes or when the degradation capacity of the RPE cells is compromised (Kennedy et al. 1995). Hypo-auto-fluorescence may result from regional destruction of the RPE cells and absence of LF related fluorophores (Holz et al. 2001). Hypo-auto-fluorescence may also occur due to the blockage of light by retinal pigments or opacities in the light pathway. Hypo auto fluorescence of the fovea and parafovea is due to the absorption of the emitted blue light by the macular pigments, lutein and zeaxanthin (Theodore et al. 2005). In any case, it should be noted that the ability of detecting pathological FAF findings is highly dependent on the quality of the images (Schmitz-Valckenberg 2008).

Among the currently recommended tests, FAF was the third to be introduced in 2006 as an effective mode for early diagnosis of CQ/HCQ retinal toxicity (Kellner et al 2006). The authors described the earliest abnormality as a “fine pericentral ring of increased FAF”. FAF appeared sensitive enough to detect toxicity in one patient with a normal clinical exam. It was a promising result because the lesion did not progress after discontinuation of the medication. In more advanced cases, ring shaped areas of mottling, mixed hyper- and hypo-fluorescence, or total loss of fluorescence enclosed by hyper-fluorescence were observed which indicated loss of RPE cells. The authors used mfERG as their reference test which proved to be more sensitive than FAF. Although one

study reported all 4 recommended tests as equally sensitive (Kellner et al. 2009), other studies described FAF to be less sensitive than sdOCT and mfERG in the early diagnosis of CQ/HCQ retinal toxicity (Marmor and Hu 2014, Marmor 2012, Kellner et al. 2006). Unlike mfERG and sdOCT, there is no inherent quantitative analysis for FAF, which might reduce its objectivity.

1.8.4- Spectral Domain Optical Coherence Tomography (sdOCT)

Optical Coherence Tomography (OCT) is an objective, non-invasive imaging technique for the evaluation of retinal structure (Figures 6&7). Spectral Domain Optical Coherence Tomography (sdOCT) provides faster data acquisition, sensitivity and resolution compared to its preceding generation, time-domain OCT. Currently, sdOCT machines are used as the mainstay for diagnosis of retinal disease in most modern ophthalmology clinics. sdOCT provides precise retinal imaging with a resolution of 5 μm , equal to histologic sections (Gabriele et al. 2011).

Of the currently recommended tests, sdOCT was the fourth and latest to be introduced for screening of CQ/HCQ retinal toxicity. Its first reports were published in 2009 (Kellner et al. 2009, Stepien et al. 2009). Kellner and colleagues reported an interruption of the photoreceptor inner segment/outer segment (IS/OS) junction in cases of mild toxicity (Kellner et al. 2009). They also described reduced outer nuclear layer thickness in all cases irrespective of disease severity. These results were very similar to those found by an earlier study (Rodriguez-Padilla et al. 2007). Chen and colleagues described a “flying saucer” sign (Figure 8B), created by a parafoveal IS/OS line disruption

with normal outer retina in the fovea, which could indicate toxicity in suspected cases (Chen et al. 2010).

Marmor reported an even earlier finding than IS/OS line disruption: disappearance of the line between RPE and photoreceptor outer segments, the *cone outer segment tip line* (Figure 6), but this line was not normally seen in the sdOCT images of all eyes (Marmor 2012). Parafoveal RPE loss was also detected in eyes with advanced toxicity (Figure 8C).

Considering its sensitivity, ease of operation, availability, and familiarity of clinicians with the interpretation of test results, sdOCT is recommended as a reliable screening tool for HCQ retinal toxicity (Marmor et al. AAO 2011).

1.9- Design and purpose of this study

A research protocol was initiated at the University of Ottawa Eye Institute in 2011 with a main purpose of calculating the sensitivity and specificity of mfERG for detection of CQ/HCQ retinal toxicity. To achieve this goal, patients taking CQ/HCQ and undergoing routine eye evaluation were enrolled. In addition to a comprehensive clinical exam, the subjects of the study were tested using the four recommended screening modalities. The hypotheses to be tested are that mfERG is more sensitive than the other recommended tests and that mfERG findings are related to histopathologic representations of CQ/HCQ retinal toxicity.

In this cross-sectional study on the collected data in the last 3 years the preliminary data were analysed. After reviewing the most updated publications, a model was developed for standardization of the data analysis using the recommended

screening protocols. The possible relationship between the clinically recorded test results and the current knowledge of cellular basis of toxicity was discussed.

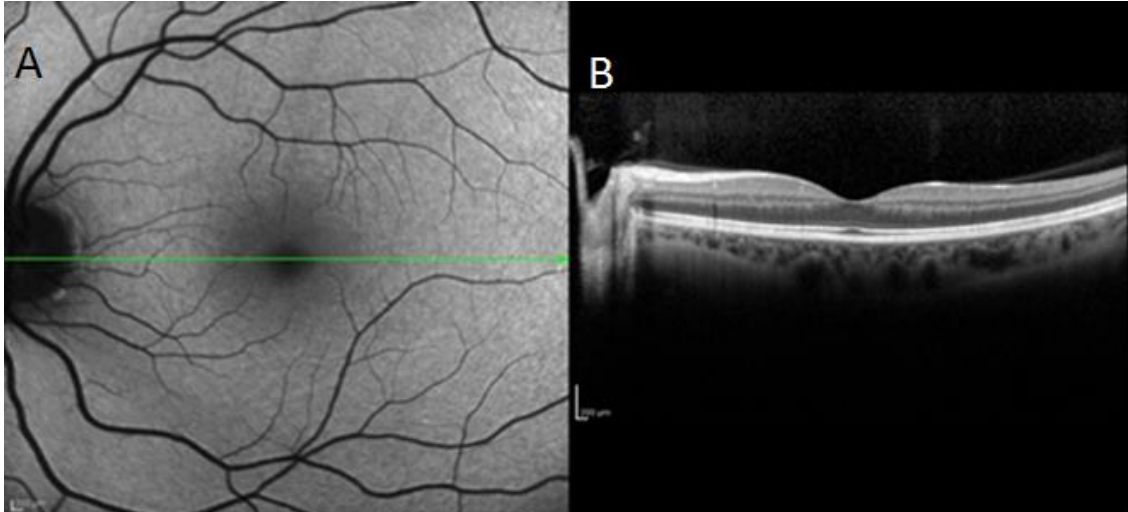


Figure 5: FAF and corresponding sdOCT images

- A: Normal fundus auto fluorescence image of the left eye. The areas with no lipofuscin - disc and vessels - produce no auto fluorescence.
- B: Foveal cross section of the sdOCT image obtained at the same time (discussed below). The green line in the FAF image corresponds to the topographical position of the sdOCT image.

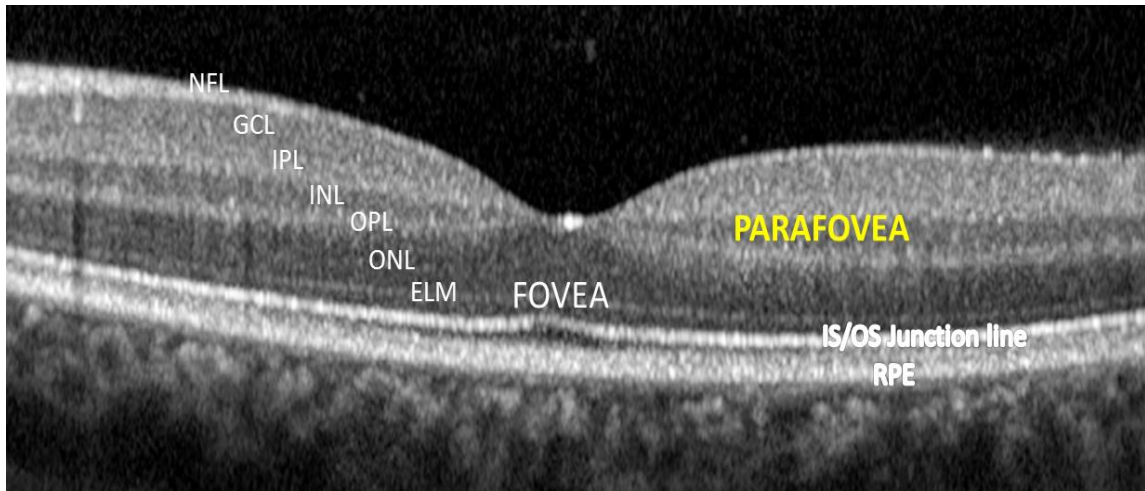


Figure 6: Normal sdOCT foveal cross section

Different retinal layers, fovea and parafoveal area are demonstrated. High resolution images obtained using this modality are almost as precise as low magnification histological sections. The retinal layers include (from inner to outer): NFL (nerve fiber layer), GCL (ganglion cell layer), IPL (inner plexiform layer), INL (inner nuclear layer), OPL (outer plexiform layer), ONL (outer nuclear layer), ELM (external limiting membrane), IS/OS junction (the line representing photoreceptors' inner segment/outer segment junction), RPE (retinal pigment epithelium). The parafoveal area has a very rich retinal capillary network while the central 400um of the fovea (symbolized by red dots), known as the Foveal Avascular Zone (FAZ), is totally devoid of retinal capillaries.

OS

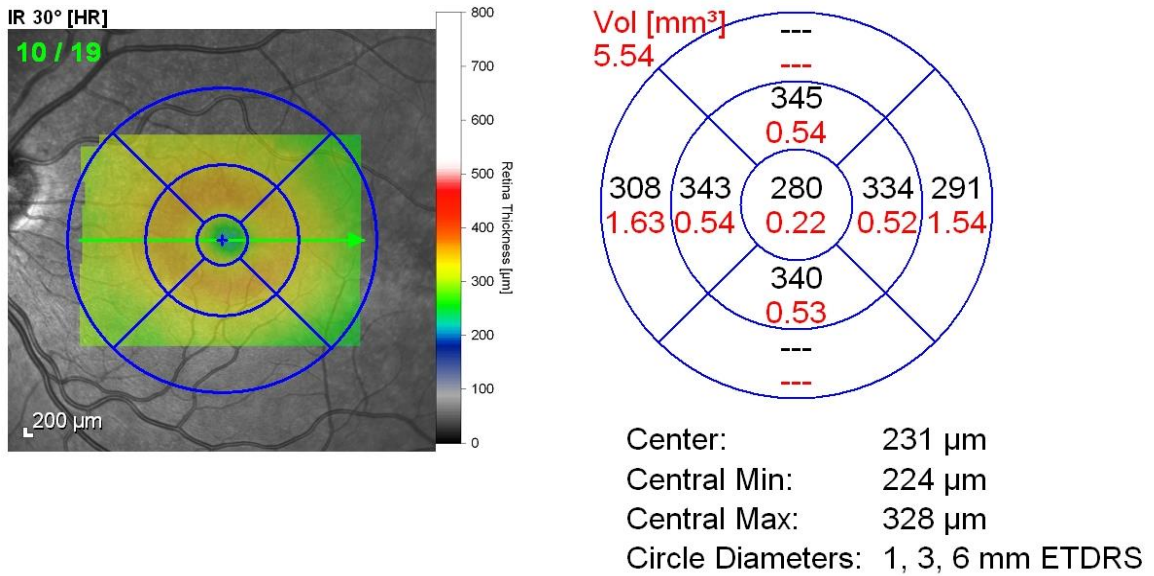


Figure 7: sdOCT retinal thickness map.

In addition to cross sectional imaging, sdOCT is capable of topographically mapping macular thickness (left). Quantitative data of retinal thickness is also presented (right). The central circle and surrounding middle and outer rings represent the fovea, parafovea, and perifovea respectively. Considering perifoveal thickness reduction, its measurement has been suggested as means of detecting early toxicity (Kahn et al. 2011).

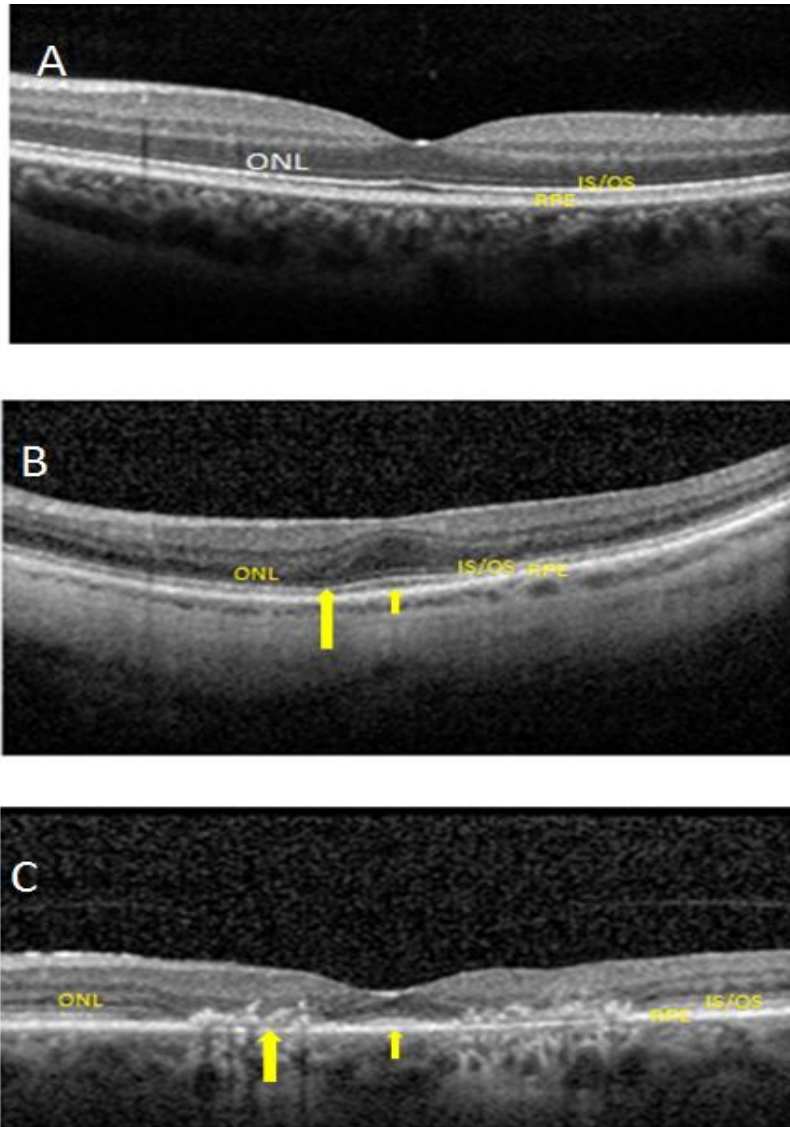


Figure 8: sdOCT image of normal retina and 2 different stages of toxicity

- A: sdOCT image illustrating the fovea of a normal eye.
- B: early parafoveal damage with localized loss of ONL and IS/OS junction line (large arrow). RPE is not damaged in this stage. Subfoveal IS/OS line is preserved (small arrow), creating “flying saucer sign”.
- C: Severe disruption of the parafoveal retina and RPE (large arrows), foveal photoreceptors and IS/OS junction line (small arrow).

