Dealing with perimetric variability in clinical glaucoma care

Junoy Montolio, Francisco Gerbert

DOI:
10.33612/diss.194778426

DEALING WITH PERIMETRIC VARIABILITY IN CLINICAL GLAUCOMA CARE

Francisco G. Junoy Montolio

COLOPHON

Cover design:	James Jardine | www.jamesjardine.nl
Layout:	James Jardine | www.jamesjardine.nl
Print:	Ridderprint | www.ridderprint.nl

The research described in this thesis was supported the University Medical Center Groningen, by a research grant from Stichting Nederlands Oogheelkundig Onderzoek (SNOO) and Uitzicht.

Publication of this thesis was financially supported by the University of Groningen, University Medical Center Groningen, Professor Mulder Stichting, Landelijke Stichting voor Blinden en Slechtzienden en Stichting Blindenhulp.

rijksuniversiteit
groningen

DEALING WITH PERIMETRIC VARIABILITY IN CLINICAL GLAUCOMA CARE

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 15 december 2021 om 11.00 uur

door

Francisco Gerbert Junoy Montolio

geboren op 30 juli 1985
te Emmen

Promotores

Prof. dr. N.M. Jansonius
Prof. dr. J.M.M. Hooymans

Beoordelingscommissie

Prof. dr. P. van Dijk
Prof. dr. D. Crabb
Prof. dr. H.J.M. Beckers

Cnt

tents

Introduction

Francisco G. Junoy Montolio

INTRODUCTION OF THE THESIS

Glaucoma and visual field are the three keywords in this thesis and are irrevocably connected with each other. To understand the connection between these keywords, some knowledge of the physiology of the eye is necessary.

Physiology of the eye

The eye is one of few sensory organs of the human body. Light that is reflected off objects meets the eye like this thesis in front of you. The first objective of the eye is to create a high quality beam of light on the retina, and the second part is to process this light beam to electrical signals that can be processed further in the brain.

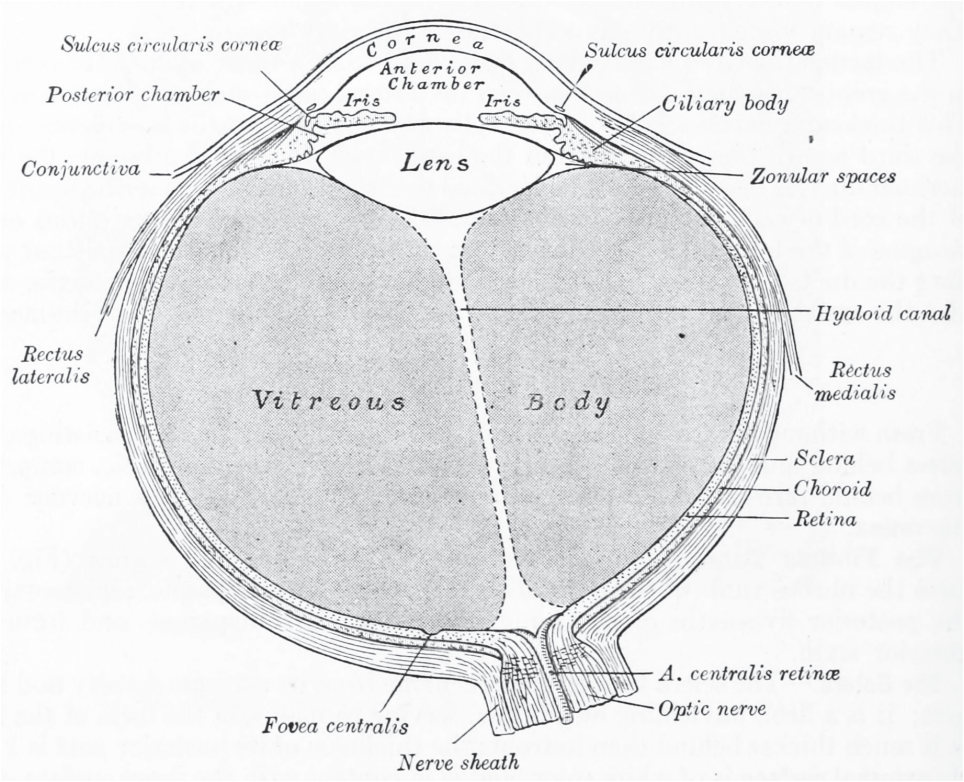


Figure 1. Anatomy of the eye.(Gray 1918)

When light is reflected off an (distant) object, these parallel beams first make contact with the **cornea** (Figure 1). The cornea accounts for two-thirds of the refractive power of the eye. The **crystalline lens** gives an additional refraction and has the ability to accommodate in case objects are in close proximity to the eye. Together they create a high quality image on one focus point on the retina. The space between the lens and cornea, called the **anterior chamber**, is filled with aqueous humor and provides a constant environment to ensure clarity and keeps the eye on tension. As will be discussed later, this fluid also plays an important role in glaucoma. The space behind the lens is filled with **vitreous humor**. This fluid consists of collagen and moisture and gives firmness and clarity to the eye. The importance of the vitreous humor is often underestimated, with the sole function as a medium in which light moves, but it also has other important functions that fall outside the scope of this thesis.(Ponsioen et al. 2010)

The refracted light is focused on the **neurosensory retina,** where the second goal of the eye starts. The light beam goes through the neurosensory retina consisting of many different layers before it is absorbed in the outermost layer of the retina, the photoreceptor layer (see Figure 2). There are two main types of photoreceptors: **rods** and **cones**. Rods perform best in dim light but cannot resolve colors, small details, or rapid changes in luminance. Cones, on the other hand, function best in bright light and are sensitive to small details and colors. A cascade of events within the photoreceptors leads to an electrical pulse that can be passed on to bipolar cells. **Bipolar cells** (illustrated within the inner nuclear layer, but not named in Figure 2.) transfer information from the photoreceptors to **ganglion cells**. The cell bodies convey the information - now coded in action potentials - through their axons. These axons start within the eye forming the optic nerve followed by the optical tract that lies within the brain and ends in the visual cortex at the back of the brain. Ganglion cells are the cells that are affected in glaucoma (see following section).

When an eye is focused on an object, the total surroundings that fall on the retina at one moment is called the visual field. The **visual field** (VF) can be divided into the central (fovea centralis) and peripheral part. The central VF has the highest density of cones, in contrast to the peripheral VF which has relatively more rods. Because of the different characteristics of cones and rods, we experience our VF as we do: distinguishing small details in the central VF compared to a blurry VF in the periphery.

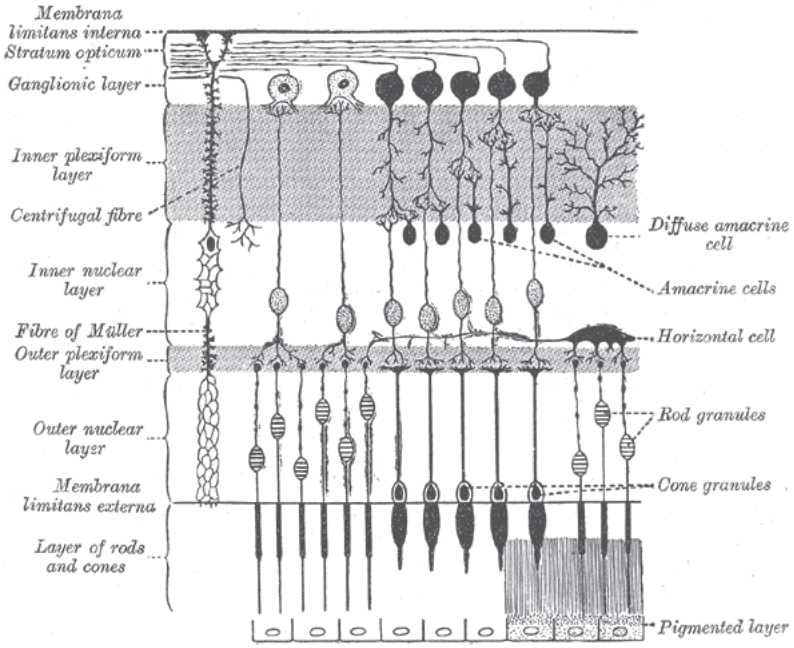


Figure 2. Anatomy of the retina.(Gray 1918)

Glaucoma

Glaucoma is a disease characterised by two irrevocably connected deviations from normal. First, there is a structural deviation: a decrease in the number of ganglion cells, more than should be expected due to ageing. Second, there is a deviation in function: a visual field defect that can only be explained by glaucoma and no other disease or structural damage.(Quigley 2011)

The loss of ganglion cells due to ageing starts even before we are born.(Ashwell & Waite 2004) Humans start with a maximum of about 4 million ganglion cells per eye at the end of the first trimester and this number decreases to about 1 million at the end of the second trimester. After that, there is a slower physiological decrease that has no influence on normal useful vision as an abundance number of ganglion cells exsist. A faster than expected loss however, as in the case of glaucoma, can cause asymptomatic visual field loss at onset and blindness as the most severe consequence. The location of damage that leads to ganglion cell loss and loss of its axons (the nerve fibers) lies at the beginning of the optic nerve. The nerve fibers pass through a fine-meshed collagenous

structure called **lamina cribrosa**. Pressure on this structure damages nerve fibers in an anterograde and retrograde manner. In addition to this mechanical cause of damage, impaired perfusion of the optic nerve head may play a role as well.

The main risk factor for the development of glaucoma is an increased intraocular pressure (IOP). Moreover, it is currently the only scientifically proven risk factor that can be addressed to cease further progression. Other risk factors are older age, myopia, positive family history, and ethnicity (particularly African descent).(Janssen et al. 2013, Marcus et al. 2011) Generally the intraocular pressure is increased by a disbalance of aqueous production in the ciliary body and the drainage of aqueous in the anterior chamber. This disbalance develops in different ways, and therefore glaucoma is subdivided into two main entities.

Primary Open Angle Glaucoma (POAG) is characterised by a higher than normal IOP due to an impaired outflow of aqueous at the level of the trabecular meshwork without known cause. POAG is a chronic, age-related disease characterised by a slow and often asymptomatic course. Most chapters in this thesis are based on patients with POAG. In **Secondary Open Angle Glaucoma** there is a clear underlying provoking factor that increases the IOP. Steroids are the most famous trigger known, but other medications can also increase the IOP. Intraocular inflammation (uveitis) or one or multiple intraocular interventions can also lead to secondary open angle glaucoma.(de Vries et al. 2016, Hoeksema et al. 2017)

Lowering the IOP is the (only) treatment to slow down or stop progression of POAG. The initial treatment is in most cases topical medication to lower production of intraocular fluid (carbonic anhydrase inhibitors, Ɪ3-blockers, and α2 adrenergic agonist) or to increase fluid drainage (prostaglandin analogs, α2 adrenergic agonist, and miotics). More than one medication may be given simultaneously. Laser treatment is possible as well, and is nowadays considered the first choice of treatment.(Gazzard et al. 2019) Different kinds of laser treatment have been developed over time, but currently the (most) used technique is the minimally invasive selective laser trabeculoplasty (SLT). (Latina et al. 1998) SLT modifies the trabecular meshwork leading to an increase in aqueous outflow. Other methods of increasing aqueous outflow are by performing surgery. So-called penetrating surgery, like the trabeculectomy and the Baerveldt glaucoma implant, are the most commonly used techniques today. Trabeculectomy, described in 1968 by Cairns,(Cairns 1968) is characterised by a kind of fistula through the sclera acting like an extra drainage point. The Baerveldt glaucoma implant also creates an extra drainage point, by redirecting the intraocular fluid from the anterior chamber through a tube to a plate in the orbital space. More recently developed less-

invasive glaucoma surgery techniques are potentially safer than the trabeculectomy and Baerveldt procedure.(Agrawal & Bradshaw 2018) Their efficacy, however, still has to be determined.