2-Materials and Methods

This is a cross-sectional, observational study performed at The University of Ottawa Eye Institute between January 2011 and September 2014. The study was approved by The Ottawa Health Science Network Research Ethics Board (OHSN-REB). 94 patients referred to Dr. Chloe Gottlieb (CG) (Ophthalmologist, retina specialist, MD, F.R.C.S.C) for detection of CQ/HCQ retinal toxicity were recruited. Patients' sex, age, height, weight, systemic disease, medication used (CQ vs. HCQ), daily dose and duration of therapy, presence of hepatic and renal failure as well as the best corrected visual acuity (BCVA) were recorded. Body mass index (BMI) was calculated using the formula:

$$\text{BMI} = \text{mass (Kg)} / (\text{height (m)})^2$$

Obesity was defined as BMI \geq 30 (WHO, BMI classification 2006).

Lean body weight (LBW) was calculated using the following formula (Hume 1966):

$$\text{For men : LBM} = (0.32810 * W) + (0.33929 * H) - 29.5336$$

$$\text{For women : LBM} = (0.29569 * W) + (0.41813 * H) - 43.2933$$

Duration of CQ or HCQ treatment was recorded based on the total years and months of drug use. For calculation of the cumulative dose this value was converted into estimated days of therapy. Complete anterior segment exam and dilated fundus exam using binocular indirect ophthalmoscopy and 78 diopter slit lamp funduscopy were performed by Dr. CG. In addition, a color fundus photograph of each eye was acquired using a Topcon TRC-50DX Fundus Camera (Topcon Positioning Systems, Inc. USA) for documentation. To blind the clinician to the test results, the clinical exam was done first and reported, and then the diagnostic workup was ordered and performed. All of the tests were done on the same day or within 1 month of each other. The chart was considered complete when the clinical exam report and the results from all 4 tests (10-2 AVF, mfERG, FAF and sdOCT) were available. Patients with incomplete charts or bilateral poor quality para-clinical test results such as myopia or hyperopia in excess of 8 diopters, coexisting retinal disease precluding appropriate evaluation of the retina for CQ/HCQ retinal toxicity, history of retinal surgery were excluded. Amblyopic eyes were also excluded due to the correlation between visual acuity loss and reduced light sensitivity particularly at the fovea (Donahue et al. 1999). When these criteria applied to only one eye of a patient, only that eye was excluded. The following tests were performed for all patients included in the study, and data was entered into a de-identified Microsoft Excel spreadsheet for analysis.

2.1- Humphrey Field Analyzer 10-2 automated perimetry

Automated 10-2 threshold static single field analysis using Humphrey Field Analyzer II (HFA II), (Carl Zeiss Meditec Inc. Germany) was performed for all patients of the study. SITA-Standard protocol, 31.5 apostilb (asb) background light illumination and white stimulus size III were used for the test. The 10 degree area around the central fixation was tested using 10-2 static strategy. Visual threshold measures in 68 points of approximately 2 degrees apart were determined. The test result was labeled either reliable or unreliable based on the measured reliability indices including false positive and false negative responses, and fixation losses. The eyes with unreliable 10-2 AVF test results were either tested again or excluded from the study.

Based on the latest AAO guidelines (Marmor et al. AAO 2011), the pattern deviation plots were scored as:

- “Probable toxicity” when there were more than 3 abnormal points with $p < 2\%$ scattered in the mid periphery (representing the visual function of the para foveal region) or central parts of the pattern deviation plot in. (Figures 9 D,E,F & Figures 10 & 11)
- “Possible toxicity” when there were up to 3 non-adjacent, abnormally recorded points with $p < 2\%$ scattered randomly in the pattern deviation plot. (Figure 9 A,B,C)
- A dense scotoma in the far periphery of the 10-2 AVF that was not characteristic of CQ/HCQ retinal toxicity was disregarded (Figure 9G).

Despite the AAO recommendation defining “probable” based on the presence of bilateral abnormalities, for the purpose of including all eyes in statistical

analyses, unilateral classical appearance of abnormalities was considered “probable toxicity” for each eye.

For this study, “possible toxicity” results required sdOCT confirmation, but “probable toxicity” cases were labeled CQ/HCQ retinal toxicity whether or not there was agreement with sdOCT.

In addition to this pattern deviation plot grading, the quantitative measures of visual field [the mean deviation (MD) and pattern standard deviation (PSD)] were analyzed for correlation to cumulative dose of CQ or HCQ.

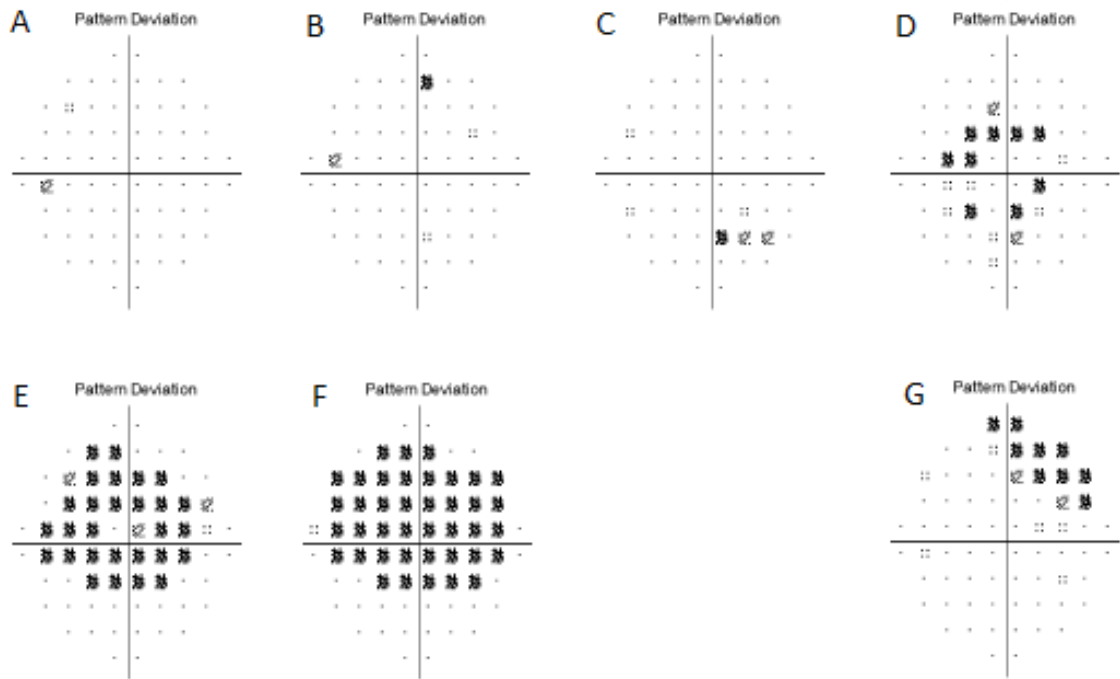


Figure 9: Humphrey 10-2 AVF Grading of the Pattern Deviation Plot

A, B, C are examples of “possible toxicity”. D, E, F are representations of “probable toxicity”. G is a peripheral quadrantic scotoma not characteristic of toxicity.

The earliest sign of toxicity on visual field may be the appearance of new abnormal spot(s) which may seem trivial and non-specific. Disregarding these findings may cause an untoward delay in diagnosis (Anderson et al. 2011). The 10-2 AVF progression map (Figure-12) is a useful tool for detecting the appearance of new abnormal points in the visual field over the course of patient follow up, which provides stronger evidence for significant abnormalities compared to the single field analysis.

Single Field Analysis

Eye: Left

Name:	DOB:
ID:	

Central 10-2 Threshold Test

Fixation Monitor: Blind Spot
 Fixation Target: Central
 Fixation Losses: 0/15
 False POS Errors: 4 %
 False NEG Errors: 0 %
 Test Duration: 05:15

Stimulus: III, White
 Background: 31.5 ASB
 Strategy: SITA-Standard

Pupil Diameter:
 Visual Acuity: 20/25
 RX: +4.50 DS DC X

Date: 07-05-2014
 Time: 1:44 PM
 Age: 79

Fovea: OFF

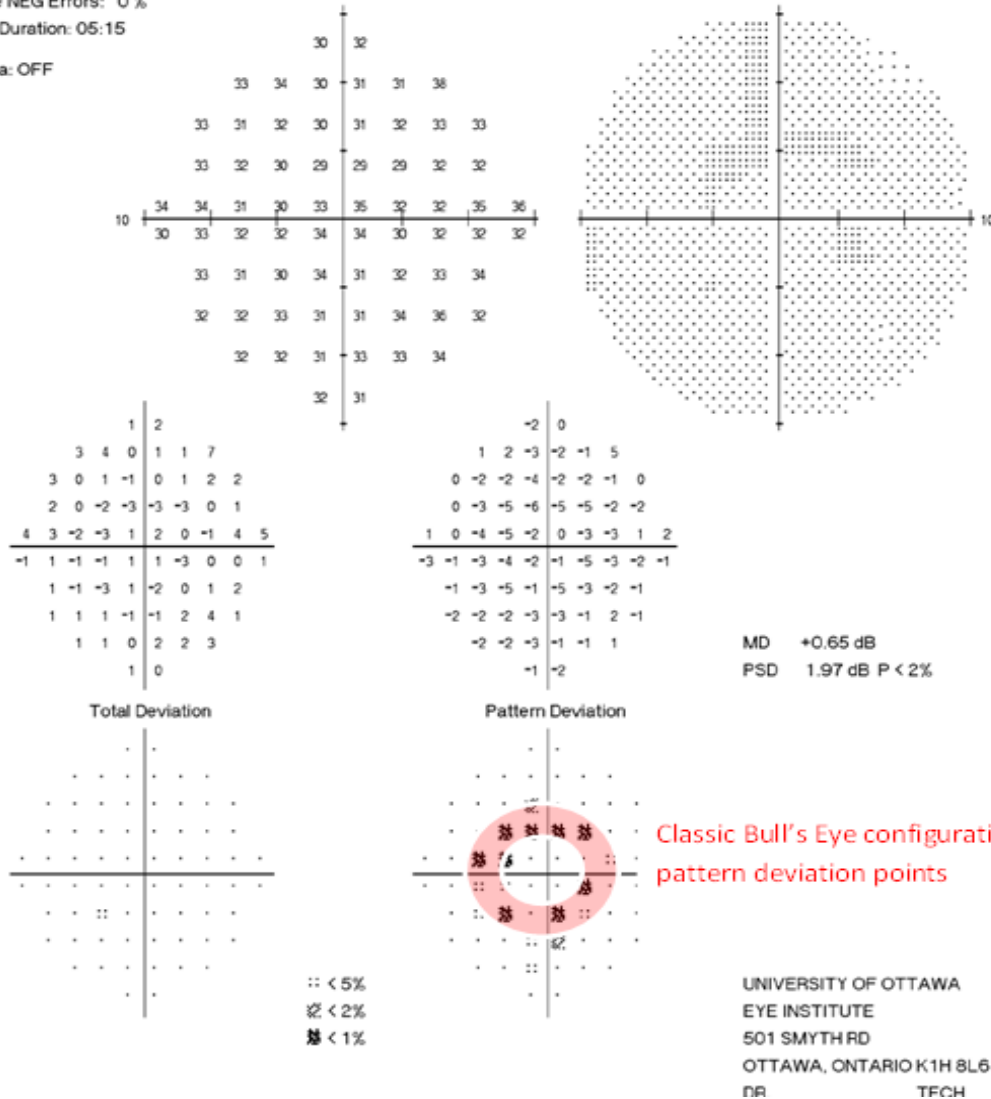


Figure 10: Single field analysis of the left eye representing a characteristic early “bull’s eye” scotoma

It consists of abnormal spots with p-value of <2% in the mid-periphery of the 10-2 visual field.

Single Field Analysis

Eye: Left

Name:	DOB:
ID:	

Central 10-2 Threshold Test

Fixation Monitor: Gaze/Blind Spot

Stimulus: III, White

Pupil Diameter:

Date: 01-10-2013

Fixation Target: Central

Background: 31.5 ASB

Visual Acuity:

Time: 8:15 AM

Fixation Losses: 1/23

Strategy: SITA-Standard

RX: +3.25 DS +1.25 DC X 169

Age: 72

False POS Errors: 2 %

False NEG Errors: 20 %

Test Duration: 09:07

Fovea: OFF

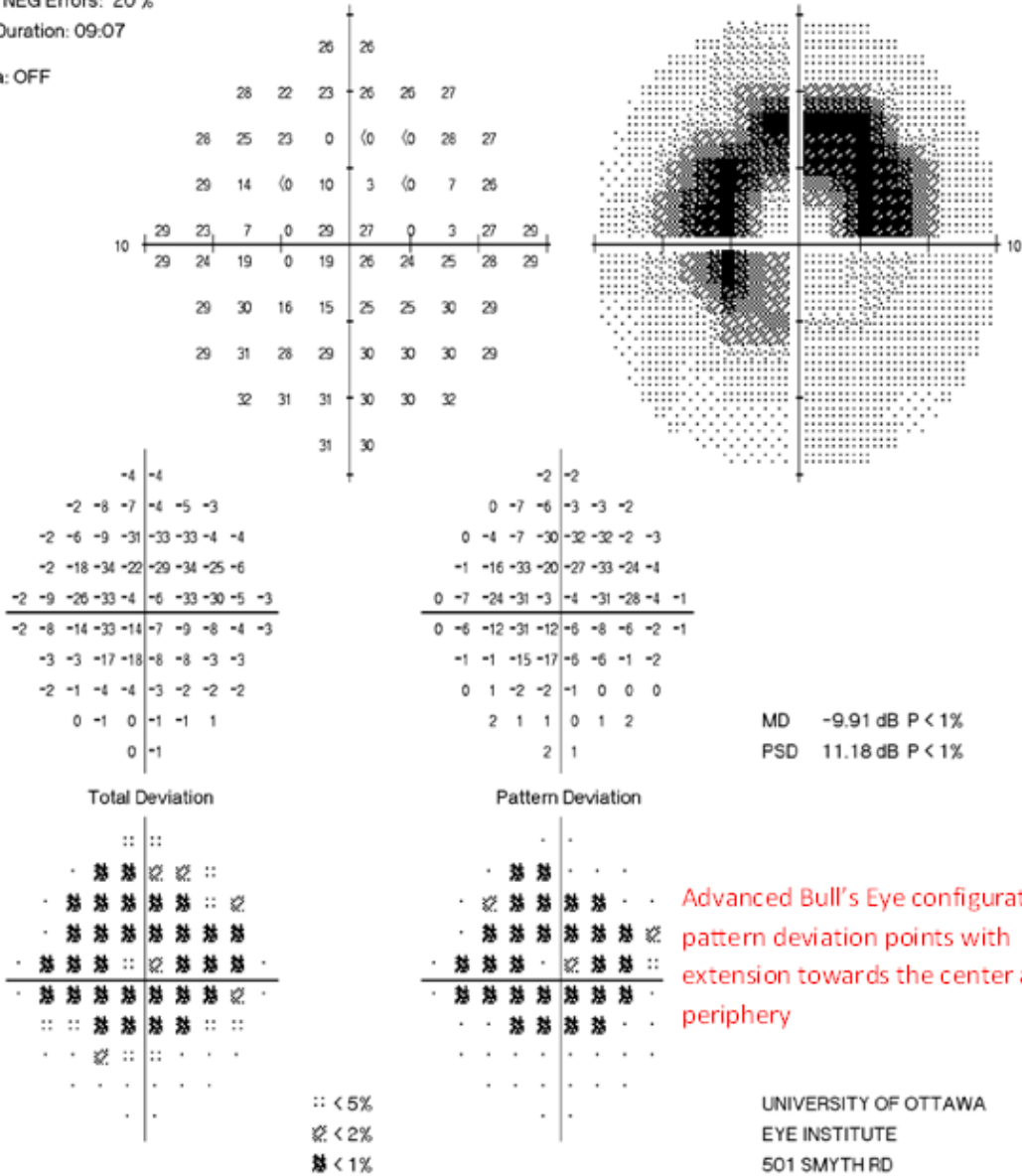


Figure 11: Humphrey 10-2 single field analysis printout illustrating advanced Hydroxychloroquine retinal toxicity

2.2- Multifocal ERG

Multifocal ERGs were recorded using an Espion Profile Multifocal System (Diagnosys, LLC, USA) according to ISCEV recommended standards (Hood 2011). The patients' eyes were anaesthetized using topical 0.5% proparacaine hydrochloride (Alcaine), and the pupils were dilated using 2.5% phenylephrine (Mydrin) and 1% tropicamide (Mydracyl). A stimulus containing 61 hexagonal elements was projected on the central 30 degrees surrounding the fovea in light adapted subjects' eyes using a LCD monitor having luminance of 1000 cd/m². Micro conductive DTL thread electrodes (Diagnosys, LLC, USA) were draped on the conjunctiva at the inferior limbus. ERG signals were extracted using the fast m-transform algorithm (m=14) in eight 30 second epochs. The test results were reviewed and reported by Dr Stuart G. Coupland (SGC), Director of Structural and Functional Retinal Imaging (Univeristy of Ottawa Eye institute). The reviewer was blind to the clinical exam and 10-2 AVF findings. The trace arrays, ring averages and response density topographic maps were evaluated, and test reliability indices were taken into account (Figures 4). Eyes with unreliable mfERG recordings were either retested or excluded from the study. Trace arrays represented 61 waveforms in 5 concentric rings (R1-R5). The software auto marked the N1 and P1 peaks, providing their implicit time and amplitude, but the technician manually adjusted the markers' positions as necessary to assure highest accuracy. The individual waveforms comprising the trace arrays were assessed for abnormally reduced amplitude or prolonged implicit times, signs of possible early toxicity. The software generated a set of 5 waveforms representing the average of the traces in the five concentric rings (Figures 4). Age-

matched normative data was available for this ring average analysis. Boxes representing 2 standard deviations around the mean were used to detect abnormally prolonged implicit time or reduced amplitudes of the P1 waves. The results were reported as either “normal”, “borderline” or “abnormal” based on the reviewer’s overall conclusion. The ring average measures, P1 implicit time in milliseconds (ms) and P1 response density amplitude in nanovolts/degree² (nV/deg²), were used in the quantitative analyses, including correlation with the cumulative dose of the drug.

2.3- Spectral domain OCT and FAF

Retinal spectral domain OCT and FAF images were obtained using a Spectralis HRA+OCT (Heidelberg Engineering GmbH, Germany) instrument through dilated pupils. Central fovea cross-sectional images were qualitatively reviewed for sdOCT abnormalities characteristic of toxicity. Disruption of para foveal IS/OS junction line, reduction of outer nuclear layer thickness creating the “flying saucer” sign, and RPE destruction in the foveal and para foveal regions (Figure 8) were considered typical sdOCT findings of toxicity (In the absence of co-existing retinal pathologies, eyes showing these abnormalities were labeled as “probable toxicity” as defined in Section 2.1).

The sdOCT scan also generated thickness maps demonstrating the average retinal thickness in 3 concentric rings with 1, 3, and 6 millimeter diameters centered on the fovea as used in the Early Treatment Diabetic Retinopathy Study (ETDRS) (Figure 7). Correlation between the recorded, average retinal thickness of each ring and the cumulative dose of CQ or HCQ was tested.

FAF images were obtained using the scanning laser ophthalmoscope by the same machine at the same session. The HRA Spectralis used an excitation wavelength of 488 nm and a barrier filter of >500nm. The images subtended 30 degrees on the posterior pole (roughly 15 degrees to each side of the foveal center). The imaging was performed in high-speed mode at a resolution of 768x768 pixels. Images were qualitatively reviewed for the presence of hyper- or hypo-auto-fluorescence (Figure 13), but the results were not used in the calculation of the sensitivity and specificity.

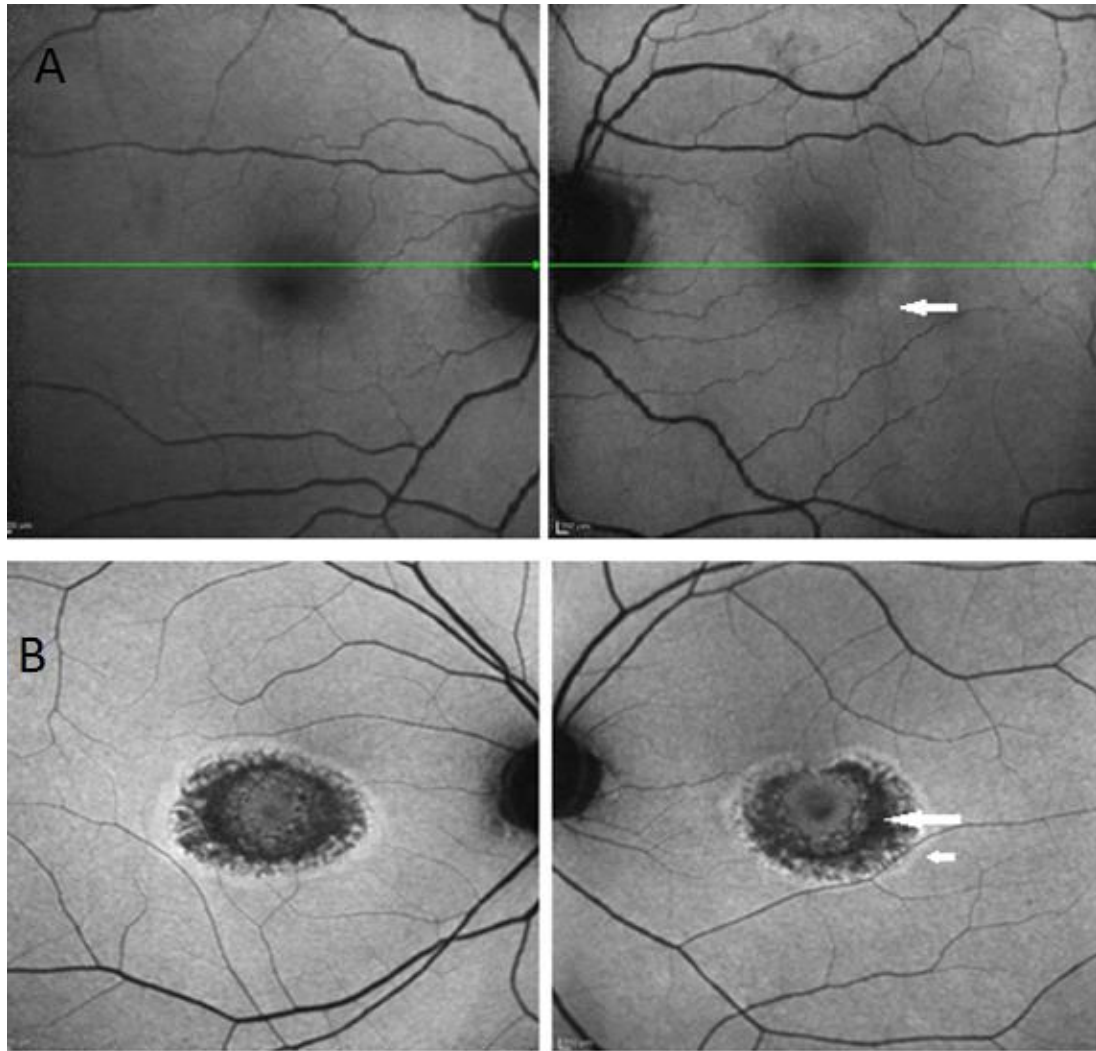


Figure 13: Early and late HCQ retinal toxicity detected by FAF.

- A: FAF image of both eyes of a patient with early HCQ retinopathy, hyper-auto-fluorescence is evident in the parafoveal area (arrow).
- B: advanced HCQ retinopathy with hypo-auto-fluorescence as a result of photoreceptor and RPE loss (large arrow) surrounded by rings of mottling and hypo-auto-fluorescence (small arrow). Foveal center is relatively preserved.

2.4- Statistical analysis

2.4.1- Relationship between target and reference test

The sensitivity and specificity of a new diagnostic test (target test) is generally calculated relative to a reference test. The reference is a test or combination of tests which is known to be the most reliable in identifying those with and without disease. The validity of the target test is calculated based on 4 possibilities in a 2 X 2 table (Figure 14).

		Reference test(s)	
		Positive	Negative
Target test	Positive	True Positive (TP)	False Positive (FP)
	Negative	False negative (FN)	True Negative (TN)

Figure 14: 2X2 table used to calculate the validity of a target test relative to the reference test

2.4.2- Sensitivity and specificity

These 2 terms are the main basis for the validity or accuracy of a test (Parikh et al. 2008):

Sensitivity (positive in disease) is the ability to correctly identify a subject with disease. Using the 2 X 2 table,

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

Specificity (negative in disease) is the ability to correctly identify a subject without disease. It is calculated as:

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

The **confidence interval (CI)** of sensitivity and specificity is 95% probability that the calculated results represent the true values of these measures in that population (Cox and Hinkley 1974).

In this study the sensitivity and specificity of mfERG as the target test was calculated using a 2X2 table. The reference test was 10-2 AVF and/or sdOCT. A positive reference test in this table indicates typically abnormal (“probable”) 10-2 AVF and/or sdOCT findings. The validity of “possible” 10-2 AVF findings was investigated based on their agreement with sdOCT.

The cases labeled as false positive mfERG results could possibly fall into two categories: First, those which were caused by technical or interpretation error. Second, those abnormal mfERG results considered normal by the reference test, reflecting a higher sensitivity of mfERG. In order to make this distinction, the eyes of patients receiving CQ and HCQ were classified into 2 groups according to mfERG results: **1- “False-Positive” Toxic** and **2- “Normal” Non-Toxic**. ANOVA with Bonferroni/Dunn post-hoc analysis was performed to determine whether a significant statistical difference exists between false-positive toxic, non-toxic normal, and control mfERG results.

mfERG P1 R5 ring ratios and implicit times recorded from the eyes in these subgroups were compared. A **control** group consisting of normal subjects who had never received CQ or HCQ was also included.

The calculation of sensitivity and specificity, and 95% confidence interval (CI) of these parameters was done using an online clinical calculator, <http://vassarstats.net/clin1.html>. ANOVA and correlation analyses were performed using STATVIEW (Abacus Concepts Inc. Berkley, CA, USA) statistical analysis software.

3-Results

94 patients were enrolled during the period of this study. 22 participants had incomplete charts at the time of analysis. In addition, 9 patients had binocular exclusion criteria: diabetic retinopathy (one patient), severe dry Age Related Macular Degeneration (ARMD) (one patient), myopia exceeding -8.0 diopters and related retinal abnormalities (3 patients), amblyopia (one patient) and unreliable AVF test results (3 patients). Therefore, 31 patients were excluded from the study. Moreover, 6 eyes of the remaining 63 patients were excluded for the following reasons: amblyopia (one eye), unilateral unreliable AVF test result (2 eyes of 2 patients), history of retinal surgery for epiretinal membrane (one eye) and retinal detachment (one eye) and one eye lost due to extensive choroidal neovascularization. Finally, the data obtained from 120 eyes of 63 patients was analyzed.

3.1- Demographics

The patients included 47 (74.6%) females and 16 (25.4%) males. The mean age of the cases was 60.6 ± 11.6 (1 SD) years (range 34-80 years). The control group consisted of 54 eyes of 20 (71.4%) females and 8 (28.6%) males with mean age 56 ± 14 (1 SD) years (range 32-84 years).

The mean refractive error of the cases was -0.25 ± 2.5 (1 SD) diopters (range -8.0 to +8.0 diopters). 109 eyes (90.8%) had less than 4 diopters of refractive error, and 11 eyes (9.2%) had moderate (4-8 diopters) refractive error. Best corrected visual acuity (BCVA) of those patients ranged from 20/20 to 20/50. RA (26 cases, 41.2%) and SLE (18

cases, 28.6%) constituted the major underlying systemic diseases mandating antimalarial therapy. The remaining cases were involved with other diseases like polymyalgia rheumatica, mixed connective tissue disease, Sjogren's syndrome and other connective tissue disorders. 6 patients (9.5%) with 12 eyes (10%) received CQ. 57 patients (90.5%) with 108 eyes (90%) were treated with HCQ.

3.2- Risk factors

Hepatic and renal dysfunctions were not associated with the development of toxicity in this series. There were 3 patients with a history of hepatic dysfunction and 3 others with some degree of renal failure who received HCQ; no sign of retinal toxicity was detected in these patients. Figure 15 illustrates the number and percentage of the CQ and HCQ patients having different risk factors with and without retinal toxicity, as defined by the findings of the reference test.

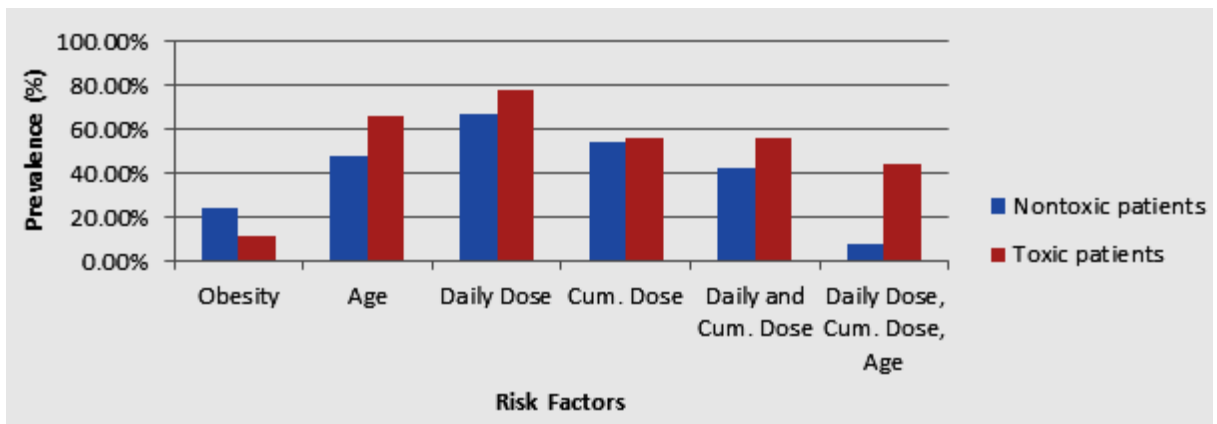


Figure 15: Prevalence of Risk Factors in Toxic and Non-toxic Patients based on the reference test results

Obesity was not more prevalent in toxic cases in this study. Each of the other risk factors was more often seen in toxic group; however their effect was much more significant when the 3 risk factors were present, making the older subjects with long-term history of receiving unsafe daily dose the most vulnerable group. Despite frequent recommendations in the literature, prescription of unsafe daily dose is still a common practice.