Primary Angle Closure Glaucoma (PACG) is a different entity where the underlying pathology is an anatomical abnormality.(Jonas et al. 2017) Three mechanisms can cause the anatomical abnormality: (1) a forward bulging of the anterior lens pole, (2) an enlarged area of contact between the posterior iris and the lens surface, and (3) an abnormal insertion of the iris root. When this abnormality is combined with an age related physiological increase in lens thickness, the anterior chamber angle closes with obstruction of the trabecular meshwork. Apart from age, risk factors include an east Asian ethnicity, female gender, and hyperopia. In **Secondary Angle Closure Glaucoma (SACG)** another disease is causing the obstruction. Uveitis and neovascularization of the iris (in the context of ischaemic retinopathies) can pull the iris into the angle creating a blockage of the trabecular meshwork. Therapy in both primary and secondary angle closure glaucoma is designed to reverse the mechanism causing the obstruction. In PACG a peripheral iridotomy performed by a laser brings the production and drainage closer together, creating an escape route. The preferred treatment to obtain a more long term result is a (clear) lens extraction, as an artificial IOL is much thinner compared to the original lens.(Azuara-Blanco et al. 2016, Chan et al. 2018, Lam et al. 2007) Additional treatment of uveitis or neovascularization, or any other underlying pathology, is needed in case of SACG.

Visual field testing

Visual field testing, and perimetry in particular, has a long history. The first mention of visual field dates back to the second century BCE.(Schiefer et al. 2005) Mariotte was the first to describe the physiologic blind spot in 1666, shortly after Porta introduced campimetry in 1593, testing the central part of the visual field using a flat screen. Perimetry (testing the visual field using a curved device) was introduced first by Von Graefe in 1856.(Von Graefe 1856) Many modifications emerged, but the biggest change was made by Goldmann in 1945.(Goldmann 1945) He developed a bowl perimeter that is, in some cases, still used to date, called kinetic perimetry because stimuli of fixed size and luminance move in a centripetal direction. This form of perimetry gave the opportunity to uniformly test different diseases like glaucoma, although a skilled and experienced perimetrist was needed to get accurate and reproducible results. The other main form of perimetry is static perimetry. Here, the stimuli have a fixed location (and size) but vary in luminance. Standard automated perimetry, a form of static perimetry, emerged in the 1970s and became the new standard. This will probably remain this way

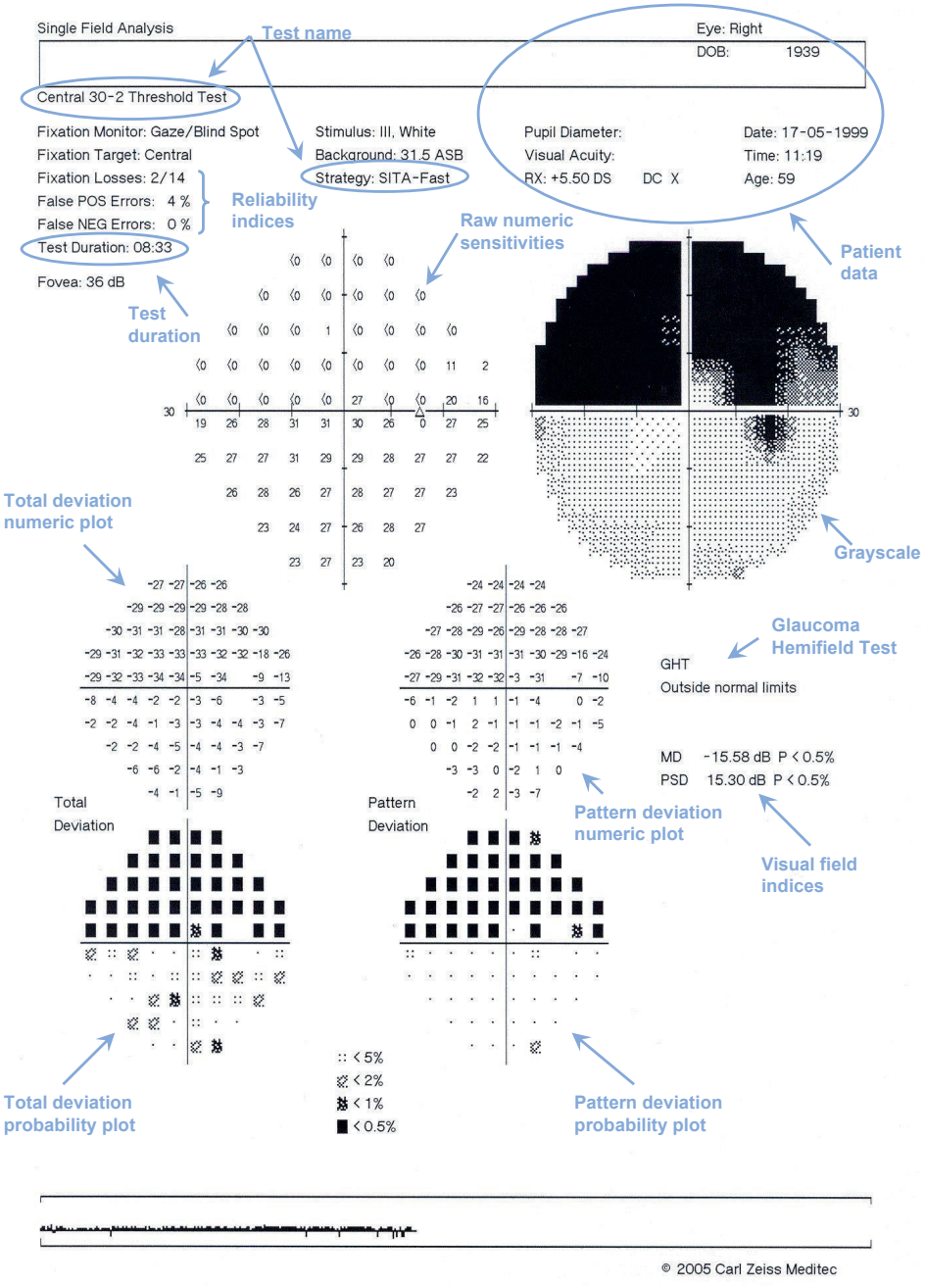


Figure 3. Print out of a visual field affected by glaucoma

for many years to come, in particular for glaucoma care.(Wall 2004) While searching for a method that is more efficient in time and has a lower test-retest variability, many alternative test methods were developed. Examples are, amongst others, ring (high-pass resolution) perimetry, frequency doubling technology perimetry, and multifocal visual evoked potential perimetry.

Within static perimetry, the Swedish Interactive Threshold Algorithm (SITA) was developed as an evolution of the time consuming Full Threshold (FT) strategy in the 1990s. Limited patient acceptance, fatigue, and test-retest variability were the issues SITA should resolve. In contrast to the FT strategy, where the sensitivity is measured in each test location (54 or 76, depending on the test grid chosen) using a staircase method, the SITA strategy makes a mathematical re-calculation of the best probability of seeing after every response of a test subject. This strategy resulted in a three-fold time reduction, particularly in healthy subjects or patients with early glaucoma. Test-retest variability, however, did not decrease at all with SITA, and this variability forms the central theme of this thesis (Figure 3: an example of a SITA fast print out of a patient with glaucoma).

The Groningen Longitudinal Glaucoma Study (GLGS)

In the previous century, a large glaucoma cohort study called the Groningen Longitudinal Glaucoma Study (GLGS) started in Groningen.(Heeg et al. 2006) Although the original objectives were (mostly) answered throughout the years, new objectives arose. This is usually the case in research, as the more knowledge you gain, the more questions arise. In addition, cohorts of glaucoma patients that are followed for more than a decade (even almost two decades) are rare, and form a wealthy source of clinical knowledge. The original objectives of the GLGS were related to case finding (the detection of conversion to glaucoma in glaucoma suspects) and progression detection (the detection of change in patients with established glaucoma). In particular, the cost-effectiveness of the frequency doubling perimeter (FDT) and laser polarimetry (GDx) was evaluated. For case finding, FDT outperformed conventional perimetry (standard automated perimetry [SAP], in our clinic the Humphrey Field Analyzer with 30-2 grid and SITA strategy) and the GDx. For progression detection, perimetry remained the primary technique. In our current clinical setting, FDT is used to screen glaucoma suspects (i.e., patients with an elevated IOP, a positive family history of glaucoma, and/or an optic nerve head appearance suggesting glaucoma).(Müskens et al. 2008, Stoutenbeek et al. 2010) FDT is also used in glaucoma patients that are not able to perform SAP.(Wesselink & Jansonius 2017) This concerns mostly elderly patients.

After answering the original objectives, the data collection of the GLGS continued and now we follow all originally included glaucoma patients in a longitudinal fashion, to quantify the rate of progression for SAP in patients treated in our outpatient department, and to detect risk factors for progression.(Wesselink 2017) As glaucoma is an indolent disease, longitudinal follow-up is indispensable to answer clinically relevant research questions and to benchmark the quality of clinical care. These longitudinal data underlie the research described in this thesis.

Variability in perimetry

The analyses of glaucoma progression relies almost completely on perimetry. An apparent shortcoming of this test is the relatively high test-retest variability. As a consequence, a visual field examination has to be performed twice or even thrice to establish a useful baseline, and, similarly, two or three tests are needed to falsify or confirm disease progression. A decrease in variability of 20% might allow a clinician to detect progression one visit earlier and 40% less variability would accelerate the detection of progression up to one year.(Turpin & McKendrick 2011) The variability seems to have a multifactorial cause: extrinsic patient factors like inattentiveness, misunderstanding of the task, fixation errors or a distracting environment are preventable causes which technicians should take into account. An important intrinsic factor causing short-term variability seems to be glaucoma itself, rather than the technique used to measure the visual field. Holmin & Krakau(Holmin & Krakau 1979) described glaucoma as an independent variability factor in one of the first computerized perimeters in already 1979; Heijl et al.(Heijl et al. 1989) described it in the Humphrey full threshold algorithm in 1989 and Bengtsson et al.(Bengtsson & Heijl 1998, Heijl et al. 1989) in the SITA strategy in 1998. It is hypothesized that a lower ganglion cell density causes variability by an increase in Ricco's area. Ricco's area is the area where threshold intensity multiplied by the area equals a constant (Ricco's Law). The conversion of partial spatial summation when the stimulus is larger than Ricco's area to total spatial summation seems to be the rationale behind this hypothesis.(Redmond et al. 2010) This is the case in glaucomatous affected areas, but also in the healthy peripheral visual field(Gardiner 2018) and in eyes with optic neuritis.(Henson et al. 2000)

AIM OF THESIS

Clinicians who provide glaucoma care, have to deal with the test-retest variability of perimetry. Considering that many different variables may affect the outcome, this can be challenging. Therefore, the aim of this thesis is to give clinicians more insight into the factors that influence perimetry, with the ultimate aim to make the interpretation of test results more valuable.