Figure-15 illustrates increased daily dose as the most prevalent single risk factor in all patients. Other than obesity, all of the risk factors were associated with increased incidence of toxicity. The most prominent difference was noted when 3 risk factors (daily dose, cumulative dose, and age) were all present in the same patient. Among the 9 patients with toxicity as detected by the reference test only one patient (11.1%) did not have any of the risk factors recorded in this study.

3.3- The influence of daily dose

To determine the influence of daily dose as a risk factor, the eyes were categorized into CQ and HCQ subgroups. All of the patients in the CQ group (6 patients, 12 eyes) received daily dose of more than 3mg/kgLBW/day, exceeding the safe limit (Marmor et al. AAO 2011). In the HCQ group (57 patients, 108 eyes), 18 patients (33 eyes) received a safe dose of ≤ 6.5 mg/kgLBW/day while 39 patients (75 eyes) were overdosed. In this series, the reference test was moderately superior to clinical exam in some HCQ cases. The frequency of detection of mfERG was greater than that of both clinical exam and the reference test in all subgroups.

6 patients (12 eyes) received CQ. All 6 patients in the CQ group were overdosed based on the safe limit of 3 mg/kgLBW/day. 4 eyes (33.3%) of 2 patients showed evidence of retinal toxicity on both clinical exam and the reference test.

Multifocal ERG was able to detect 2 more eyes (one patient) in the CQ subgroup on top of these 4 as abnormal [6 eyes (50%)].

57 patients (108 eyes) received HCQ. The HCQ eyes were also further subdivided into those overdosed and not overdosed based on the safe limit of 6.5mg/kgLBW/day. Again, mfERG was able to detect abnormality at a higher rate in both HCQ subgroups compared to clinical exam and the reference test.

Figure 16 describes the detection frequency of clinical exam, the reference test, and mfERG for HCQ/CQ retinal toxicity.

3.4- Correlation between cumulative dose and test results

In this study, the minimum cumulative dose associated with retinal toxicity was 450g for CQ and 875g for HCQ cases. Among the 12 eyes with HCQ retinal toxicity, only 7 (58.3%) eyes were exposed to a toxic cumulative dose more than 1000g. We found that in the absence of a high cumulative dose, other risk factors may be involved. 2 of the 5 remaining eyes (1 patient) without a toxic cumulative dose had two other risk factors: age greater than 65 years and toxic daily dose. 1 eye had high age as a risk factor while 2 eyes (1 patient) had none of the risk factors reviewed in this study.

Regression analyses were performed to evaluate the correlation between the cumulative dose and different test results. The results of the correlation analysis in the CQ group were not statistically meaningful due to the low number of eyes in this group.

In the HCQ group, mfERG P1 amplitude R5 ring ratio measures were correlated with the cumulative HCQ dosage ($P < 5\%$). In this analysis, the eyes with advanced,

clinically evident toxicity were excluded to focus on the preclinical stage of the disease. This correlation was statistically significant for mfERG only, with the p-value being most significant in the R2/R5 (parafoveal ring). The results did not show such a correlation between the HCQ cumulative dose and 10-2 AVF or sdOCT quantitative measures (Table-1).

Table-1 outlines the regression plot p-values for cumulative dose and test results.

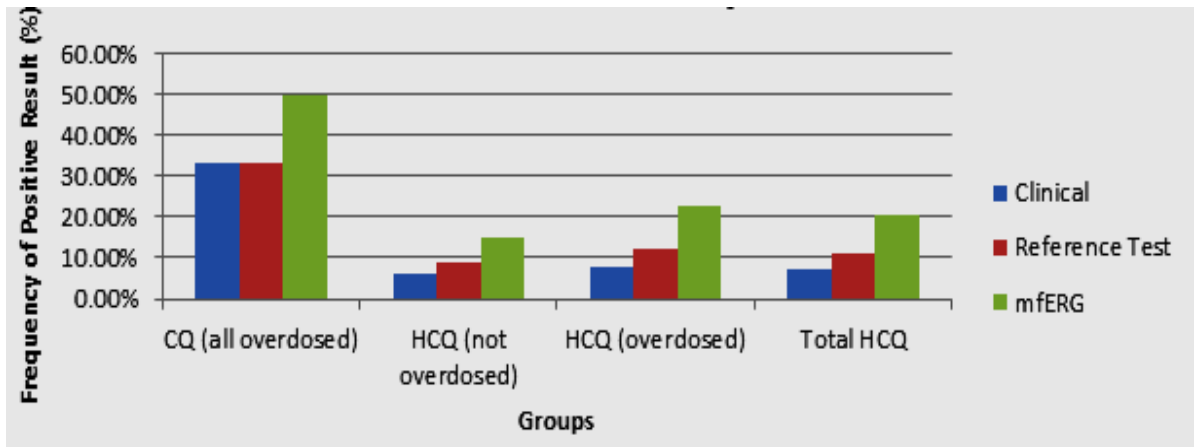


Figure 16: Frequency of Positive Test Results in CQ & HCQ Groups

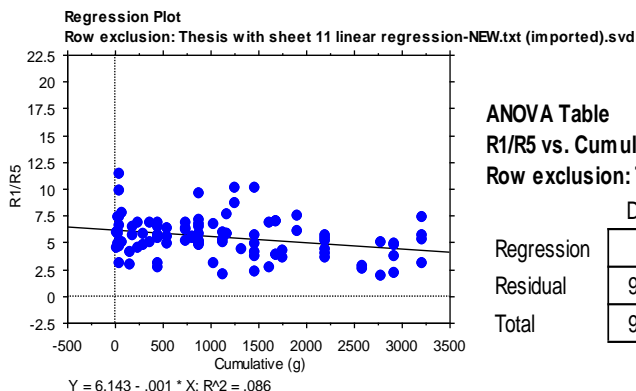
mfERG showed more positive test results in all groups.

	Multifocal ERG R5 ring ratio				10-2 AVF		sdOCT thickness		
	R1/R5	R2/R5	R3/R5	R4/R5	MD	PSD	Center	Middle	Outer
Regression P-Value	0.0040	0.0006	0.0443	0.0111	0.1826	0.7687	0.5599	0.4461	0.1427

Table 1: The regression analysis p-values between cumulative dose and test results

Analysis showed significant correlation between cumulative dose and mfERG amplitude in all rings, correlation being much more significant in the parafoveal ring 2. Such a correlation was not detected between the cumulative dose and the quantitative results of the reference test.

A- Regression Analysis: R1/R5 vs. Cumulative Dose



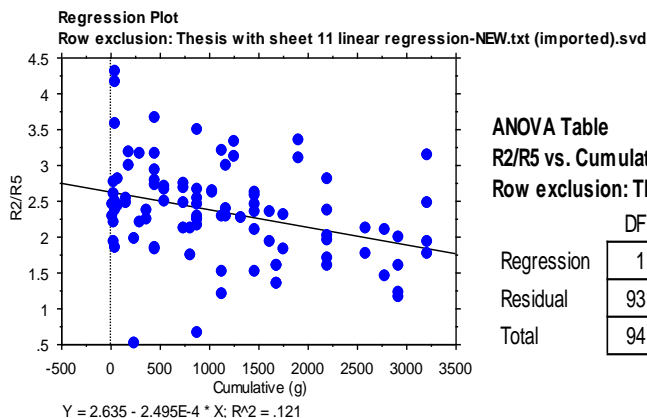
ANOVA Table

R1/R5 vs. Cumulative (g)

Row exclusion: Thesis with sheet 11 linear regression-NEW.txt (imported).svd

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	27.292	27.292	8.733	.0040
Residual	93	290.635	3.125		
Total	94	317.928			

B- Regression Analysis: R2/R5 vs. Cumulative



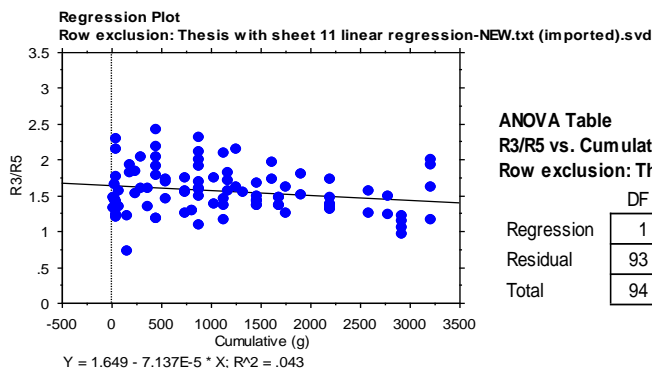
ANOVA Table

R2/R5 vs. Cumulative (g)

Row exclusion: Thesis with sheet 11 linear regression-NEW.txt (imported).svd

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	5.063	5.063	12.767	.0006
Residual	93	36.883	.397		
Total	94	41.946			

C- Regression Analysis: R3/R5 vs. Cumulative Dose



ANOVA Table

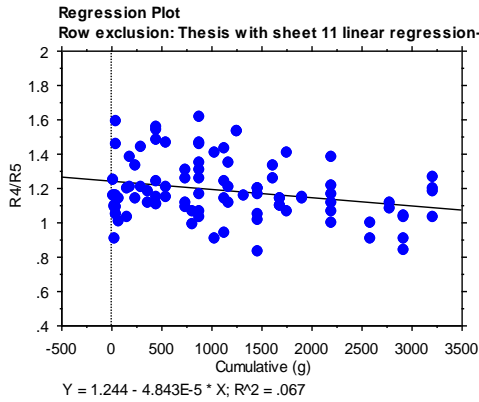
R3/R5 vs. Cumulative (g)

Row exclusion: Thesis with sheet 11 linear regression-NEW.txt (imported).svd

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	.414	.414	4.158	.0443
Residual	93	9.265	.100		
Total	94	9.679			

Figure 17 A, B&C: Regression analysis between R1,R2&R3/R5 ring ratios and HCQ cumulative dose

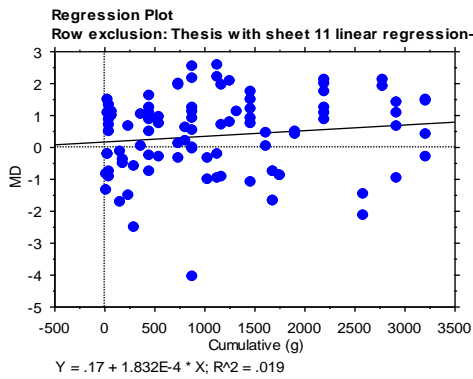
D- Regression Analysis: R4/R5 vs. Cumulative Dose



ANOVA Table
R4/R5 vs. Cumulative (g)
Row exclusion: Thesis with sheet 11 linear regression-NEW.txt (imported).svd

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	.191	.191	6.710	.0111
Residual	93	2.644	.028		
Total	94	2.834			

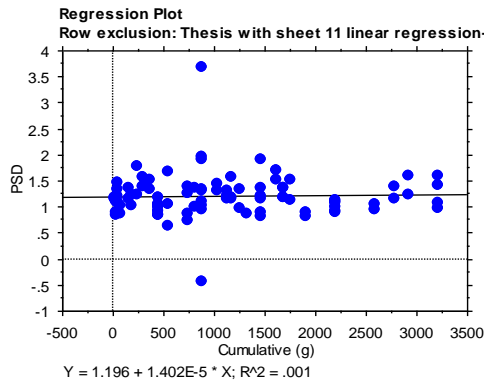
E- Regression Analysis: AVF MD vs. Cumulative Dose



ANOVA Table
MD vs. Cumulative (g)
Row exclusion: Thesis with sheet 11 linear regression-NEW.txt (imported).svd

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	2.729	2.729	1.803	.1826
Residual	93	140.747	1.513		
Total	94	143.476			

F- Regression Analysis: AVF PSD vs. Cumulative



ANOVA Table
PSD vs. Cumulative (g)
Row exclusion: Thesis with sheet 11 linear regression-NEW.txt (imported).svd

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	.015	.015	.087	.7687
Residual	91	15.394	.169		
Total	92	15.409			

Figure 17 D, E&F: Regression analysis between R4 /R5 ring ratio, MD&PSD and HCQ cumulative dose

G- Regression Analysis: Avg Thickness vs. Cumulative

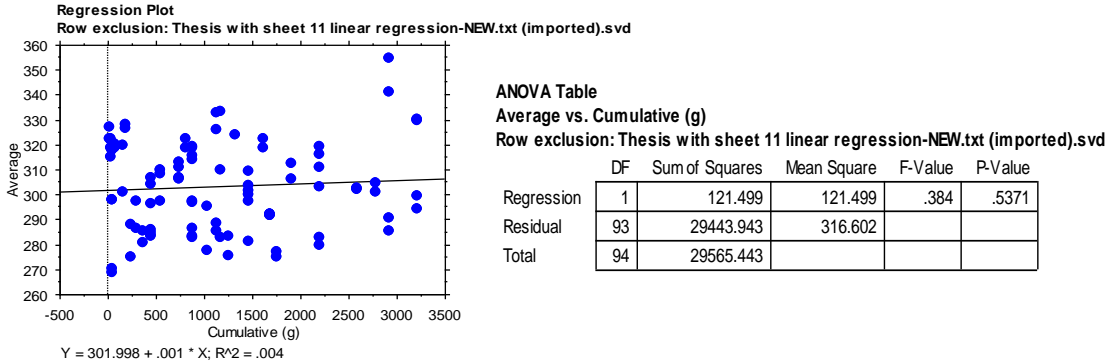


Figure 17G: Regression analysis between sdOCT average thickness and HCQ cumulative dose

Figure 17A-G: Regression and ANOVA analysis between cumulative dose and quantitative data of the tests results

(Only the average sdOCT macular thickness data is presented; the p-value for the analysis of the central circle and surrounding rings can be found in Table 1)

3.5- Sensitivity and specificity of mfERG

Since the clinical and diagnostic findings of CQ and HCQ retinal toxicity are identical, all eyes of patients receiving both drug types were grouped together in the statistical analysis for estimating sensitivity and specificity of mfERG. Using a 2 X 2 table the sensitivity of mfERG for CQ and HCQ retinal toxicity was calculated as 87.5% (95% CI: 60.4 to 97.8%). The specificity of this test was 86.5% (95% CI: 78.5 to 92.2%).