OUTLINE OF THESIS

Chapter II: The efficiency of SAP has increased significantly with the introduction of the SITA. However, SITA has a remarkable property: the test time increases disproportionately if visual field defects are present. This chapter describes the influence of disease stage on test time, as well as a new method to avoid this increase in test time in patients with a visual field defect.

Chapter III: An ideal test gives a reliable outcome; that is, an outcome that is only influenced by the disease studied, in this case glaucoma. If this were true, a decline in visual field would indicate inadequate treatment and a stable visual field sufficient treatment. A visual field test, however, is a subjective test and many known and unknown factors influence the test result. This chapter points out several important factors and gives an indication to what extent they influence the test result. The factors studied are the experience of the technician who guides the patient through the test, time of day, and season. In addition, we study to what extent the test reliability indices given by the perimeter are a true measure of reliability.

Chapter IV: The optics of the eye may also influence perimetric test results. Perimetric test results are compared to the normative data obtained from surgically untouched eyes.(Bengtsson & Heijl 1999) As such, it is not self-evident that eyes who underwent cataract surgery have the same perimetric sensitivity as phakic eyes with a clear lens. In addition, different types of intraocular lenses (IOLs) are implanted nowadays, with different optical characteristics. This chapter investigates the influence of mono- and multifocal IOLs on perimetry and the implications for testing glaucoma patients.

Chapter V: In this chapter we apply our knowledge of perimetry to assess glaucomatous visual field progression, before and after glaucoma surgery. Filtering or drainage surgery can be performed if additional intraocular pressure lowering is needed to slow-down progression. Patients, however, often complain of deterioration of visual

acuity and/or their visual field after surgery, the "cost" of surgery. The purpose of this chapter is to compare the cost of surgery to the benefit of arresting the progression of glaucoma. An additional meta-analysis was performed to strengthen our results.

Chapter VI: The previous chapters aim to explain many factors that influence perimetry, performed using the default method for thresholding in standard automated perimetry: the staircase procedure. Although it is important to investigate the contribution of all these factors using a single method, the method itself could also be a contributing factor that influences the outcome. Therefore, we compare perimetric variability between the gold standard and an alternative thresholding method, the Continuous Light Increment perimetry (CLIP) strategy.(International Perimetric Society. Meeting, Wall & Mills 2001)

Chapter VII: A striking issue in perimetry is that the visual field may still look normal whereas the optic nerve head is already damaged. One explanation for this apparent discrepancy between function and structure is that the ganglion cells density is highest in the macular area, whereas, in standard automated perimetry (SAP), the macular area contains only a few test locations.(De Moraes et al. 2018, Schiefer et al. 2010) As such, centrally located glaucomatous damage may result in normal or near normal perimetric test results. In this chapter we explore an alternative explanation for this apparent discrepancy: loss of lateral inhibition.

REFERENCES

Agrawal P & Bradshaw SE (2018): Systematic literature review of clinical and economic outcomes of micro-invasive glaucoma surgery (MIGS) in primary open-angle glaucoma. Ophthalmol Ther. 7: 49–73.

Ashwell KWS & Waite PME (2004): Development of the peripheral nervous system. The Human Nervous System. Amsterdam: Elsevier/Academic Press. 95–110.

Azuara-Blanco A, Burr J, Ramsay C, Cooper D, Foster PJ, Friedman DS, Scotland G, Javanbakht M, Cochrane C, Norrie J & EAGLE study group. (2016): Effectiveness of early lens extraction for the treatment of primary angle-closure glaucoma (EAGLE): a randomised controlled trial. Lancet. 388: 1389–1397.

Bengtsson B & Heijl A (1998): Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. Acta Ophthalmol Scand. 76: 268–272.

Bengtsson B & Heijl A (1999): Inter-subject variability and normal limits of the SITA Standard, SITA Fast, and the Humphrey Full Threshold computerized perimetry strategies, SITA STATPAC. Acta Ophthalmol Scand. 77: 125–129.

Cairns JE (1968): Trabeculectomy. Preliminary report of a new method. Am J Ophthalmol. 66: 673–679.

Chan PP, Pang JC & Tham CC (2018): Acute primary angle closure–treatment strategies, evidences and economical considerations. Eye. 33: 110–119.

De Moraes CG, Sun A, Jarukasetphon R, Rajshekhar R, Shi L, Blumberg DM, Liebmann JM, Ritch R & Hood DC (2018): Association of macular visual field measurements with glaucoma staging systems. JAMA Ophthalmol. 137: 139-145.

de Vries MM, Müskens RPHM, Renardel de Lavalette VW, Hooymans JMM & Jansonius NM (2016): Glaucoma drainage device surgery after vitreoretinal surgery: incidence and risk factors. Acta Ophthalmol. 94: 135–139.

Gardiner SK (2018): Differences in the relation between perimetric sensitivity and variability between locations across the visual field. Invest Ophthalmol Vis Sci. 59: 3667–3674.

Gazzard G, Konstantakopoulou E, Garway-Heath D, Garg A, Vickerstaff V, Hunter R, Ambler G, Bunce C, Wormald R, Nathwani N, Barton K, Rubin G, Morris S & Buszewicz M (2019): Selective laser trabeculoplasty versus drops for newly diagnosed ocular hypertension and glaucoma: the LiGHT RCT. Health Technol Assess. 23: 1–102.

Goldmann H (1945): Grundlagen exakter Perimetrie. Ophthalmologica. 109: 57–70.

Graefe A (1856): Ueber die Untersuchung des Gesichtsfeldes bei amblyopischen Affectionen. Albrecht von Graefes Archiv für Ophthalmologie. 2: 258–298.

Gray H (1918): Anatomy of the Human Body.

Heeg GP (2006): The Groningen longitudinal glaucoma study: baseline characteristics of the included cohort, screening performance of frequency doubling perimetry and scanning laser polarimetry, and the prediction of glaucomatous visual field loss (Doctoral dissertation).

Heijl A, Lindgren A & Lindgren G (1989): Test-retest variability in glaucomatous visual fields. Am J of Ophthalmol. 108: 130–135.

Henson DB, Chaudry S, Artes PH, Faragher EB & Ansons A (2000): Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. Invest Ophthalmol Vis Sci. 41: 417–421.

Hoeksema L, Jansonius NM & Los LI (2017): Risk factors for secondary glaucoma in herpetic anterior uveitis. Am J Ophthalmol 181: 55–60.

Holmin C & Krakau CE (1979): Variability of glaucomatous visual field defects in computerized perimetry. Albrecht Von Graefes Arch Klin Exp Ophthalmol. 210: 235–250.

Wall M & Mills RP (2001): Perimetry Update 2000/2001. International Perimetric Society. Meeting, Kugler Publications.

Janssen SF, Gorgels TGMF, Ramdas WD, Klaver CCW, van Duijn CM, Jansonius NM & Bergen AAB (2013): The vast complexity of primary open angle glaucoma: disease genes, risks, molecular mechanisms and pathobiology. Prog Retin Eye Res. 37: 31–67.

Jonas JB, Aung T, Bourne RR, Bron AM, Ritch R & Panda-Jonas S (2017): Glaucoma. Lancet. 390: 2183–2193.

Lam DSC, Tham CCY, Lai JSM & Leung DYL (2007): Current approaches to the management of acute primary angle closure. Curr Opin Ophthalmol. 18: 146–151.

Latina MA, Sibayan SA, Shin DH, Noecker RJ & Marcellino G (1998): Q-switched 532-nm Nd:YAG laser trabeculoplasty (selective laser trabeculoplasty): a multicenter, pilot, clinical study. Ophthalmology. 105: 2082–8; discussion 2089–90.

Marcus MW, de Vries MM, Junoy Montolio FG & Jansonius NM (2011): Myopia as a risk factor for open-angle glaucoma: a systematic review and meta-analysis. Ophthalmology. 118: 1989–1994.e2.

Müskens RPHM (2008): Some diagnostic and therapeutic controversies in glaucoma addressed. (Doctoral dissertation)

Ponsioen TL, Hooymans JMM & Los LI (2010): Remodelling of the human vitreous and vitreoretinal interface--a dynamic process. Prog Retin Eye Res. 29: 580–595.

Quigley HA (2011): Glaucoma. Lancet. 377: 1367–1377.

Redmond T, Garway-Heath DF, Zlatkova MB & Anderson RS (2010): Sensitivity loss in early glaucoma can be mapped to an enlargement of the area of complete spatial summation. Invest Ophthalmol Vis Sci. 51: 6540-6548.

Schiefer U, Papageorgiou E, Sample PA, Pascual JP, Selig B, Krapp E & Paetzold J (2010): Spatial pattern of glaucomatous visual field loss obtained with regionally condensed stimulus arrangements. Invest Ophthalmol Vis Sci. 51: 5685–5689.

Schiefer U, Pätzold J & Dannheim F (2005): Conventional perimetry I: introduction--basics. Ophthalmologe. 102: 627-44; quiz 645-6.

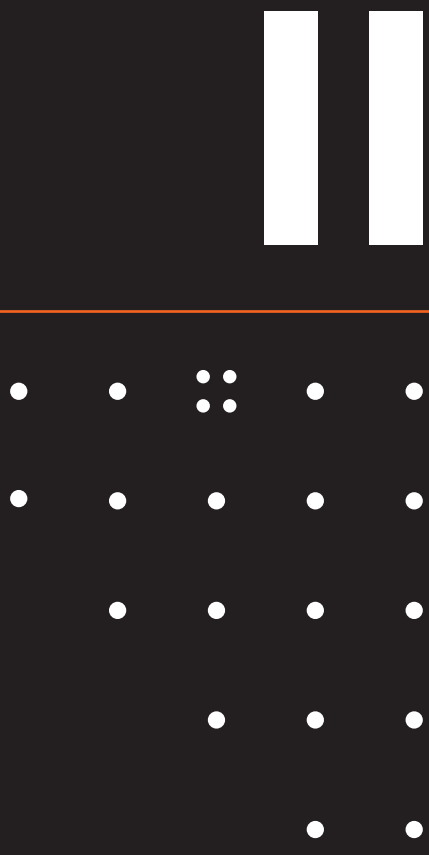
Stoutenbeek R (2010): Population Based Glaucoma Screening. (Doctoral dissertation).