		Reference test sdOCT and 10-2 AVF		Totals
		Positive	Negative	
Target test mfERG	Positive	14 (TP)	14(FP)	28
	Negative	2(FN)	90 (TN)	92
		16	104	120

Table 2: Four different categories for mfERG test results relative to the reference test

	Estimated value	95% Confidence interval	
		Lower limit	Upper limit
Sensitivity	87.5%	60.4%	97.8%
Specificity	86.5%	78.5%	92.2%

Table 3: Calculated sensitivity and specificity of mfERG with 95% confidence interval

3.6- Relevance of mfERG test results to CQ/HCQ retinal toxicity

Analysis of Variance (ANOVA) with Bonferroni/Dunn post-hoc analysis was performed to determine any statistical differences between **1- Toxic (abnormal), 2- Borderline, 3- Non-toxic (normal) and 4- Control** mfERG results. Analysis showed that mfERG parafoveal R2/R5 and R3/R5 amplitude were significantly reduced in the toxic group compared to non-toxic and control groups. This finding is consistent with the classic anatomical configuration of CQ/HCQ retinal toxicity. The non-toxic group was not significantly different from the control group in any of the rings.

Similar analysis of mfERG implicit timing was also performed (significant difference: $p < 0.0083$). Ring average implicit timing in the toxic group was prolonged and different from non-toxic eyes in R2 ($p < 0.0065$), however, difference was more significant in outer rings R4 ($p < 0.0001$) and R5 ($p < 0.0001$). These measures were not significantly different in R1 ($p < 0.0705$) and R3 ($p < 0.1179$). Therefore, prolonged implicit timing was inconsistent with characteristic CQ/HCQ retinal toxicity.

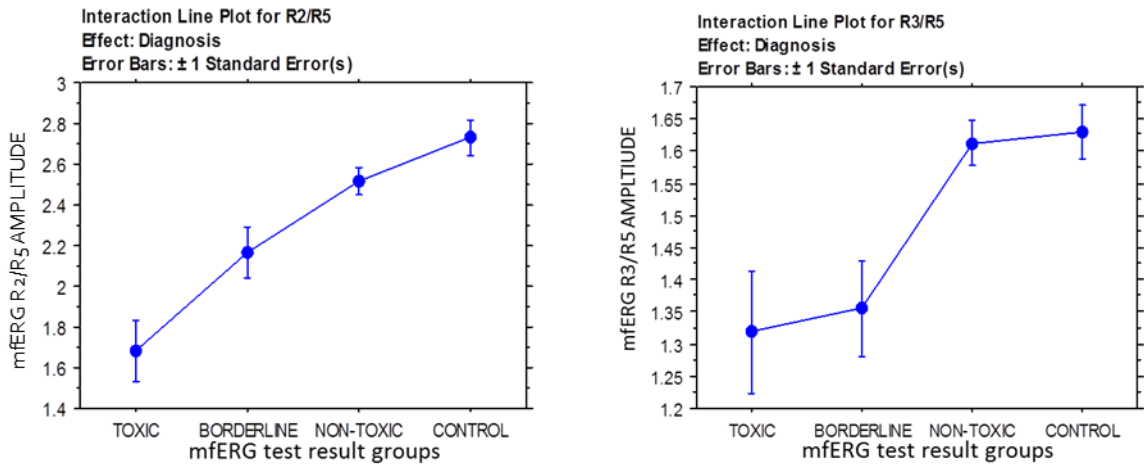


Figure 18: Interaction line plots demonstrating statistically significant differences in P1 amplitude R2/R5 (Left) and R3/R5 (Right) between the toxic and non-toxic eyes

The borderline group also shows amplitude reduction, but its difference is not significant.

	P-Value			
	R1/R5	R2/R5	R3/R5	R4/R5
Toxic, Non-Toxic	0.4047	<0.0001 (S)	0.0002 (S)	0.0092
Borderline, Non-Toxic	0.1354	0.0678	0.0143	0.0494
Non-Toxic, Control	0.5908	0.0634	0.7664	0.7726
<i>Comparisons in this table are not significant unless the corresponding p-value is less than 0.0083</i>				

Table 4: ANOVA P-value results demonstrating mfERG P1 amplitude R5 ring ratio differences between reportedly toxic, borderline, non-toxic and control groups. Statistically significant differences are limited to the P1 amplitude R5 ring ratios of the parafoveal rings, R2 and R3 between the toxic and non-toxic groups.

3.7- Analysis of mfERG false positives

Disagreement between the mfERG and the reference test (sdOCT and 10-2 AVF) generated 14 false positive and 2 false negative results. The number of false positives was not negligible and needed to be explained. False positive reports could reflect either a higher sensitivity of mfERG relative to the reference test, or from mfERG changes unrelated to actual toxicity. In order to make this distinction, the eyes of patients receiving CQ and HCQ were classified into 2 groups according to mfERG results: **1- False-Positive Toxic** and **2- Non-Toxic “Normal”**. ANOVA with Bonferroni/Dunn post-hoc analysis was performed to determine whether a significant statistical difference exists between false-positive toxic and non-toxic, normal mfERG results. Analysis revealed a statistically significant reduction in R2/R5 P1 amplitude of the “false positive” (those patients with abnormal mfERG and normal reference test) group relative to the non-toxic and control groups (Table 5). There were no other significant amplitude measurement differences between these groups in any of the other rings, and there were no differences between the non-toxic and control groups. This finding indicates that the false positive group showed typical signs of toxicity and could suggest a higher sensitivity of mfERG compared to the reference test. The analysis of implicit times was inconsistent with characteristic toxicity.

	P-Value			
	R1/R5	R2/R5	R3/R5	R4/R5
Toxic, Non-Toxic	0.2307	0.0002 (S)	0.0563	0.0291
Toxic, Control	0.1264	<0.0001 (S)	0.0414	0.0218
Non-Toxic, Control	0.5266	0.0568	0.7375	0.7507
<i>Comparisons in this table are not significant unless the corresponding p-value is less than 0.0167</i>				

Table 5: ANOVA P-value results demonstrating the P1 amplitude R5 ring ratio differences between the “False positive”, non-toxic and control groups

“False positive” group showed statistically significant P1 amplitude reduction in the parafoveal ring R2 compared to both non-toxic and control groups.

3.8- FAF findings

FAF results in the thesis study were inferior to the other three tests and slightly less sensitive than clinical fundus exam. FAF failed to detect any abnormality in 5 out of 16 (31.3%) toxic cases detected by the reference test. It also did not show abnormal fluorescence in 2 eyes with clinically detected toxicity.

4- Discussion

Generally it is not easy to determine the sensitivity and specificity of a test when it is expected to detect early stages of an uncommon disease. Despite the prolonged period of data collection, the results may still suffer from small sample size. In addition, the early signs of toxicity as detected by a very sensitive test may seem non-specific or irrelevant to the target disease. Calculation of sensitivity and specificity of mfERG in early detection of CQ/HCQ retinal toxicity is not an exception. The number of positive cases in this study was not adequate to provide strong enough evidence required to significantly influence clinical practice patterns; however, there was enough data to establish a relationship between the clinically recordable abnormalities and the cellular mechanisms of toxicity. Current knowledge regarding the accumulation of the drugs in the retinal neurons and RPE cells, and the topographical location of toxicity are crucial clues to understand the outcomes of the clinical tests. This subject is discussed in more detail here.

Retina vs. RPE: The concept of long term damage of the retina by the melanin bound CQ in the RPE has been strongly debated and there is evidence that retinal toxicity mainly arises from the intracellular interactions of the drug with the lysosomal activities.

Retina: Electron microscopic studies indicate the presence of *membranous cytoplasmic bodies (MCBs)* in the retinal neurons of CQ treated rats (Abraham and Hendy 1970). In another study on rats, Mahon and colleagues (Mahon et al. 2004) described the ultra-structural changes related to CQ toxicity. They reported that the pathologic findings could be seen both in neural retina and RPE. Among the inner retinal neurons, accumulation of MCBs was more significant in ganglion cells compared to bipolar, amacrine and horizontal cells (Dina et al. 2007, Mahon et al, 2004).

Rosental and colleagues reported that in CQ treated rhesus monkeys, ganglion cells revealed the earliest pathologic changes, which were reversible upon discontinuation of the drug up to 6 months after treatment onset, with total disappearance of the darkly staining MCBs. With long-term CQ treatment beyond 6 months, permanent damage and degeneration of the ganglion cells was evident (Rosental et al. 1978).

The results of another study have shown that CQ treatment of rats causes severe *lipoidosis* in inner retinal elements, ganglion and bipolar cells, which is associated with significant reduction of the ERG b-wave amplitude. At this stage, there was no significant involvement of photoreceptors or reduction of ERG a-wave amplitude. *Lipoidosis* of the inner retinal cells was reversed after cessation of the drug but photoreceptor degeneration progressed if it was already started (Dumker and Beredehorn 1995).

While photoreceptors were involved later than ganglion cells, their destruction was more severe and irreversible (Rosental et al. 1978). There were also differences in the

susceptibility to CQ damage between the two photoreceptor cells types. MCBs build-up was more prominent in cones than in rods (Dina et al. 2007, Mahon et al. 2004). Damage to the photoreceptors and other retinal neurons became evident somewhat later but degenerative changes of these cells could finally be demonstrated in more prolonged use of CQ. Reduction of parafoveal retinal thickness was a consistent early histologic finding (Dina et al. 2007, Rosental et al. 1978).

RPE: Involvement of RPE cells due to high affinity of CQ to melanin could theoretically explain “the secondary damage to the retina”. Rosental and colleagues demonstrated that despite accumulation of the CQ by-products in the RPE cells, toxic damage in these cells occurred later than that in retinal neurons (Rosental et al. 1978). They suggested that toxic damage to retina precedes destruction of the RPE cells. Mahon and colleagues (Mahon et al. 2004) on the other hand, detected some early changes in RPE cells in the form of increased number of “lysosome associated organelles (LAOs)”. They proposed that the lysosomotropic properties of CQ with resultant generalized dysfunction of the lysosomal system in both retina and RPE is the basis of CQ retinal toxicity.

There are theories both in favor and against the contribution of melanin to the development and advancement of toxicity. Melanin binding and accumulation of CQ in the RPE cells might explain progression of retinal neuron toxicity despite their cessation or conversely, it may play a protective role by attracting a significant amount of toxic material away from the retina (Marmor et al. AAO 2011).

Although in its very advanced form, CQ retinal toxicity can be associated with widespread atrophy of the retina, RPE and choroid (Hobbs et al. 1959, Rosental et al. 1978), any theory for the primary role of retina vs. RPE in the genesis of toxicity should ideally explain the characteristic anatomical localization of the lesion in less severe stages. The most prominent histologic feature of the parafoveal area is the multi-layered organization of retinal ganglion cells. As mentioned earlier, these cells show the earliest histologic signs of toxicity. Another study has shown that chronic use of HCQ is associated with selective thinning of the inner retina, particularly ganglion cells and the inner plexiform layer even in the absence of clinically detectable functional and structural changes in the photoreceptors and RPE (Pasadhika et al. 2010). Authors of this study hypothesize that while all retinal ganglion cells accumulate HCQ, effects are most prominent in the para-foveal area where the ganglion cells are most abundant. Despite the differences in ganglion cell types and their distribution in the retina, differential cell type susceptibility to CQ has not been mentioned in the references that were reviewed. Therefore, it seems that the abundance of parafoveal ganglion cells, rather than their specific cell type, is the major determinant of parafoveal localization of early toxicity. While these authors did not detect any concomitant reduction of the nerve fiber layer thickness, this phenomenon was reported in another study (Xiaoyun et al. 2010).

The central 400 μ m of the fovea known as the foveal avascular zone (FAZ) is totally devoid of retinal capillaries. Its surrounding perifoveal area, in addition to having the thickest ganglion cell layer, contains a dense and deep retinal capillary network. This vascular architecture may also contribute to the bull's eye configuration of early toxicity;

however, to date, no histologic study has linked it with macular vasculature (Browning 2014).

The theory for RPE-basis of toxicity fails to explain the anatomical location of the lesion since this tissue does not have specific characteristics in that topographical location.

Overall, the theory for retinal-basis of toxicity and particularly involving inner retinal elements, at least in its early stages, has received more support than that of RPE.

4.1- The influence of risk factors in the development of toxicity

Genetic predisposition as a risk factor for CQ/HCQ retinal toxicity was not considered in this study. Among the other known risk factors, obesity, hepatic dysfunction and renal failure were not associated with increased incidence of toxicity. Age, daily dose and cumulative dose, however, whether alone or in combination, could play a role in the development of toxicity. Patients with a combination of these three risk factors had the highest chance of developing toxicity (Figure 15).

Unsafe daily doses of CQ and HCQ are still commonly prescribed. Mackenzie (Mackenzie 1983) found that in a population of 900 RA patients, no eye disease developed when daily dose was limited to less than 6.5 mg/kg of HCQ and 4.0mg/kg of CQ in a mean of about 7 years. Considering the very low amount of fat distribution of these drugs, the author recommended that this limit should be adjusted based on lean body weight (LBW), rather than actual weight. The study did not find a significant correlation between toxicity and cumulative dose. Berstein (Berstein 1992) reported

almost identical results. There are reports that have failed to find a relationship between daily dose and the development of retinal effects (Yaylali et al. 2013), however, the bulk of current data indicates that in most cases of retinal toxicity these limits have been exceeded (Tsank et al. 2014, Pyne et al. 2011, Walvick et al. 2011, Michaelidis et al. 2011). Since the dose should be titrated based on LBW, obese patients are often overdosed, but overdosing is not exclusive to obesity. For instance, the lean body weight of men shorter than 5 feet 4 inches and women shorter than 5 feet 6 inches is generally less than 135 pounds. A conventional dose of 400 mg of HCQ (2 tablets) per day predisposes them to an increased risk of retinal toxicity (Lahley 2008). Due to the dosage of available preparations, generally an initial “standard” dose of 400mg of HCQ and 250mg of CQ is prescribed and continued for an extended period of several years. With the long half-life of these medications, other options should be considered to avoid unnecessary exposure to retinal toxicity, for example, keeping the patient off the drug for one or two days per week, or alternate day therapy with 400 and 200 mg doses of HCQ, particularly in the care of patients over 65 years of age.

In this study, daily dose exceeding the safe limits was the most prevalent risk factor both in the toxic and non-toxic groups. Much consideration should be given to daily dose not only because it is a frequently reported risk factor, but also because it is the one which can be most easily adjusted. In addition, over the long-term course of CQ/HCQ therapy, a lower daily dose can diminish another significant risk factor, the cumulative dose.