Turpin A & McKendrick AM (2011): What reduction in standard automated perimetry variability would improve the detection of visual field progression? Invest Ophthalmol Vis Sci. 3237.

Wall M (2004): What's New in Perimetry. J Neuroophthalmol. 24: 46–55.

Wesselink C (2017): Glaucoma care optimised in an ageing population (Doctoral dissertation).

Wesselink C & Jansonius NM (2017): Glaucoma progression detection with frequency doubling technology (FDT) compared to standard automated perimetry (SAP) in the Groningen Longitudinal Glaucoma Study. Ophthalmic Physiol Opt. 37: 594–601.



Persistence, spatial distribution and implications for progression detection of blind parts of the visual field in glaucoma:

A clinical cohort study

Francisco G. Junoy Montolio*, Christiaan Wesselink*, Nomdo M. Jansonius

PLoS One 2012;7(7):e41211

** = contributed equally to this chapter*

ABSTRACT

Purpose: Visual field testing is an essential part of glaucoma care. It is hampered by variability related to the disease itself, response errors and fatigue. In glaucoma, blind parts of the visual field contribute to the diagnosis but - once established – not to progression detection; they only increase testing time. The aims of this study were to describe the persistence and spatial distribution of blind test locations in standard automated perimetry in glaucoma and to explore how the omission of presumed blind test locations would affect progression detection.

Methods: Data from 221 eyes of 221 patients from a cohort study with the Humphrey Field Analyzer with 30-2 grid were used. Patients were stratified according to baseline mean deviation (MD) in six strata of 5 dB width each. The probability of measuring a ≤0 dB sensitivity as a function of the number of preceding consecutive ≤0 dB values in the same test location was calculated. The effect of assuming 0 dB based on the ≤0 dB history in the same location on MD was determined.

Results: For one, two, three and four consecutive <0 dB sensitivities in the same test location in a series of baseline tests, the probabilities to observe <0 dB again in the concerning test location in a follow-up test were 76, 86, 88 and 90%, respectively. For <10 dB, the probabilities were 88, 95, 97 and 98%, respectively. Median (interquartile range) percentages of test locations with three consecutive <0 dB sensitivities were 0(0-0), 0(0-2), 4(0-9), 17(8-27), 27(20-40) and 60(50-70)% for the six MD strata. Similar percentages were found for a subset of test locations within 10 degree eccentricity (P>0.1 for all strata). Omitting test locations with three consecutive <0 dB sensitivities at baseline did not affect the performance of the MD-based Nonparametric Progression Analysis progression detection algorithm.

Conclusion: Test locations that have been shown to be reproducibly blind tend to display a reasonable blindness persistence and do no longer contribute to progression detection. There is no clinically useful universal MD cut-off value beyond which the testing can be limited to 10 degree eccentricity.

INTRODUCTION

Glaucoma is a progressive disease that may cause irreversible blindness. Monitoring of the disease with perimetry is an essential part of glaucoma care, unless patients have a short life expectancy and little glaucomatous damage. Variability hampers the use of perimetry in detecting small changes in visual function. In glaucoma, variability is presumably related to response errors, fatigue effects(Bengtsson & Heijl 1998, Hudson et al. 1994) and a flatter frequency-of-seeing curve in regions with a reduced sensitivity. (Chauhan et al. 1993, Wall 2004) The development of the Swedish Interactive Threshold Algorithm (SITA) strategies for the Humphrey Field Analyzer (HFA) has partially resolved the fatigue issue by reducing the test time.(Bengtsson & Heijl 1998)

SITA reduces the test time, amongst others, by predicting the sensitivity in a test location from the sensitivity in neighboring test locations and by incorporating general knowledge on glaucomatous visual field (VF) patterns. However, SITA ignores an obvious other source of prior knowledge, being the previous test result. The use of the previous test result can reduce test time(Frankhauser et al. 1977, Schiefer et al. 2009, Turpin et al. 2007) and test-retest variability.(Schiefer et al. 2009, Turpin et al. 2007) To illustrate this, for a typical glaucomatous visual field, that is, a blind superior hemifield together with an intact inferior hemifield, the test time of SITA is about 1.5 times longer than for a normal field. Hence, to establish blindness in a test location takes twice as long as establishing a normal sensitivity – and thus a 33% test-time reduction should be possible by incorporating information from previous tests. This is in agreement with earlier findings.(Schiefer et al. 2009) To go one step further, if the superior hemifield would have been unresponsive on several consecutive occasions, it makes no sense to test it again: only the inferior hemifield needs to be tested to monitor the eye. Hence, a 67% test-time reduction would ultimately be possible in this case.

The aims of this study were (1) to describe the persistence and spatial distribution of blind test locations in standard automated perimetry in glaucoma and (2) to explore how the omission of presumed blind test locations would affect progression detection. For the first aim, we determined the probability to observe a sensitivity below a certain value as a function of the number of preceding consecutive sensitivities below that value in the concerning test location. This was evaluated for <0, <5, <10 and <20 dB. The value <0 dB corresponds to the maximum stimulus intensity of the HFA perimeter; the values <5 and <10 dB approximately to the maximum stimulus intensities of the Octopus and Oculus perimeters, respectively. Subsequently, we compared the percentages of blind test locations between the regular standard automated perimetry 30-2 grid (with test locations up to 30 degree eccentricity) and the subset of test locations falling within

the 10-2 grid (up to 10 degree eccentricity), as a function of disease stage as defined by the mean deviation (MD). The aim here was to determine a clinically useful MD cut-off value for preferring 10-2 testing over 30-2 testing in advanced glaucoma. After all, although glaucoma sometimes starts close to fixation,(Hood et al. 2011, Schiefer et al. 2010) it is conceptually a disease affecting the peripheral visual field first and thus a transition from 30-2 to 10-2 testing would be the easiest way to avoid uninformative testing of unresponsive parts of the visual field in advanced disease. For the second aim, we studied the performance of an MD-based progression detection algorithm with and without assuming blind test locations as established at baseline to be blind in all follow-up fields.

METHODS

Ethics statement

The study protocol was approved by the ethics board of the University Medical Center Groningen. This board approved that for the current study no informed consent had to be obtained because the study comprised a retrospective anonymous analysis of visual field data collected during regular glaucoma care. To ensure a proper glaucoma diagnosis of the included patients, we limited the study population of this study to glaucoma patients that had been included in the Groningen Longitudinal Glaucoma Study (GLGS) in the past. In the GLGS, all glaucoma patients and glaucoma suspects who visited the glaucoma outpatient service of the University Medical Center Groningen between July 1, 2000, and June 30, 2001, and who provided informed consent were included in an observational study with conventional perimetry, frequency-doubling perimetry (FDT; Carl Zeiss Meditec AG, Jena, Germany) and laser polarimetry (GDx; Laser Diagnostic Technologies, San Diego, California, USA). Patients received written information at home at least two weeks before their regular care visit that was flagged as the baseline visit of the study. The receipt of the information and agreement to participate was checked verbally during the concerning visit. The aim of the study was explained; participation was voluntary and participation could be stopped also after having agreed to participate. The study essentially comprised the collection of regular care data obtained during regular visits and an additional FDT and GDx test embedded in a regular visit. FDT and GDx are non-invasive diagnostic tests with a very limited additional burden and no additional risk for the patient. The protocol of the original GLGS was approved by the department of Medical Technology Assessment of the University of Groningen. The original health technology assessment research

question was if it was possible to replace, in glaucoma patients and/or glaucoma suspects, the lengthy and cumbersome conventional perimetry by FDT and/or GDx. The study followed the tenets of the Declaration of Helsinki.

Study population

Details of the GLGS have been described earlier.(Heeg et al. 2005, Wesselink et al. 2009) In short, after the initial health technology assessment study described above, we continued performing conventional perimetry in glaucoma patients and moved to FDT/GDx in glaucoma suspects in our regular care. The GLGS continued as an ongoing anonymous gathering of all information from glaucoma patients and glaucoma suspects obtained during regular care. For the present study, we used data from a subpopulation of the GLGS cohort: patients had to have (1) glaucoma at baseline (for criteria see below) and (2) at least four (five with discarded learning test) standard automated perimetry tests (HFA; Carl Zeiss Meditec Inc., Dublin, CA).

Perimetry

Perimetry was performed using the HFA 30-2 SITA fast strategy. For glaucoma, two consecutive, reliable tests had to have defects according to previously published criteria.(Heeg et al. 2005, Wesselink et al. 2009) For being reproducible, defects had to be in the same hemifield and at least one depressed test point of these defects had to have exactly the same location on both tests. Moreover, defects had to be compatible with glaucoma and without any other explanation (for example, cataract, macular degeneration or lesions of the central visual pathways). Prior to these two tests, another test had to be made and this test was excluded to reduce the influence of learning. During the follow-up period, perimetry was performed at a frequency of one test per year. In case of suspected progression or unreliable test results, clinicians could increase the frequency of testing. This was a subjective decision; no formal tools or rules were given (observational study design).

Data analysis

One eye per patient was included. If both eyes met the above-described criteria, one eye was chosen randomly. For anatomical representation, all left-eye threshold data were converted to a right-eye format. Thresholds representing the blind-spot were excluded from the analysis, leaving 74 tests locations for analysis.

Persistence of blindness

For this analysis we only included patients who (1) performed at least eight tests and (2) had at least one test location showing a <0 dB sensitivity on four consecutive

baseline tests (most stringent criterion for blindness). We defined four subgroups of test locations, based on the first four tests and named VF4<0, VF3-4<0, VF2-4<0 and VF1-4<0. A test location VF4<0 had to have a sensitivity of <0 dB in the fourth visual field test. A test location VF3-4<0 had to have a sensitivity of <0 dB in both the third and the fourth test, and so on. For VF4<0, the sensitivity of the test location in the third test may or may not be <0 dB. Hence, VF3-4<0 is a subset of VF4<0, and so on. We took the fourth test as a reference in order to be able to vary the number of baseline tests without the need of changing the selection of the four follow-up tests, which were the fifth to eighth test.

For all test locations with a sensitivity of <0 dB in the fourth test, we analyzed the corresponding sensitivities in the four follow-up tests. Outcome measures were (1) the percentage of follow-up tests showing a sensitivity of <0 dB and (2) the mean sensitivity. Here, test locations with <0 dB were set at -2 dB. This is the arbitrary interpretation of <0 dB as chosen by the manufacturer. For patients, the difference between 0 dB and <0 dB implies seeing the maximum light stimulus of the perimeter (0 dB) or not (<0 dB).

Test locations within a single subject cannot be considered independent. Therefore, to avoid that a few patients with many blind test locations would dominate the results, we first determined the averages and corresponding standard deviations of the outcome measures within each patient for each subgroup of test locations (VF4<0, VF3-4<0, VF2-4<0 and VF1-4<0). Subsequently, the averages coming from all patients were presented using nonparametric descriptive statistics and the standard deviations were averaged over all patients and presented as the "mean within-patient standard deviation".