The effect of cumulative dose as a risk factor for toxicity was slightly less prominent in this study. Almost the same percentage of patients with and without toxicity had already received a high risk cumulative dose. Nevertheless, the findings of the studies concerning the cellular mechanisms of CQ retinal toxicity (Rosenthal et al. 1978, Mahon et al. 2004) indicate that cumulative dose, which is the product of daily dose and duration of therapy, should influence test results. The findings of this study support this idea, but the effect of cumulative dose on the development of toxicity and test results is discussed later.

4.1.1- Correlation between test results and cumulative dose

It is known that the accumulation of CQ-related intracellular inclusion bodies and lysosomal dysfunction is dependent on the dose and duration of therapy (Mahon et al. 2004, Rosenthal et al. 1978). Therefore, it can be expected that a relationship might exist between the cumulative dose and disturbances of retinal electrical activity. In fact, a significant correlation between HCQ cumulative dose and reduced mfERG absolute amplitudes in all rings has already been reported (Lai et al. 2005). This thesis study has found similar results. Moreover, possibly as a result of analysing highly sensitive R5 ring ratios, the impairment was shown to be characteristically more severe in the parafoveal R2 area (Figure 17 and Table 1). This finding also indicates that the detectable electrophysiological disturbances are initially more severe in but not exclusive to the parafoveal area. Figure 19 illustrates mfERG deviation map, sdOCT foveal cross sections and associated FAF images of patient 94 of this study. As seen, early toxicity is evident in both eyes. mfERG has detected some degree of reduced amplitude in all rings which is

most pronounced in ring 2. Meanwhile, sdOCT shows normal IS/OS junction line and mild reduction of parafoveal ONL thickness. No abnormality is evident in FAF image.

No correlation was observed between HCQ cumulative dose and mfERG implicit time, sdOCT thickness measurements and 10-2 AVF MD and PSD.

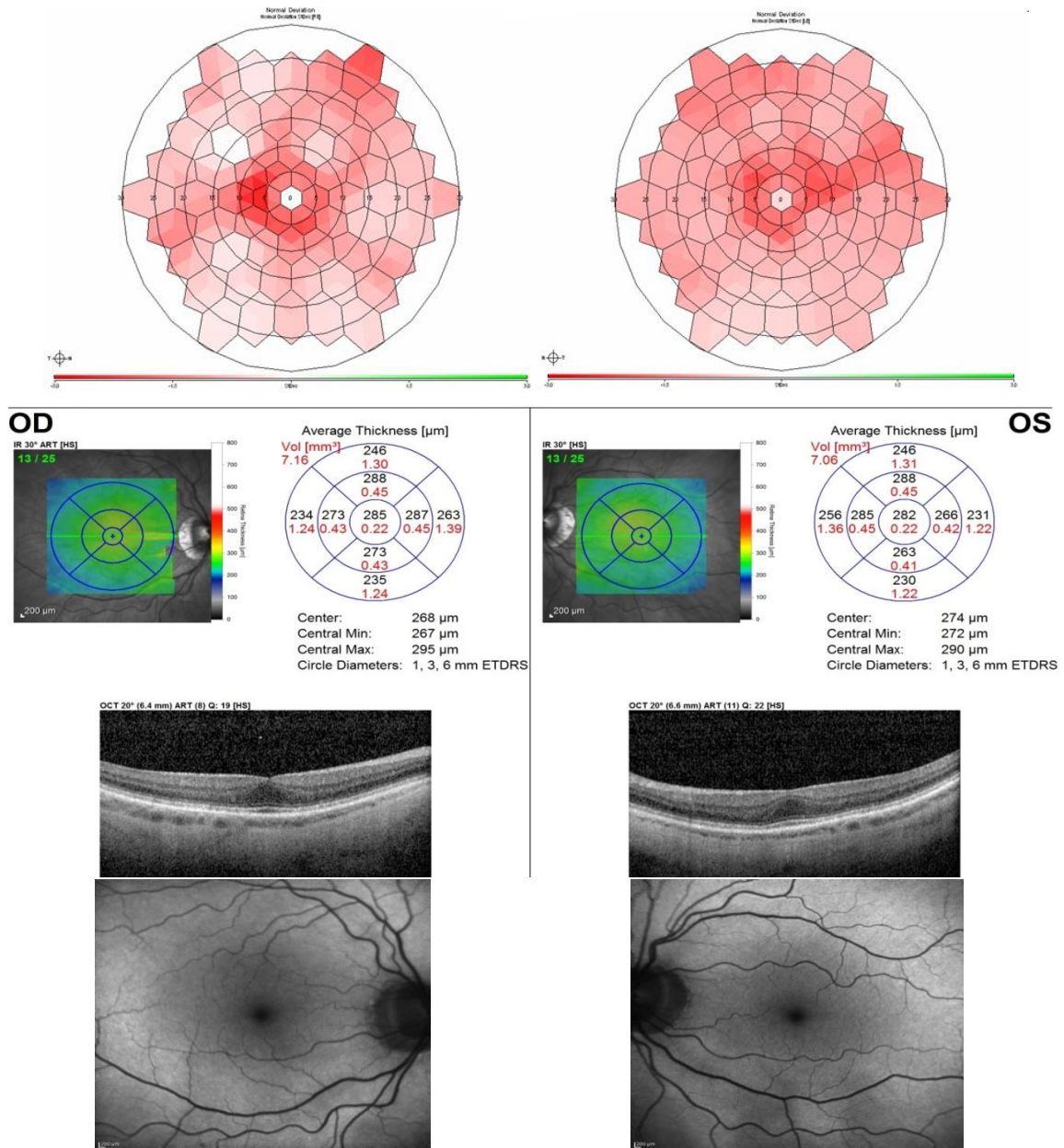


Figure 19: mfERG deviation map (Top), sdOCT imaging (Middle) and FAF (Bottom) of both eyes of a patient with HCQ retinal toxicity

mfERG has recorded diffuse amplitude reduction in 5 rings, which is most prominent in the parafoveal region. This finding is very consistent with histologic findings of toxicity. Meanwhile, abnormalities in the structural images are limited to mild parafoveal retinal thinning in the sdOCT cross sections.

4.2- The model for mfERG sensitivity and specificity calculation

While there is evidence in the literature that mfERG may be more sensitive than the other recommended screening tests, their sensitivity and specificity have never been determined (Marmor et al. AAO 2011). A model was therefore required and developed using a combination of 10-2 AVF and sdOCT as the reference test with inclusion of both eyes.

4.2.1- Inclusion of one or both eyes

In ophthalmic studies, something which needs to be established is whether to include only one or both eligible eyes of binocular patients. CQ/HCQ retinal toxicity is a bilateral disease, at least in its later stages. Since the results may be correlated when both eyes are included, some studies have considered only one eye of each patient in the evaluation of the CQ/HCQ retinal toxicity to enhance the precision of statistical analysis (Mißner and Kellner 2012, Adam et al. 2011, Browning and Lee 2014). This approach has both advantages and disadvantages. The major disadvantages of including only one eye are: non-random selection bias, non-optimal power and precision of the analysis, and loss of information (Murdoch et al. 1998). It should be noted that while involvement of both eyes is a characteristic finding in later stages of the disease, test data may appear asymmetrical between the two eyes of the same patient in earlier stages (Figure 20) (Kellner et al. 2006, Michaelidis et al. 2011, Sheroyer et al. 2001). Diagnosis of CQ/HCQ retinal toxicity can be made with more certainty when abnormal findings are bilateral, but including both eyes in the quantitative analysis may increase the power of the results. Many of the previously published studies on CQ/HCQ retinal

toxicity have included both eyes (Pasadhika et al. 2010, Rodriguez-Padilla et al. 2007, Lyons and Severns 2007, Lai et al. 2006).

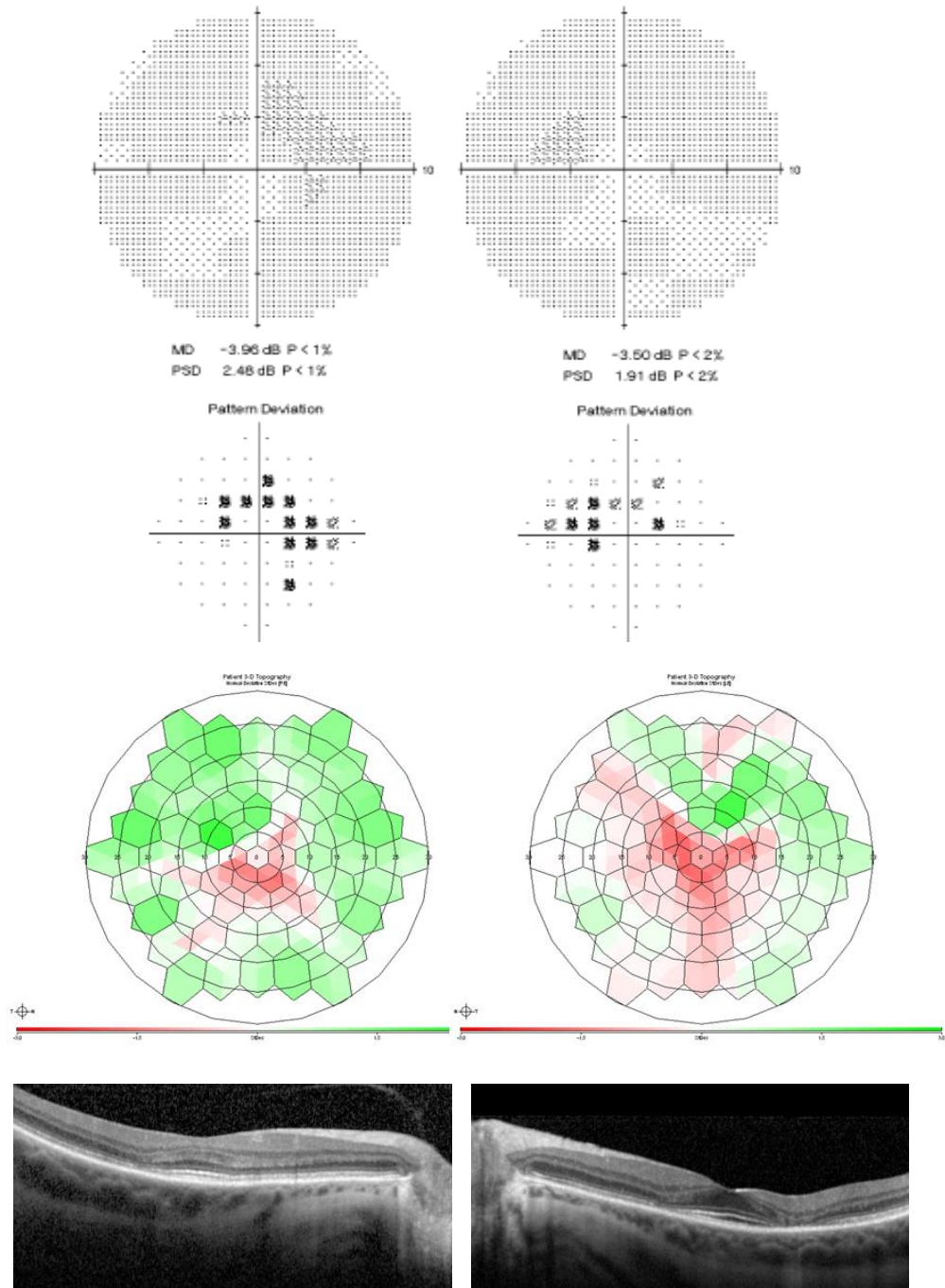


Figure 20: Asymmetry of retinal damage in HCQ retinal toxicity between two eyes using all test modalities

The lesion may show asymmetries relative to the horizontal and vertical meridians.

4.2.2- Humphrey 10-2 AVF as the mainstay of screening

Visual field testing has long been the mainstay in the diagnosis of CQ/HCC retinal toxicity. Even the first description of the disease was associated with documented para foveal field loss (Hobbs et al. 1959). In a prospective study of 1500 patients, 10-2 AVF was more sensitive than color vision testing and fluorescein angiography (Easterbrook 1988). 10-2 AVF could detect toxicity early enough to stop the disease progression after cessation of therapy in a significant number of eyes, whereas the other tests could only detect toxicity in its late, irreversible stage. In later years, AVF became the standard clinical diagnostic tool (Marmor et al. AAO 2002) and reference test in many studies to which newer diagnostic modalities were compared (Browning 2002, So et al. 2003, Lai et al. 2005, Lai et al. 2006, Lyons and Severns 2006, Rodriguez-Padilla et al. 2007, Stepien et al. 2009).

Retrospective studies have shown that early signs of retinopathy detected by AVF have been disregarded based on their p-value or small size of the scotoma. This practice has recently been discouraged since AVF abnormalities may become apparent with subtle loss of threshold many months or even years before a diagnosis can be made clinically (Figure-12) (Anderson et al. 2011, Marmor et al. AAO 2011). The AAO 2011 document recommends that upon observation of persistent, minor AVF abnormalities, an objective test (FAF, mfERG, or sdOCT as available) should be ordered to evaluate it better. The pattern deviation map is the standard item to review (Marmor et al. 2013),

but AVF abnormalities are not limited to those found in the deviation maps; Lyons has reported that pattern standard deviation is also a reliable, quantitative index of early HCQ retinal toxicity (Lyons 2013).

The debate of red versus white target stimuli has recently been re-evaluated. It appears that a white stimulus, despite being slightly less sensitive in the detection of early toxicity, is more specific and more consistent over repeated testing (Marmor et al. 2013). Overall, the authors suggested that there is no preference about the color of the target stimulus.

The approach used in this study to categorize the cases for sensitivity and specificity calculation was based on the AAO 2011 guidelines (Figure-9). More than 3 abnormal points in the center or mid periphery of the 10-2 AVF deviation map proved to be an acceptable sign of “probable” toxicity. Notably, all of the 10-2 AVF test results that were labeled “probable” toxicity also showed characteristic toxic signs in sdOCT. None of the 10-2 AVF “possible” toxicity cases had abnormal sdOCT. The percentage of disagreement between 10-2 AVF and sdOCT in a recently published study was not significant. In a review of 150 cases with definitive paraclinical signs of HCQ retinal toxicity, 10% of eyes with “pathognomonic” 10-2 AVF results were totally normal on sdOCT (Marmor and Melles 2014). The authors denoted that these results were obtained in a retrospective study that excluded poorly reliable AVF test results. Another recently published study also denotes correlation between functional abnormalities detected by AVF testing and histological abnormalities resulting from this toxicity (Bae et al. 2014).

Hence, their findings are not necessarily indicative of a lower sensitivity of sdOCT or earlier appearance of functional changes. Therefore, the reference test of our study seemed appropriate and in accordance with the AAO 2011 recommendations.

4.3- Sensitivity and specificity of mfERG

The sensitivity and specificity of mfERG as compared to the reference test were 87% (95% CI: 60.4% to 97.8%) and 86.5% (95% CI: 78.5% to 92.2%) respectively. The relatively wide confidence interval can be explained by the low number of positive cases. This problem is frequently encountered when calculating test validity parameters of uncommon conditions. One approach to narrow the CI range is to include more patients, which can require substantially more time. Therefore, statistical analysis of the present data could help towards making a conclusion about the calculated values.