Table 1 gives an example of two patients as represented in the database. These patients are present in the VF4<0 subgroup with two and eight test locations, respectively. The first patient is also present in the VF3-4<0, VF2-4<0 and VF1-4<0 subgroups, with one test location. The second patient is present in these subgroups with four, one and one test locations, respectively. For the first patient, blindness persistence was 25% for the VF4<0 subgroup and 50% for the VF3-4<0, VF2-4<0 and VF1-4<0 subgroups. For the second patient, this was 53, 69, 75 and 75% for the VF4<0, VF3-4<0, VF2-4<0 and VF1-4<0 subgroups, respectively.

The analyzes were repeated with blindness of a test location defined as a sensitivity of <5, <10 and <20 dB instead of <0 dB. The influence of the number of consecutive tests (1, 2, 3 or 4) showing blindness in a test location and the definition of blindness (<0,

<5, <10 or <20 dB) on the persistence of blindness was analyzed with ANOVA, with the persistence of blindness (average percentage of follow-up tests showing blindness in the concerning test locations) as the dependent variable.

Table 1. Example of two patients as represented in the database, with two and eight test locations with a sensitivity of <0 dB in the fourth visual field, respectively

Patient	VF1 (dB)	VF2 (dB)	VF3 (dB)	VF4 (dB)	VF5 (dB)	VF6 (dB)	VF7 (dB)	VF8 (dB)	Position on VF	% <0 dB	Mean (dB)
1	4	4	6	<0	0	0	0	11	4	0	2.75
1	<0	<0	<0	<0	<0	2	<0	18	9	50	4.00
2	12	3	11	<0	<0	<0	13	10	1	50	4.75
2	20	0	5	<0	<0	<0	9	0	2	50	1.25
2	<0	<0	<0	<0	<0	3	<0	<0	19	75	-0.75
2	15	12	<0	<0	11	<0	<0	4	20	50	2.75
2	20	2	<0	<0	<0	1	<0	<0	30	75	-1.25
2	21	<0	6	<0	16	<0	<0	4	31	50	4.00
2	26	12	4	<0	10	4	5	0	33	0	4.75
2	<0	3	<0	<0	<0	10	<0	<0	35	75	1.00

VF = visual field; columns VF1-VF4 refer to baseline, VF5-VF8 to follow-up; last two columns depict the data analysis as applied to the follow-up data (for details see text).

Spatial distribution of perimetrically blind test locations

For this analysis, we included all patients who performed at least four tests. Patients were stratified according to baseline MD in six strata, being above -5 dB, from -5 to -10 dB, -10 to -15 dB, -15 to -20 dB, -20 to -25 dB and beyond -25 dB. We plotted the test locations considered blind based on their sensitivity history and calculated the percentages of these test locations, for all test locations of the 30-2 grid and for a subset laying within the 10-2 area. Percentages were compared with a nonparametric paired test (Wilcoxon).

A commonly used progression detection algorithm, the Glaucoma Progression Analysis (GPA),(Leske et al. 1999) has its own built-in criterion for blindness: a cross on the printout indicates that the test location is 'out of range' and not used for

progression detection by the software. We compared – for all six MD strata - the spatial distributions and percentages of test locations flagged as 'out of range' by GPA with that of test locations considered blind based on their sensitivity history. Percentages were compared using a nonparametric paired test (Wilcoxon).

Influence of assuming blindness on progression detection

If the sensitivity of a test location has been below a certain value on a number of consecutive tests, it might be an efficient approach to consider such a test location blind in all future tests – in glaucoma – without actually retesting it. This might result in a (slight) underestimation of the MD and thus might affect MD-based progression detection algorithms. To determine the influence of this approach on clinical decision making, we classified all included eyes as stable or progressing according to the MD-based Non-parametric Progression Analysis algorithm (NPA), with progression defined as at least possible progression at the end of the follow-up (NPA is based on a nonparametric ranking of MD values; for possible progression, the MDs of the last two tests have to be lower than the lower MD of two baseline tests).(Wesselink et al. 2009) Subsequently, we repeated this after assuming test locations to be perimetrically blind based on their sensitivity history. Here, we excluded test locations from the analysis if they were blind on the first three tests, according to four different definitions of blindness: <0, <5, <10 and <20 dB. For all four definitions, both classifications were compared with a McNemar test. Because the MD is an average weighed to test location eccentricity, and the weigh factors are unpublished, we applied the NPA criterion to the eccentricity-uncorrected average sensitivity of all test locations (mean sensitivity).

Calculations and statistical analyses were performed using SPSS Statistics 18.0 (SPSS Inc., Chicago, IL); the ANOVA was performed using MrF (http://psy.otago.ac.nz/miller/).

RESULTS

Table 2 shows the patient characteristics. Two-hundred-twenty-one patients were included of which 53 performed at least eight tests and had at least one test location showing a <0 dB sensitivity on four consecutive baseline tests. The average follow-up durations were 6.4 and 6.9 years, respectively, with median MD values at baseline of -7.0 and -14.7 dB.

Table 2. Patient characteristics for all included 221 eyes of 221 patients and for the subset of 53 eyes of 53 patients with at least eight visual field tests and at least one test location showing a <0 dB sensitivity on four consecutive baseline tests (mean with standard deviation between brackets unless stated otherwise)

	N = 221	N = 53
Baseline		
Age (years)	65.1 (12.3)	65.1 (10.3)
Gender (% male)	55.2	45.3
Right eye (%)	50.7	52.8
Mean Deviation (median [interquartile range]; dB)	-7.0 (-14.5 to -3.0)	-14.7 (-18.7 to -10.9)
Follow-up		
Follow-up duration (years)	6.4 (1.2)	6.9 (1.0)
Rate of progression (median [interquartile range]; dB/year)	-0.1 (-0.5 to +0.1)	-0.2 (-0.5 to 0.0)
Square root of the residual mean square of Mean Deviation (dB)	1.1 (0.7)	1.2 (0.7)

Figure 1 shows the blindness persistence characteristics as a function of the number of consecutive baseline sensitivities below <0, <5, <10 and <20 dB. The boxplots visualize the between-patient variability; the corresponding mean within-patient standard deviations were, following the sequence of Figure 1 from left to right, 25, 23, 21, 20, 16, 15, 13, 12, 14, 10, 9, 10, 13, 9, 9 and 8%. If the number of consecutive baseline tests on which a test location was blind increased, the probability of being blind during follow-up increased. The increase in blindness persistence appeared to saturate at three consecutive baseline sensitivities below the concerning value. Blindness persistence appeared to be highest for <10 and <20 dB and lowest for <0 dB. Table 3 shows that blindness persistence depended significantly on both the number of consecutive tests showing blindness in a test location (P<0.001) and the definition of blindness (P<0.001). Figure 2 presents the corresponding mean sensitivity as recorded during the four follow-up tests in the presumed blind test locations.

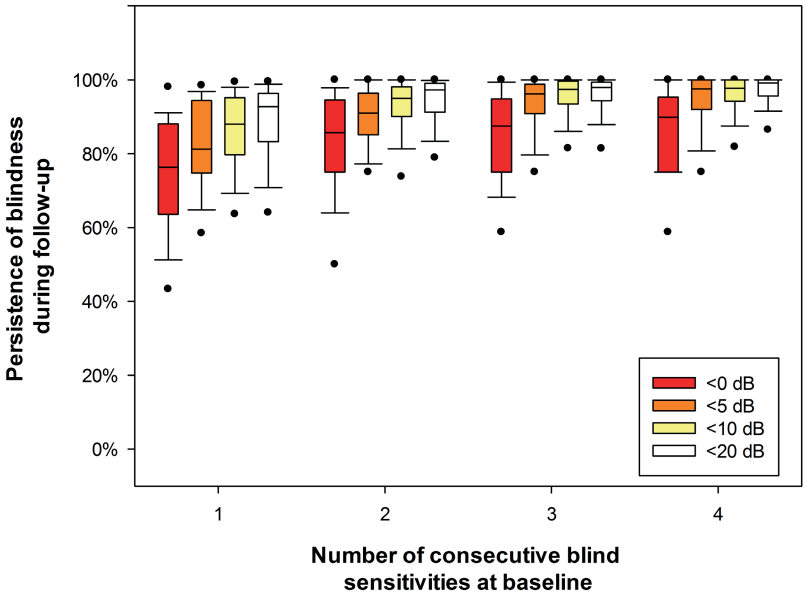


Figure 1. Percentage of follow-up sensitivities being <0, <5, <10 and <20 dB as a function of the number of consecutive <0, <5, <10 and <20 dB (1, 2, 3 and 4, respectively) baseline sensitivities. Boxplots show median, interquartile range, and 5th, 10th, 90th and 95th percentiles.

Table 3. Analysis of variance with the number of consecutive tests showing blindness in a test location (N) and the definition of 'perimetrically blind' (B) as within-subject factors and the persistence of blindness as dependent variable

	df	MS	dfe	MSe	F	*P*-value
mean	1	680.65	52	0.08460	8045	<0.001
N	3	0.47354	156	0.00702	67	<0.001
B	3	0.57498	156	0.01089	53	<0.001
N*B	9	0.00547	468	0.00161	3	<0.001

df = degrees of freedom; MS = mean squares (MS = SS/df with SS = sum of squares); dfe is df for error; MSe = mean squares for error.

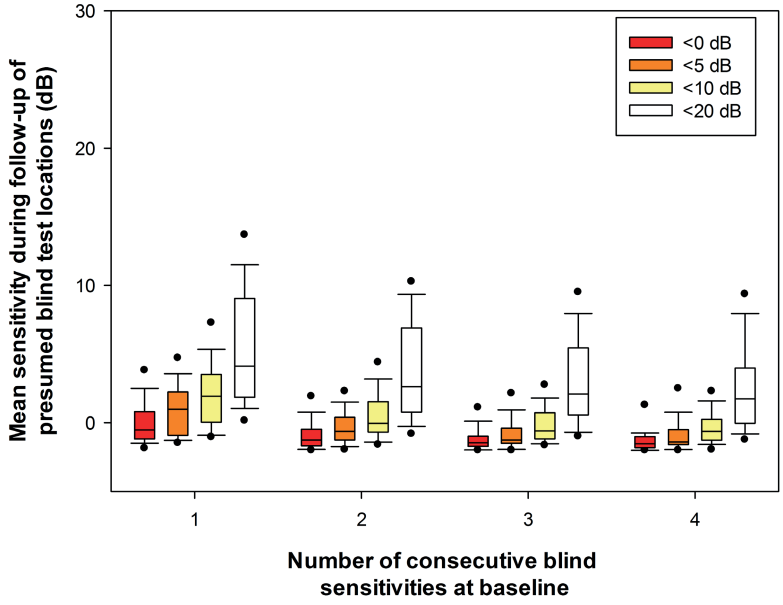


Figure 2. Mean sensitivity during follow-up in test locations with 1, 2, 3 and 4 consecutive <0, <5, <10 and <20 dB baseline sensitivities. Boxplots show median, interquartile range, and 5th, 10th, 90th and 95th percentiles.