4.3.1- Interpretation of the mfERG “false positive” cases

As seen in table 2 the number of false positive mfERG test results is unexpectedly high and equal to the number of true positive cases. These are the reportedly “abnormal” mfERG cases that did not show agreement with the reference test.

An explanation for this finding might be that the target test is more sensitive than the reference test, and so some of the true positive cases of the target test are inadvertently categorized as false. If this occurs, both the sensitivity and specificity of the target test are erroneously underestimated. The most appropriate approach in dealing

with this situation might be to perform a longitudinal study on the false positive cases. Over time, the reference test may show evidence of actual toxicity in cases that were previously labeled as false positive. This longitudinal approach is not always practical however, especially in uncommon diseases or when the disease follows a slow and prolonged course. Other factors making longitudinal study challenging include loss of follow up due to individual or familial restrictions, shifting to other medications, or migration to a distant area. Analysis of the present data could give an instant answer about the validity of these results. The findings of this study showed that mfERG “false positive” cases were significantly affected in the characteristic topographical zone of early toxicity, indicating a higher sensitivity of mfERG compared to the reference test.

4.3.2- Current knowledge of sensitivity and specificity of the screening tests

The sensitivity and specificity of each currently recommended screening test was not known when the 2011 AAO document was published. These parameters for 10-2 AVF, mfERG and sdOCT have been presented in a 2014 publication (Browning and Lee 2014) and are summarized in Table 6. FAF was omitted due to its lower sensitivity in early stages of toxicity.

	10-2 AVF	mfERG	sdOCT
Sensitivity	85.7%	92.9%	78.6%
Specificity	92.5%	86.9%	98.1%

Table 6: Sensitivity and specificity of 10-2 AVF, mfERG, sdOCT as determined by Browning and Lee (Browning and Lee 2014)

The authors used a 2 X 2 table approach. The gold standard reference test was *“the total evidence for toxicity which resulted in the decision to stop the medication by the ophthalmologist”*. While their article has the advantage of estimating the sensitivity of all 3 tests in one study, the reference test - decision to stop the medication - is not clear to the reader and might have been variable between the participating ophthalmologists. Nevertheless, despite differences in the methods used, their mfERG findings are similar to those of this thesis study. Additionally, if the analysis of the mfERG “false positive” cases is considered, the actual sensitivity of this test can be higher than what was presented by both studies.

4.3.3- Validity of mfERG independent of the reference test

In this study 28 eyes (23.3%) in both CQ and HCQ groups showed “abnormal” mfERG results. A relatively high number of abnormal mfERG tests results have also been observed in some previously published studies (So et al. 2003, Moschos et al. 2004, Lai et al. 2005, Lyons and Severns 2006).

Browning and Lee have argued that mfERG is not a perfect objective test because its results are affected by the subjectivity in interpretation (Browning and Lee 2014); however, mfERG amplitude and implicit time measures are quantitative and standard protocols for clinical interpretation can be effectively developed and implemented. In this study, to validate the mfERG results independent of the reference test, the reportedly abnormal mfERG cases were compared to reportedly normal cases

and controls using ANOVA analysis. The abnormal mfERG group showed significant amplitude reduction in the classic area affected by toxicity (Figure 18 and Table 4). mfERG implicit timing data were not consistent nor typical of toxicity.

These findings indicate that the reference test is not sensitive enough to be the gold standard for sensitivity and specificity calculation. Therefore, other statistical methods such as *Latent Class Analysis (LCA)* may prove superior to 2 X 2 tables for estimating the target test validity (Rutjes et al. 2007). A major advantage of this type of analysis is that it does not require a gold standard reference test and as a result, no characteristically positive test result is erroneously labeled as false positive. LCA was not possible in this study due to the low number of positive cases. The prospective arm of this study may eventually enroll an adequate number of affected eyes to give a better estimate of mfERG validity.

5- Conclusion

Long term use of CQ and HCQ causes gradual disruptions of the vital cellular mechanisms in retinal neurons and RPE cells. However, it remains unclear why only a small fraction of users develop clinically evident toxicity. Risk factors play a prominent role, but cannot explain all cases of toxicity development. The understanding of the cellular mechanisms and characteristic topographical localization of early abnormalities can aid in interpreting the clinical test results and validating their findings. Currently recommended diagnostic tests have shown to be capable of demonstrating the signs of cellular stress and injury in a preclinical stage of toxicity. The results of this thesis study indicate that mfERG is capable of detecting characteristic functional abnormalities before they can be detected by other recommended tests. Multifocal ERG P1 amplitude when normalized using R5 ring ratios is a reliable indicator of early retinal toxicity induced by these drugs. A gold standard test to be used as the reference for evaluation of the sensitivity and specificity of mfERG does not exist and statistical methods other than 2X2 tables should be used for this calculation.

References

- 1 Abraham R, Hendy RJ, 1970, Irreversible lysosomal damage induced by chloroquine in the retinae of pigmented and albino rats. *Exp Mol Pathol.* Apr;12(2):185-200.
- 2 Adam MK, Covert DJ, Stepien KE, Han DP, 2012, Quantitative assessment of the 103-hexagon multifocal electroretinogram in detection of hydroxychloroquine retinal toxicity. *Br J Ophthalmol.* May;96(5):723-9.
- 3 Anderson C, Blaha GR, Marx JL, 2011, Humphrey visual field findings in hydroxychloroquine toxicity. *Eye (Lond).* Dec;25(12):1535-45.
- 4 Aylward JM, 1993, Hydroxychloroquine and chloroquine: assessing the risk of retinal toxicity. *J Am Optom Assoc.* Nov;64(11):787-97.
- 24 Bae EJ, Kim KR, Tsang SH, Park SP, Chang S, 2014, Retinal Damage in Chloroquine Maculopathy, Revealed by High Resolution Imaging: A Case Report Utilizing Adaptive Optics Scanning Laser Ophthalmoscopy, *Korean J Ophthalmol.* Feb; 28(1): 100–107
- 5 Bellmann C, Neveu MM, Kousoulides L, Sloper JJ, Bird AC, Holder GE, 2004, Potential diagnostic dilemmas using the multifocal electroretinogram in intermittent exotropia. *Br J Ophthalmol.* Sep;88(9):1223-4.
- 6 Bernstein HN. 1992, Ocular safety of hydroxychloroquine sulfate (Plaquenil). *South Med J.* Mar;85(3):274-9.
- 7 Brinkley JR Jr, Dubois EL, Ryan SJ, 1979 Long-term course of chloroquine retinopathy after cessation of medication. *Am J Ophthalmol.* Jul;88(1):1-11.
- 8 "Brotten L, Aviña-Zubieta JA, Lacaille D, Joseph L, Hanly JG, Lix L, O'Donnell S, Barnabe C, Fortin PR, Hudson M, Jean S, Peschken C, Edworthy SM, Svenson L, Pineau CA, Clarke AE, Smith M, Bélisle P, Badley EM, Bergeron L, Bernatsky S, 2014, Systemic autoimmune

rheumatic disease prevalence in Canada: updated analyses across 7 provinces. *J Rheumatol.* Apr;41(4):673-9. doi: 10.3899/jrheum.130667.

- 9 Browning DJ, 2002, Hydroxychloroquine and chloroquine retinopathy: screening for drug toxicity. *Am J Ophthalmol.* May;133(5):649-56.
- 11 Browning DJ, 2014 *Hydroxychloroquine and Chloroquine Retinopathy*, Springer New York, May 10.
- 10 Browning DJ, Lee C, 2014, Relative sensitivity and specificity of 10-2 visual fields, multifocal electroretinography, and spectral domain optical coherence tomography in detecting hydroxychloroquine and chloroquine retinopathy. *Clin Ophthalmol.* Jul 25;8:1389-99.
- 12 Browning DJ, 2013, Impact of the revised american academy of ophthalmology guidelines regarding hydroxychloroquine screening on actual practice. *Am J Ophthalmol.* Mar;155(3):418-428.e1.
- 13 Chen E, Brown DM, Benz MS, Fish RH, Wong TP, Kim RY, Major JC, 2010, Spectral domain optical coherence tomography as an effective screening test for hydroxychloroquine retinopathy (the "flying saucer" sign). *Clin Ophthalmol.* Oct 21;4:1151-8.
- 14 Cox DR, Hinkley DV, 1974 *Theoretical Statistics*, Chapman & Hall, p49, p209
- 15 Creel DJ, 2011, Multifocal electroretinograms. *J Vis Exp.* Dec 4;(58).
- 16 David J. Browning, 2014, *Hydroxychloroquine and Chloroquine Retinopathy*. books.google.com/books?isbn=149390597X
- 17 Davis MJ, Dawes PT, Fowler PD, Clarke S, Fisher J, Shadforth MF, 1991, Should disease-modifying agents be used in mild rheumatoid arthritis? *Br J Rheumatol.* Dec;30(6):451-4.
- 18 Dina H. Abdel Kader, Safinaz Salah El Din Sayed and Tarek A. El-Ghamrawy , 2007. Chloroquine-Induced Retinopathy in the Rat. *Immunohistochemical and Ultrastructural*

Study. *Journal of Medical Sciences*, 7: 1225-1238.

- 19 Dosso A, Rungger-Brändle E, 2007, In vivo confocal microscopy in hydroxychloroquine-induced keratopathy. *Graefes Arch Clin Exp Ophthalmol*. Feb;245(2):318-20.
- 20 Duncker G, Bredehorn T, 1995, Chloroquine-induced lipidosis in the rat retina: functional and morphological changes after withdrawal of the drug. *Graefe's Archive for Clinical and Experimental Ophthalmology* Vol: 234(6):378-381
- 21 Easterbrook M, 1988, Ocular effects and safety of antimalarial agents. *Am J Med*. Oct 14;85(4A):23-9.
- 22 Easterbrook M, 1999, Detection and prevention of maculopathy associated with antimalarial agents. *Int Ophthalmol Clin*. Spring;39(2):49-57.
- 23 Ehrenfeld M, Neshor R, Merin S, 1986, Delayed-onset chloroquine retinopathy. *Br J Ophthalmol*. Apr;70(4):281-3.
- 25 Fielder A, Graham E, Jones S, Silman A, Tullo A. Royal College of Ophthalmologists guidelines: ocular toxicity and hydroxychloroquine. *Eye (Lond)*. 12 (Pt 6):907-9.
- 26 Finbloom DS, Silver K, Newsome DA, Gunkel R., 1985 Comparison of hydroxychloroquine and chloroquine use and the development of retinal toxicity. *Rheumatol*. Aug;12(4):692-4.
- 27 Fox RI, 1993, Mechanism of action of hydroxychloroquine as an antirheumatic drug. *Semin Arthritis Rheum*. Oct;23(2 Suppl 1):82-91.
- 28 Gabriele ML, Wollstein G, Ishikawa H, Kagemann L, Xu J, Folio LS, Schuman JS. 2011, Optical coherence tomography: history, current status, and laboratory work. *Invest Ophthalmol Vis Sci*. Apr 14;52(5):2425-36.
- 29 Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, Liang MH, Kremers HM, Mayes MD, Merkel PA, Pillemer SR, Reveille JD, Stone JH;National Arthritis Data Workgroup, 2008 Estimates of the prevalence of arthritis and other rheumatic conditions

in the United States. Part I. Arthritis Rheum. Jan;58(1):15-25. doi: 10.1002/art.23177

- 30 Hobbs HE, Sorsby A, Freedman A, 1959, Retinopathy following chloroquine therapy. Lancet. Oct 3;2(7101):478-80.
- 31 Holz FG, Bellman C, Staudt S, Schütt F, Völcker HE, 2001, Fundus autofluorescence and development of geographic atrophy in age-related macular degeneration. Invest Ophthalmol Vis Sci. Apr;42(5):1051-6.
- 32 Hood DC, 2011, ISCEV standard for clinical multifocal electroretinography. iscev.org/.../ISC...
- 33 Hood DC, Seiple W, Holopigian K, Greenstein V, 1997, A comparison of the components of the multifocal and full-field ERGs. Vis Neurosci. 1997 May-Jun;14(3):533-44.
- 34 Hume R, 1966, Prediction of lean body mass from height and weight. Journal of clinical pathology 19 (4): 389-91
- 35 Humphrey Field Analyzer Manual (5.1, for Series II instruments) - Nov 11, 2012 , Free ebook download as PDF File (.pdf),
- 36 Jivrajka RV, Genead MA, McAnany JJ, Chow CC, Mieler WF, 2013, Microperimetric sensitivity in patients on hydroxychloroquine (Plaquenil) therapy. Eye (Lond). Sep;27(9):1044-52.
- 37 Kahn JB, Haberman ID, Reddy S, 2011, Spectral-domain optical coherence tomography as a screening technique for chloroquine and hydroxychloroquine retinal toxicity. Ophthalmic Surg Lasers Imaging. Nov-Dec;42(6):493-7.
- 67 Kazi MS, Saurabh K, Rishi P, and Rishi E, 2014, Delayed Onset Chloroquine Retinopathy Presenting 10 Years after Long-Term Usage of Chloroquine. Clin Ophthalmol. Jul 25;8:1389-99.
- 38 Kellner S, Weinitz S, Farmand G, Kellner U. 2014, Cystoid macular oedema and epiretinal membrane formation during progression of chloroquine retinopathy after drug cessation. Br J Ophthalmol. Feb;98(2):200-6.
- 39 Kellner S, Weinitz S, Kellner U, 2009, Spectral domain optical coherence tomography detects early stages of chloroquine retinopathy similar to multifocal electroretinography,

fundus autofluorescence and near-infrared autofluorescence. *Br J Ophthalmol.* Nov;93(11):1444-7.

- 40 Kellner U, Kraus H, Foerster MH, 2000, Multifocal ERG in chloroquine retinopathy: regional variance of retinal dysfunction. *Graefes Arch Clin Exp Ophthalmol.* Jan;238(1):94-7.
- 41 Kellner U, Renner AB, Tillack H, 2006, Fundus autofluorescence and mfERG for early detection of retinal alterations in patients using chloroquine/hydroxychloroquine. *Invest Ophthalmol Vis Sci.* Aug;47(8):3531-8.
- 42 Kennedy CJ, Rakoczy PE, Constable IJ, 1995, Lipofuscin of the retinal pigment epithelium: a review. *Eye (Lond).* 9 (Pt 6):763-71.
- 43 Lahey JM, Lehmer JM, Walvick MD, et al. 2008 Incidence and characteristics of plaquenil toxicity in a large population-based study. Presented as a poster at the American Academy of Ophthalmology Annual Meeting, November, Atlanta, Georgia.
- 44 Lai TY, Chan WM, Li H, Lai RY, Lam DS, 2005, Multifocal electroretinographic changes in patients receiving hydroxychloroquine therapy. *Am J Ophthalmol.* Nov;140(5):794-807.
- 45 Lai TY, Ngai JW, Chan WM, Lam DS, 2006, Visual field and multifocal electroretinography and their correlations in patients on hydroxychloroquine therapy. *Doc Ophthalmol.* May;112(3):177-87.
- 46 Lyons JS, 2013, Impact of the revised American Academy of Ophthalmology guidelines regarding hydroxychloroquine screening on actual practice. *Am J Ophthalmol.* Aug;156(2):410.
- 47 Lyons JS, Severns ML, 2007, Detection of early hydroxychloroquine retinal toxicity enhanced by ring ratio analysis of multifocal electroretinography. *Am J Ophthalmol.* May;143(5):801-809.
- 48 Lyons JS, Severns ML, 2009, Using multifocal ERG ring ratios to detect and follow Plaquenil retinal toxicity: a review : Review of mfERG ring ratios in Plaquenil toxicity. *Doc Ophthalmol.* Feb;118(1):29-36.
- 49 Mackenzie AH, 1983, Dose refinements in long-term therapy of rheumatoid arthritis with antimalarials. *Am J Med.* Jul 18;75(1A):40-5.

- 50 Mahon GJ, Anderson HR, Gardiner TA, McFarlane S, Archer DB, Stitt AW, 2004, Chloroquine causes lysosomal dysfunction in neural retina and RPE: implications for retinopathy. *Curr Eye Res. Apr*;28(4):277-84.
- 51 Marmor MF, 2005, The dilemma of hydroxychloroquine screening: new information from the multifocal ERG. *Am J Ophthalmol. Nov*;140(5):894-5
- 52 Marmor MF, 2012, Comparison of screening procedures in hydroxychloroquine toxicity. *Arch Ophthalmol. Apr*;130(4):461-9
- 53 Marmor MF, Hu J, 2014, Effect of disease stage on progression of hydroxychloroquine retinopathy. *JAMA Ophthalmol. Sep 1*;132(9):1105-12.
- 54 Marmor MF, Melles RB, 2014, Disparity between visual fields and optical coherence tomography in hydroxychloroquine retinopathy. *Ophthalmology. Jun*;121(6):1257-62.
- 55 Marmor MF, Carr RE, Easterbrook M, Farjo AA, Mieler WF; 2002, American Academy of Ophthalmology. Recommendations on screening for chloroquine and hydroxychloroquine retinopathy: a report by the American Academy of Ophthalmology. *Ophthalmology. Jul*;109(7):1377-82.
- 56 Marmor MF, Chien FY, Johnson MW, 2013, Value of red targets and pattern deviation plots in visual field screening for hydroxychloroquine retinopathy. *JAMA Ophthalmol. Apr*;131(4):476-80.
- 57 Marmor MF, Hood DC, Keating D, Kondo M, Seeliger MW, Miyake Y; International Society for Clinical Electrophysiology of Vision, 2003, Guidelines for basic multifocal electroretinography (mfERG). *Doc Ophthalmol. Mar*;106(2):105-15.
- 58 Marmor MF, Hu J, 2014, Effect of Disease Stage on Progression of Hydroxychloroquine Retinopathy. *JAMA Ophthalmol. 2014 Jun 12*. doi: 10.1001/jamaophthalmol. 1099
- 59 Marmor MF, Kellner U, Lai TY, Lyons JS, Mieler WF; American Academy of Ophthalmology, 2011, Revised recommendations on screening for chloroquine and hydroxychloroquine retinopathy. *Ophthalmology. Feb*;118(2):415-22

- 60 Marmor MF, Melles RB, 2014 Disparity between visual fields and optical coherence tomography in hydroxychloroquine retinopathy. *Ophthalmology*. Jun;121(6):1257-62
- 61 Maturi RK, Yu M, Weleber RG, 2004, Multifocal electroretinographic evaluation of long-term hydroxychloroquine users. *Arch Ophthalmol*. Jul;122(7):973-81.
- 62 Mavrikakis M, Papazoglou S, Sfikakis PP, Vaiopoulos G, Rougas K, 1996, Retinal toxicity in long term hydroxychloroquine treatment. *Ann Rheum Dis*. Mar;55(3):187-9.
- 63 McChesney EW, 1983, Animal toxicity and pharmacokinetics of hydroxychloroquine sulfate. *Am J Med*. Jul 18;75(1A):11-8
- 64 McChesney EW, Banks WF Jr, Fabian RJ, 1967, Tissue distribution of chloroquine, hydroxychloroquine, and desethylchloroquine in the rat. *Toxicol Appl Pharmacol*. May;10(3):501-13
- 65 Michaelides M, Stover NB, Francis PJ, 2011, Weleber RG, Retinal toxicity associated with hydroxychloroquine and chloroquine: risk factors, screening, and progression despite cessation of therapy. *Arch Ophthalmol*. Jan;129(1):30-9.
- 66 Mißner S, Kellner U. Comparison of different screening methods for chloroquine/hydroxychloroquine retinopathy: multifocal electroretinography, color vision, perimetry, ophthalmoscopy, and fluorescein angiography. *Graefes Arch Clin Exp Ophthalmol*. Mar;250(3):319-25
- 68 Morris SJ, Wasko MC, Antohe JL, Sartorius JA, Kirchner HL, Dancea S, Bili A, 2011, Hydroxychloroquine use associated with improvement in lipid profiles in rheumatoid arthritis patients. *Arthritis Care Res (Hoboken)*. Apr;63(4):530-4. doi: 10.1002/acr.20393.
- 69 Moschos MN, Moschos MM, Apostolopoulos M, Mallias JA, Bouros C, Theodossiadis GP, 2004, Assessing hydroxychloroquine toxicity by the multifocal ERG. *Doc Ophthalmol*. Jan;108(1):47-53.
- 70 Murdoch IE, Morris SS, 1998, Cousens SN. People and eyes: statistical approaches in ophthalmology. *Br J Ophthalmol*. Aug;82(8):971-3.

- 71 Nebbioso M, Grenga R, Karavitis P. 2009 Early detection of macular changes with multifocal ERG in patients on antimalarial drug therapy. *J Ocul Pharmacol Ther.* Jun;25(3):249-58.
- 72 Nika M, Blachley TS1, Edwards P2, Lee PP1, Stein JD, 2014, Regular Examinations for Toxic Maculopathy in Long-term Chloroquine or Hydroxychloroquine Users.*JAMA Ophthalmol.* Jun 26. doi: 10.1001/jamaophthalmol.2014.1720
- 73 Ochsendorf FR, 2010, Use of antimalarials in dermatology. *J Dtsch Dermatol Ges.* Oct;8(10):829-44; quiz 845
- 74 O'Dell JR, Haire CE, Erikson N, Drymalski W, Palmer W, Eckhoff PJ, Garwood V, Maloley P, Klassen LW, Wees S, Klein H, Moore GF, 1996.Treatment of rheumatoid arthritis with methotrexate alone, sulfasalazine and hydroxychloroquine, or a combination of all three medications. *N Engl J Med.* May 16;334(20):1287-91.
- 75 O'Dell JR, Leff R, Paulsen G, Haire C, Mallek J, Eckhoff PJ, 2002, Treatment of rheumatoid arthritis with methotrexate and hydroxychloroquine, methotrexate and sulfasalazine, or a combination of the three medications: results of a two-year, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum.* May;46(5):1164-70
- 76 Ono C, Yamada M, Tanaka M, 2003, Absorption, distribution and excretion of ¹⁴C-chloroquine after single oral administration in albino and pigmented rats: binding characteristics of chloroquine-related radioactivity to melanin in-vivo. *J Pharm Pharmacol.* Dec;55(12):1647-54.
- 77 Palma Sánchez D, Rubio Velazquez E, Soro Marín S, Reyes García R, 2013, Retinal toxicity due to antimalarials: frequency and risk factors. *Reumatol Clin.* Sep-Oct;9(5):259-62
- 78 Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R, 2008, Understanding and using sensitivity, specificity and predictive values. *Indian J Ophthalmol.* Jan-Feb;56(1):45-50.
- 79 Pasadhika S, Fishman GA, Choi D, Shahidi M, 2010, Selective thinning of the perifoveal inner retina as an early sign of hydroxychloroquine retinal toxicity. *Eye (Lond).* May;24(5):756-62; quiz 763

- 80 Payne JF, Hubbard GB 3rd, Aaberg TM Sr, Yan J, 2011, Clinical characteristics of hydroxychloroquine retinopathy. *Br J Ophthalmol.* Feb;95(2):245-50.
- 81 Penrose PJ, Tzekov RT, Sutter EE, Fu AD, Allen AW Jr, Fung WE, Oxford KW, 2003, Multifocal electroretinography evaluation for early detection of retinal dysfunction in patients taking hydroxychloroquine. *Retina.* Aug;23(4):503-12.
- 82 Raines MF, Bhargava SK, Rosen ES, 1989, The blood-retinal barrier in chloroquine retinopathy. *Invest Ophthalmol Vis Sci.* Aug;30(8):1726-31.
- 83 Rodriguez-Padilla JA, Hedges TR 3rd, Monson B, Srinivasan V, Wojtkowski M, Reichel E, Duker JS, Schuman JS, Fujimoto JG, 2007, High-speed ultra-high-resolution optical coherence tomography findings in hydroxychloroquine retinopathy. *Arch Ophthalmol.* Jun;125(6):775-80.
- 84 Rosenthal AR, Kolb H, Bergsma D, Huxsoll D, Hopkins JL, 1978, Chloroquine retinopathy in the rhesus monkey. *Invest Ophthalmol Vis Sci.* Dec;17(12):1158-75
- 85 Ruiz-Irastorza G, Ramos-Casals M, Brito-Zeron P, Khamashta MA, 2010 Clinical efficacy and side effects of antimalarials in systemic lupus erythematosus: a systematic review. *Ann Rheum Dis.* Jan;69(1):20-8. doi: 10.1136/ard.2008.101766.
- 86 Rutjes AW, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM, 2007, Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess.* Dec;11(50):iii, ix-51.
- 87 Sauer M, Hofkens J, and Enderlein J, 2011, Handbook of Fluorescence Spectroscopy and Imaging.-Chapter 1- Basic Principles of Fluorescence Spectroscopy. Copyright WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
- 88 Scherbel AL, Harrison JW, Atdjian M, 1958 Further observations on the use of 4-aminoquinoline compounds in patients with rheumatoid arthritis or related diseases. *Cleve Clin Q.* Apr;25(2):95-111

- 89 Scherbel AL. M.D, 1983, Use of Synthetic Antimalarial Drugs and Other Agents for Rheumatoid Arthritis: Historic and Therapeutic Perspectives. July 18, The American Journal of Medicine
- 90 Schlesinger P H, Krogstad D J, and Herwaldt B L, 1988, Antimalarial agents: mechanisms of action. Antimicrob Agents Chemother. Jun; 32(6): 793–798.
- 91 "Schmitz-Valckenberg S, Holz FG, Bird AC, Spaide RF, 2008, Fundus autofluorescence imaging: review and perspectives. Retina. 2008 Mar;28(3):385-409.
- 92 Shroyer, Noah F ; Lewis, Richard Alan ; Lupski, James R, 2001, Analysis of the ABCR (ABCA4) gene in 4-aminoquinoline retinopathy: is retinal toxicity by chloroquine and hydroxychloroquine related to Stargardt disease? American Journal of Ophthalmology, Vol.131(6), pp.761-766
- 93 Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken). May;64(5):625-39. doi: 10.1002/acr.21641.
- 94 So SC, Hedges TR, Schuman JS, Quireza ML, 2003, Evaluation of hydroxychloroquine retinopathy with multifocal electroretinography. Ophthalmic Surg Lasers Imaging. May-Jun;34(3):251-8.
- 95 Stepien KE, Han DP, Schell J, Godara P, Rha J, Carroll J, 2009, Spectral-domain optical coherence tomography and adaptive optics may detect hydroxychloroquine retinal toxicity before symptomatic vision loss. Trans Am Ophthalmol Soc. Dec;107:28-33.
- 96 Suarez-Almazor ME1, Belseck E, Shea B, Homik J, Wells G, Tugwell P, 2000, Antimalarials for treating rheumatoid arthritis.Cochrane Database Syst Rev.:(4):CD000959
- 97 Sutter E E, 2010, Noninvasive Testing Methods: Multifocal Electrophysiology. The Smith - Kettlewell Eye Research Institute, San Francisco, CA, USA Elsevier

- 98 Tett SE, Cutler D J, Day R O, Brown KF, 1989, Bioavailability of hydroxychloroquine tablets in healthy volunteers. *Br J Clin Pharmacol*. Jun; 27(6): 771–779.
- 99 Theodore Smith P, Jan P. Koniarek, Jackie Chan, Takayuki Nagasaki, Janet R. Sparrow, Langton Kevin, 2005, Autofluorescence Characteristics of Normal Foveas and Reconstruction of Foveal Autofluorescence from Limited Data Subsets. *Invest Ophthalmol Vis Sci*. Aug; 46(8): 2940–2946.
- 100 Tsang-A-Sjoe MW, Bultink IE, Voskuyl AE, 2014, Long-term evaluation of antimalarials in a Dutch SLE cohort: intolerance and other reasons for non-use. *Clin Exp Rheumatol*. Jan-Feb;32(1):95-100
- 101 Tzekov RT, Gerth C, Werner JS, 2004, Senescence of human multifocal electroretinogram components: a localized approach. *Graefes Arch Clin Exp Ophthalmol*. Jul;242(7):549-60.
- 102 Tzekov RT, Serrato A, Marmor MF, 2004, ERG findings in patients using hydroxychloroquine. *Doc Ophthalmol*. Jan;108(1):87-97.
- 103 Walvick MD, Walvick MP, Tongson E, Ngo CH, 2011, Hydroxychloroquine: lean body weight dosing. *Ophthalmology*. Oct;118(10):2100
- 104 Weiner A, Sandberg MA, Gaudio AR, Kini MM, Berson EL, 1991, Hydroxychloroquine retinopathy. *Am J Ophthalmol*. Nov 15;112(5):528-34.
- 105 Wilson JMG, Jungner G. 1968, Principles and practice of screening for disease (large pdf). WHO Chronicle Geneva:World Health Organization.22(11):473
- 106 Wolfe F, Marmor MF, 2010, Rates and predictors of hydroxychloroquine retinal toxicity in patients with rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Care Res (Hoboken)*. Jun;62(6):775-84.
- 107 World Health Organization. 2006 "BMI Classification" Global Database on Body Mass Index.

- 108 Xiaoyun MA, Dongyi HE, Linping HE. 2010, Assessing Chloroquine toxicity in RA patients using retinal nerve fibre layer thickness, multifocal electroretinography and visual field test. *Br j Ophthalmol*; 94: 1632-1636.
- 109 Yam JC, Kwok AK, 2006, Ocular toxicity of hydroxychloroquine. *Hong Kong Med J*. Aug;12(4):294-304.
- 110 Yaylali SA, Sadigov F, Erbil H, Ekinci A, 2013, Akcakaya AA, Chloroquine and hydroxychloroquine retinopathy-related risk factors in a Turkish cohort. *Int Ophthalmol*. Dec;33(6):627-34.