Figure 3 illustrates the spatial distributions of test locations that met the requirement of three consecutive sensitivities of <0 (A), <10 (B) and <20 dB (C), and that were 'out of range' according to the GPA (D), as a function of baseline MD. The number of blind test locations increased monotonically with MD for all criteria of blindness except for GPA; for GPA the number of test locations flagged as 'out of range' decreased again with advanced glaucoma. As a consequence, significantly less sensitivities were 'out of range' according to GPA compared to blindness at <0 dB for baseline MD values below -25 dB *(P*<0.001), while the opposite was the case for all other strata (*P*<0.001). For the three consecutive <0 dB criterion (Figure 3A), the median (interquartile range) percentages of blind test locations were 0(0-0), 0(0-2), 4(0-9), 17(8-27), 27(20-40) and 60(50-70)% for the six MD strata.

Figure 4 presents a scatter plot showing the percentages of blind test locations according to the three consecutive <0 dB criterion for all test locations of the 30-2 grid versus a subset of 12 test locations located within the 10-2 grid. For the subset, the median (interquartile range) percentages of blind test locations were 0(0-0), 0(0-2),

6(4-11), 8(4-18), 23(13-48) and 50(40-70)% for the six MD strata. These percentages were similar to the corresponding percentages for the 30-2 grid (listed above) for all six MD strata (*P*=0.32, 0.34, 0.11, 0.44, 0.17 and 0.23, respectively).

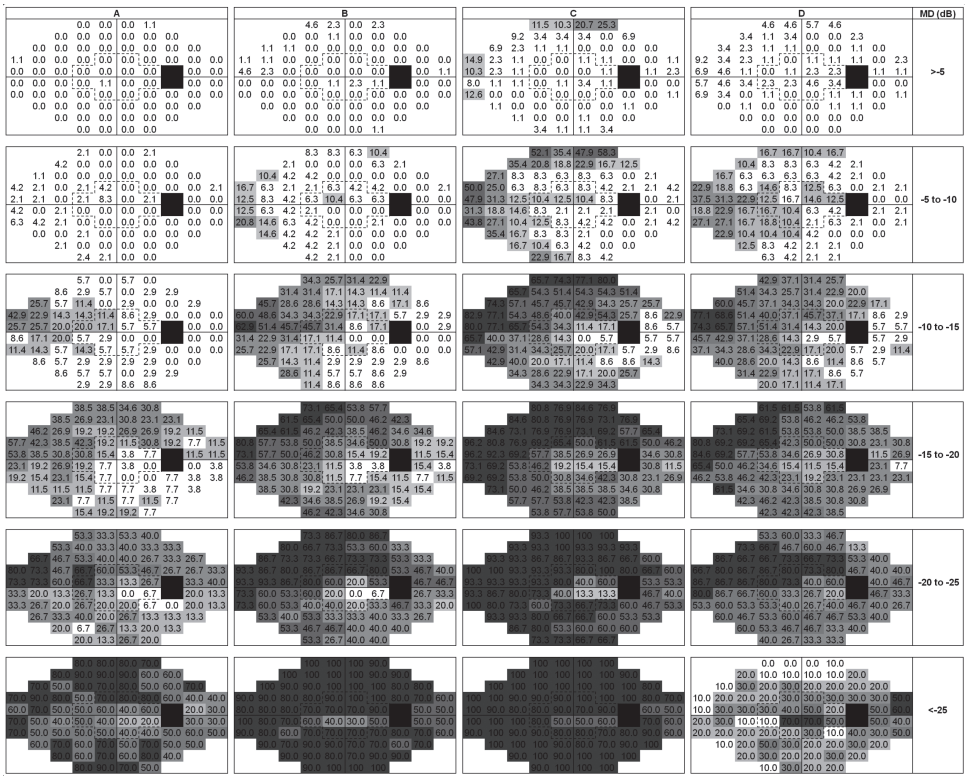


Figure 3. Spatial distributions of test locations that met the requirement of three consecutive sensitivities of <0 (A), <10 (B) and <20 dB (C), and that were 'out of range' according to the Glaucoma Progression Analysis (GPA; D), as a function of baseline mean deviation (MD). Black squares are the blind spot; white squares are test locations flagged as blind in 0-10% of the patients. The remaining intermediate four gray scales denote, from light to dark, blindness in 10-20%, 20-40%, 40-60% and above 60% of the patients.

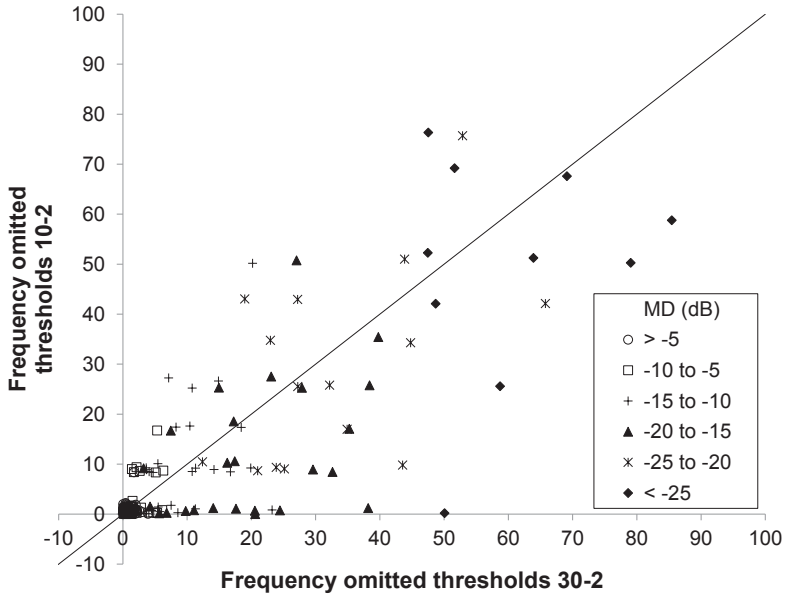


Figure 4. Percentage of blind test locations according to the three consecutive <0 dB criterion for all test locations within the 30-2 grid (x-axis) versus a subset of test locations within the 10-2 grid (y-axis). Symbols indicate stratification according to baseline mean deviation (MD) in six strata, being up to -5 dB, from -5 to -10 dB, -10 to -15 dB, -15 to -20 dB, -20 to -25 dB en beyond -25 dB. Noise with a standard deviation of 1% was added in order to avoid overlapping data points.

Figure 5 shows Venn diagrams indicating the number of eyes with at least possible progression at the end of the follow-up according NPA versus NPA after removing all test locations that were blind on the first three tests (NPA-RONI, where RONI is regions of no interest) for four different definitions of blindness: <0, <5, <10 and <20 dB. There was no significant difference between the classifications by both approaches (P=0.25, P=1.0, P=1.0 and P=0.26 for <0, <5, <10 and <20 dB, respectively). Similar findings were done in the subset of 53 eyes (P=0.25, P=1.0, P=1.0 and P=0.75 for <0, <5, <10 and <20 dB, respectively).

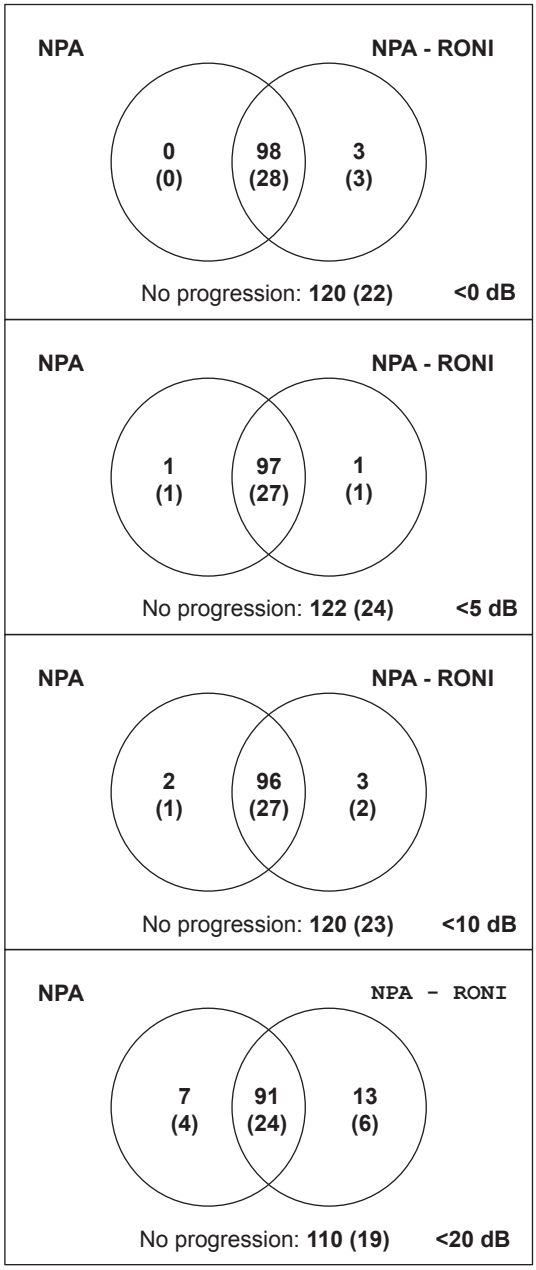


Figure 5. Venn diagrams showing progression according Nonparametric Progression Analysis (NPA) versus NPA after removing all test locations that were blind on the first three tests (NPA-RONI, where RONI is regions of no interest) for four different definitions of blindness: <0, <5, <10 and <20 dB. Results for all 221 subjects with results for the subset of 53 subjects between brackets.

DISCUSSION

Test locations with a sensitivity below a certain value on three consecutive occasions are unlikely to show a substantially higher sensitivity later on. Hence, if the concerning value corresponds to the maximum stimulus intensity of the perimeter used, these test locations do no longer contribute to progression detection. Omitting these locations from future tests will result in time saving without hampering progression detection. Obviously, the number of blind test locations (and thus the potential time saving) increases with increasing disease severity. Interestingly, the percentages of blind test locations appeared to be similar for 30-2 and 10-2 grids for all disease stages.

With the introduction of the SITA strategies in the late ninety's of the previous century, the examination time of standard automated perimetry decreased substantially. (Bengtsson et al. 1997) Unfortunately, this advantage over the full-threshold strategy is largely lost in severe glaucoma. Older Octopus strategies and the German Adaptive Threshold Estimation (GATE) algorithm overcome this increase in test time by using information from previous test results to determine more appropriate starting values for the stair-case procedure.(Frankhauser et al. 1977, Schiefer et al. 2010) We would suggest a further step by entirely omitting test locations that were shown to be blind at earlier occasions ('regions of no interest'). This enables more time saving but obviously limits the application of our approach to irreversible eye diseases. Leaving out test locations may seem crude, but this is what is actually done by clinicians who exchange the default 30-2 grid by a 10-2 grid in advanced glaucoma and by clinicians who rely on GPA for progression detection. Interestingly, GPA ignores even more test locations than we propose to do with our 'regions of no interest' approach (see below and Results section). As GPA leaves them out in the analysis phase only, however, no time saving is obtained.

The time gained by the suggested approach should be interpreted and weighed correctly. Obviously, if the time saving is compared to the total time spent in the hospital, the saving is negligible. However, not testing blind test locations refrains a patient with moderate or advanced glaucoma from long time periods in which he or she does not observe any stimulus but has to stay alert nonetheless. This should increase concentration, thus increasing the reliability of the test result. Second, long time periods without any visible stimulus increase patient frustration by emphasizing not seeing things. Third, the saved time can be used to study the remaining parts of the visual field in more detail without additional visits or costs. This can be done by either adding test locations or determining thresholds more accurately. Obviously, to allow for a reliable progression detection throughout the follow-up, only the test locations

belonging to the original grid should contribute to the MD. The added test locations, however, may be analyzed separately and may yield important information.(Hood et al. 2011, Schiefer et al. 2010)

A caveat of incorporating our regions-of-no-interest approach is that it may cause propagation of blindness through the visual field if applied to strategies that use some form of spatial smoothing (that is, do not determine a formal threshold in all individual test locations) in order to reduce test time (as possibly occurs in SITA). This will not occur in strategies that use neighboring sensitivities only for estimating a starting value for determining a threshold.

The classical picture of glaucoma deterioration is the development of visual field defects initially in the periphery, leaving vision unaltered centrally until the latest stages of the disease. Albeit this picture has been challenged recently,(Hood et al. 2011, Schiefer et al. 2010) the clinical translation of this picture is starting with 30-2 testing with a transition to 10-2 somewhere along the line - the easiest way to get rid of unresponsive parts of the visual field in advanced disease. One of the aims of this study was to develop a clinically useful guideline, that is, an MD cut-off value, for preferring 10-2 testing over 30-2 testing in advanced glaucoma. Interestingly, no such MD value appeared to exist – the median percentage of blind test locations was essentially identical for 30-2 and 10-2 grids for all disease stages. With a closer look at our data, this corresponded to the three clinically well known patterns of visual field loss in severe glaucoma: (1) a central island without a peripheral (temporal) island, (2) a temporal island without a central island, and (3) both a central and a temporal island. This is also visible in Figure 4. Hence, in many patients a transition from 30-2 to 10-2 testing will never become a meaningful change. It is important to realize that we did not actually measure a 10-2 grid – a form of high spatial resolution perimetry(Weber et al. 1989, Westcott et al. 1997) - but analyzed a subset of 30-2 test locations that lay within the 10-2 area. Here, the assumption is that this can be considered a representative (unbiased) sample. Also, inclusion of a patient in this study implied the presence of 30-2 fields. This might have induced a selection bias, as patients with only a central island might be underrepresented because they were at baseline already monitored with 10-2 testing – and thus excluded. This is unlikely, however, as at the baseline of the GLGS Goldmann perimetry and not 10-2 testing was the default escape in advanced glaucoma(Heeg et al. 2005) – suggesting an underrepresentation of temporal islands rather than of central islands in this study. To conclude, the transition from 30-2 to 10-2 testing should be individualized and the advantage of a more detailed monitoring of a

central island should be weighed against the need of building a new baseline and the loss of monitoring of any peripheral island. After all, it is not unlikely that progression in the periphery predicts future central loss.

Figure 3A-C actually depicts the "average" glaucoma progression pattern. Not unexpectedly, the glaucomatous deterioration starts nasal-superiorly. In agreement with the findings discussed in the paragraph above, both a central and a temporal island survived until the last MD stratum. With GPA, the number of test locations with a cross (indicating that the software ignores these locations for progression detection) increases with disease progression up to an MD of about -20 dB but decreases beyond that point (Figure 3D). Although this pattern is identical to what is observed in the pattern deviation plot and is in agreement with the idea that GPA is based on pattern deviation analysis,(Bengtsson et al. 1997) it might mislead the clinician as it suggests erroneously that test locations that are actually blind are still monitored.

The absence of a response to the maximum stimulus intensity is not identical to blindness. The dynamic range of the perimeter can be increased by replacing stimulus size III by size V. Interestingly, this appears to reduce the test-retest variability.(Wall et al. 1997, Wall et al. 2009, Wall et al. 2010) Until now, however, the time-saving SITA strategy is not available for size V. Within a given stimulus size, it is not self-evidently beneficial to increase the dynamic range by increasing the maximum stimulus intensity. Although the well-known pointwise test-retest variability plot (for our data shown in Figure 6) suggests a reduced variability close to the maximum stimulus intensity, this is merely a floor effect. If we look in an alternative way to the same data (Figure 1), it might be the case that the extended dynamic range as used in HFA compared to Octopus and Oculus corresponds to a reduced reproducibility of blindness (Table 3). This is in line with the idea that a high test-retest variability is related to ganglion cell saturation,(Swanson et al. 2011) but requires further study.

The exclusion of test locations with a sensitivity of <0, <5 or <10 dB at baseline did not affect progression detection with NPA (Figure 5). Only for <20 dB some difference (albeit statistically not significant) appeared to occur. Here, progression according to NPA but not according to NPA-RONI might reflect deepening of existing defects (that is test locations with a sensitivity already <20 dB at baseline); progression according to NPA-RONI but not according to NPA might be caused by a reduced variability in the calculated mean sensitivity for the RONI approach, which results in an increase in NPA sensitivity.(Jansonius 2005, Wesselink et al. 2012) These observations are in line

with the findings described in the previous paragraph. It might be possible that other progression detection algorithms would be affected differently. This requires further study.

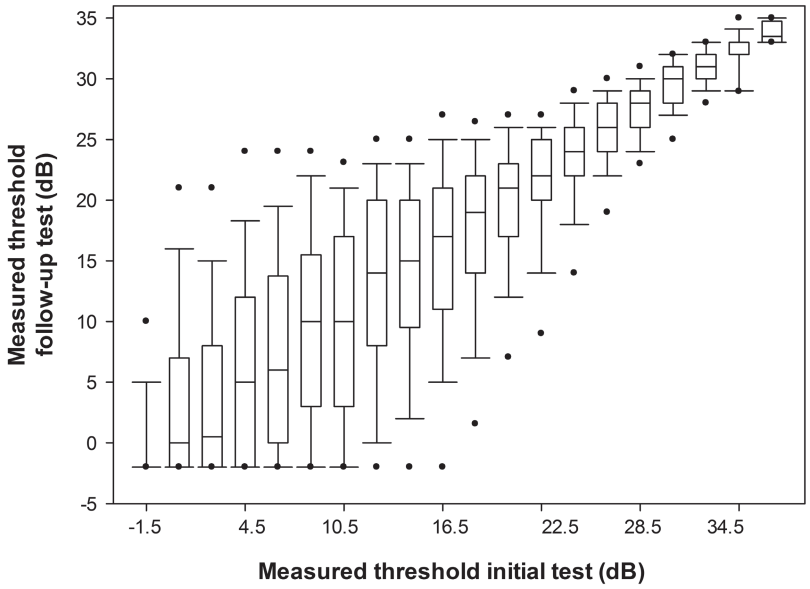


Figure 6. Pointwise test-retest variability. Data presented in strata of 2 dB, except for <0 dB which was set to -1.5 dB in one box. Boxplots show median, interquartile range, and 5th, 10th, 90th and 95th percentiles.

Originally, the SITA fast strategy, as used in the GLGS, was considered a time-saving improvement of the SITA standard strategy and for that reason we adopted it in our study designed in 1999. Later it became clear that the strategies performed slightly different. Two studies reported a slightly higher sensitivity for SITA standard in comparison with SITA fast;(Budenz et al. 2002, Delgado et al. 2002) one study reported a higher sensitivity for SITA fast.(Pierre-Filho et al. 2006) These differences – if any – are not relevant to the current study. More relevant to the current study is the finding that SITA fast seems to have a higher test-retest variability in areas with a reduced sensitivity in comparison with SITA standard.(Artes et al. 2002) This tentatively suggests that blindness reproducibility might be better in SITA standard and thus our criterion – three consecutive <0 dB readings – should be applicable to SITA standard as well.

In conclusion, current perimetric strategies share the inconvenient property that test-time increases in advanced glaucoma, while a smaller residual visual field has to be tested. A more clever customizing to what has to be tested than a default change to 10-2 testing should allow for an improved and uninterrupted long-term monitoring of glaucoma patients with standard automated perimetry.

REFERENCES

Artes PH, Iwase A, Ohno Y, Kitazawa Y & Chauhan BC (2002): Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. Invest Ophthalmol Vis Sci. 43: 2654–2659.

Bengtsson B & Heijl A (1998): Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. Acta Ophthalmol Scand. 76: 268–272.

Bengtsson B & Heijl A (1998): SITA Fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma. Acta Ophthalmol Scand. 76: 431–437.

Bengtsson B, Lindgren A, Heijl A, Lindgren G, Asman P & Patella M (1997): Perimetric probability maps to separate change caused by glaucoma from that caused by cataract. Acta Ophthalmol Scand. 75: 184–188.

Bengtsson B, Olsson J, Heijl A & Rootzén H (1997): A new generation of algorithms for computerized threshold perimetry, SITA. Acta Ophthalmol Scand. 75: 368-375.

Budenz DL, Rhee P, Feuer WJ, McSoley J, Johnson CA & Anderson DR (2002): Sensitivity and specificity of the Swedish Interactive Threshold Algorithm for glaucomatous visual field defects. Ophthalmology. 109: 1052–1058.

Chauhan BC, Tompkins JD, LeBlanc RP & McCormick TA (1993): Characteristics of frequency-of-seeing curves in normal subjects, patients with suspected glaucoma, and patients with glaucoma. Invest Ophthalmol Vis Sci. 34: 3534–3540.

Delgado MF, Nguyen NTA, Cox TA, Singh K, Lee DA, Dueker DK, Fechtner RD, Juzych MS, Lin SC, Netland PA, Pastor SA, Schuman JS, Samples JR & American Academy of Ophthalmology. Ophthalmic Technology Assessment Committee 2001-2002 Glaucoma Panel (2002): Automated perimetry: a report by the American Academy of Ophthalmology. Ophthalmology. 109: 2362–2374.

Frankhauser F, Spahr J & Bebie H (1977): Some aspects of the automation of perimetry. Surv Ophthalmol. 22: 131–141.

Heeg GP, Blanksma LJ, Hardus PLLJ & Jansonius NM (2005): The Groningen Longitudinal Glaucoma Study. I. Baseline sensitivity and specificity of the frequency doubling perimeter and the GDx nerve fibre analyser. Acta Ophthalmol Scand. 83: 46–52.

Hood DC, Raza AS, de Moraes CGV, Odel JG, Greenstein VC, Liebmann JM & Ritch R (2011): Initial arcuate defects within the central 10 degrees in glaucoma. Invest Ophthalmol Vis Sci. 52: 940–946.

Hudson C, Wild JM & O'Neill EC (1994): Fatigue effects during a single session of automated static threshold perimetry. Invest Ophthalmol Vis Sci. 35: 268–280.

Jansonius NM (2005): Bayes' theorem applied to perimetric progression detection in glaucoma: from specificity to positive predictive value. Graefes Arch Clin Exp Ophthal. 243: 433–437.

Leske MC, Heijl A, Hyman L & Bengtsson B (1999): Early Manifest Glaucoma Trial: design and baseline data. Ophthalmology. 106: 2144–2153.

Pierre-Filho P de TP, Schimiti RB, de Vasconcellos JPC & Costa VP (2006): Sensitivity and specificity of frequency-doubling technology, tendency-oriented perimetry, SITA Standard and SITA Fast perimetry in perimetrically inexperienced individuals. Acta Ophthalmol Scand. 84: 345–350.

Schiefer U, Papageorgiou E, Sample PA, Pascual JP, Selig B, Krapp E & Paetzold J (2010): Spatial pattern of glaucomatous visual field loss obtained with regionally condensed stimulus arrangements. Invest Ophthalmol Vis Sci. 51: 5685–5689.

Schiefer U, Pascual JP, Edmunds B, Feudner E, Hoffmann ES, Johnson CA, Lagrèze WA, Pfeiffer N, Sample PA, Staubach F, Weleber RG, Vonthein R, Krapp E & Paetzold J. (2009): Comparison of the new perimetric GATE strategy with conventional full-threshold and SITA standard strategies. Invest Ophthalmol Vis Sci. 50: 488–494.

Swanson WH, Sun H, Lee BB & Cao D (2011): Responses of primate retinal ganglion cells to perimetric stimuli. Invest Ophthalmol Vis Sci. 52: 764–771.

Turpin A, Jankovic D & McKendrick AM (2007): Retesting visual fields: utilizing prior information to decrease test-retest variability in glaucoma. Invest Ophthalmol Vis Sci. 48: 1627–1634.

Wall M (2004): What's new in Perimetry. Journal NeuroOphthalmol. 24: 46–55.

Wall M, Kutzko KE & Chauhan BC (1997): Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. Invest Ophthalmol Vis Sci. 38: 426–435.

Wall M, Woodward KR, Doyle CK & Artes PH (2009): Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. Invest Ophthalmol Vis Sci. 50: 974–979.

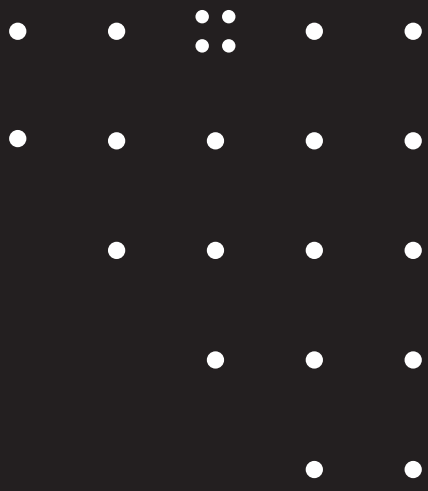
Wall M, Woodward KR, Doyle CK & Zamba G (2010): The effective dynamic ranges of standard automated perimetry sizes III and V and motion and matrix perimetry. Arch Ophthalmol. 128: 570–576.

Weber J, Schultze T & Ulrich H (1989): The visual field in advanced glaucoma. Int Ophthalmol. 13: 47–50.

Wesselink C, Heeg GP, Jansonius NM (2009): Glaucoma Monitoring in a clinical setting: glaucoma progression analysis vs nonparametric progression analysis in the Groningen Longitudinal Glaucoma Study. Arch Ophthalmol. 127: 270-274.

Wesselink C, Marcus MW & Jansonius NM (2012): Risk factors for visual field progression in the groningen longitudinal glaucoma study: a comparison of different statistical approaches. J Glaucoma. 21: 579–585.

Westcott MC, McNaught AI, Crabb DP, Fitzke FW & Hitchings RA (1997): High spatial resolution automated perimetry in glaucoma. Br J Ophthalmol. 81: 452–459.



Factors that influence standard automated perimetry test results in glaucoma:

Test reliability, technician experience, time of day and season

Francisco G. Junoy Montolio, Christiaan Wesselink, Marijke Gordijn, Nomdo M. Jansonius

Investigative Ophthalmology & Visual Science (IOVS). 2012;53(11):7010-7

ABSTRACT

Purpose: To determine the influence of several factors on standard automated perimetry test results in glaucoma.

Methods: Longitudinal Humphrey Field Analyzer 30-2 Swedish interactive threshold algorithm data from 160 eyes of 160 glaucoma patients were used. The influence of technician experience, time of day and season on the mean deviation (MD) was determined by performing linear regression analysis of MD against time on series of visual fields and subsequently performing a multiple linear regression analysis with the MD residuals as dependent variable and the factors mentioned above as independent variables. Analyses were performed with and without adjustment for the test reliability (fixation losses and false-positive and false-negative answers) and with and without stratification according to disease stage (baseline MD).

Results: Mean follow-up was 9.4 years with on average 10.8 tests per patient. Technician experience, time of day and season were associated with the MD. Approximately 0.2 dB lower MD values were found for inexperienced technicians ($P<0.001$), tests performed after lunch ($P<0.001$) and tests performed in the summer or autumn ($P<0.001$). The effects of time of day and season appeared to depend on disease stage. Independent of these effects, the percentage of false-positive answers strongly influenced the MD with a 1 dB increase in MD per 10% increase in false-positive answers.

Conclusion: Technician experience, time of day, season and the percentage of false-positive answers have a significant influence on the MD of standard automated perimetry.

INTRODUCTION

Standard automated perimetry (SAP) is an invaluable test for the diagnosis and monitoring of glaucoma. However, as SAP is a subjective test that aims to measure a sensitivity threshold in a living organism, SAP test results are prone to variability. Knowing the sources of the variability might enable reducing it, and this may improve progression detection in glaucoma. With less variability, smaller changes can be picked-up and fewer tests are needed.

The variability of SAP test results depends on many factors. A well-known factor is disease stage.(Bengtsson & Heijl 2000, Blumenthal et al. 2003) Other factors are patient motivation and technician performance, the latter via appropriateness of refraction and patient instruction, reassurance before the test, and patient monitoring during the test.(Anderson et al. 2001, Kutzko et al. 2000, Mutlukan 1994)

Circadian rhythms and seasonal influences may also contribute to the variability. These periodicities may influence SAP test results in at least two different ways. First, patients need a good cognitive function to perform perimetry. In an extensive review, Blatter & Cajochen described a daily variation in cognitive function in humans.(Blatter & Cajochen 2007) Cognitive function deteriorated over the day in elderly (at the typical age of a glaucoma patient), while the opposite is the case in young subjects. Second, SAP measures a threshold sensitivity of the visual system. Earlier studies in humans showed a daily variation in retinal visual sensitivity, suggesting lower sensitivities in the early morning.(Roenneberg et al. 1992, Tassi et al. 2000, Tuunainen et al. 2001) Besides a daily variation, a seasonal (circannual) variation might also be present. After all, exposure to light varies per season and prior light history affects light sensitivity. (Hébert et al. 2002) Indeed, retinal sensitivity seems to be highest in the spring.(Bassi & Powers 1986, Sweeney et al. 1960)

Variability is entangled with reliability. Reliability is commonly assessed by reliability indices, and displayed as the percentages of fixation losses (FL) and false-positive (FP) and false negative (FN) answers. Two issues concerning reliability are relevant to this study. First, whether reliability as assessed by the reliability indices influences variability. Second, whether other factors that influence variability act through the reliability or have their own, independent influence.

The aim of this study was to determine the influence of several factors on SAP test results in glaucoma. For this purpose, we analyzed the influence of technician experience, time of day, season and the percentages FL and FP and FN answers on the mean deviation (MD), a commonly used summary measure of SAP test results.

METHODS

Study population

The present study was performed using data from the Groningen Longitudinal Glaucoma Study (GLGS), a prospective observational cohort study performed in a clinical setting.(Heeg et al. 2005, Wesselink et al. 2009) In the GLGS, all glaucoma patients and glaucoma suspects who visited the glaucoma outpatient service of the University Medical Center Groningen between July 1, 2000, and June 30, 2001, and who provided informed consent were included in an observational study with conventional perimetry, frequency-doubling perimetry (FDT; Carl Zeiss Meditec AG, Jena, Germany) and laser polarimetry (GDx; Laser Diagnostic Technologies, San Diego, California, USA). The study comprised both glaucoma patients and glaucoma suspect patients and was originally designed as a health technology assessment study. The original aim was to determine if it was possible to replace, in glaucoma patients and/or glaucoma suspects, the lengthy and cumbersome conventional perimetry by FDT and/or GDx. After the initial health technology assessment study, we continued performing conventional perimetry in glaucoma patients and moved to FDT/GDx in glaucoma suspects in our regular care. The GLGS continued as an ongoing anonymous gathering of all information from all glaucoma patients and glaucoma suspects obtained during regular care.

For the present study, we used data from a subpopulation of the GLGS cohort: we only included patients who had (1) glaucoma at baseline (for criteria see below), (2) series of at least eight (nine with discarded learning test) visual field tests made with standard automated perimetry (Humphrey field analyzer [HFA]; Carl Zeiss Meditec Inc., Dublin, CA), and (3) a follow-up of at least five years.(Jansonius 2010) The study protocol was approved by the ethics board of the University Medical Center Groningen and followed the tenets of the Declaration of Helsinki.

Perimetry and glaucoma

Perimetry was performed using the HFA 30-2 Swedish interactive threshold algorithm (SITA) fast strategy. An abnormal test result was defined as any one of the following:

a glaucoma hemifield test "outside of normal limits", a pattern standard deviation with $P<0.05$, or three adjacent non-edge points with $P<0.05$ in the pattern deviation probability plot, of which at least one point reached $P<0.01$, with all points being on the same side of the horizontal meridian. As the influence of reliability was part of the outcome of the study, we did not exclude test results based on reliability indices.

For glaucomatous visual field loss at baseline, two consecutive perimetry test results had to be abnormal in at least one eye. Defects had to be in the same hemifield, and at least one depressed test point of these defects had to have exactly the same location on both fields. Moreover, defects had to be compatible with glaucoma and without any other explanation. This was judged by one of three glaucoma specialists involved in the baseline of the study.(Heeg et al. 2005) A visual field test prior to the two baseline tests was discarded in order to reduce the influence of learning. Hence, at least three tests had to be performed at baseline before glaucomatous visual field loss could be diagnosed.

For being a glaucoma patient at baseline, glaucomatous visual field loss had to be present. Neither the intraocular pressure (IOP) nor the aspect of the optic disk was a formal part of the glaucoma diagnosis. However, the mean (standard deviation [SD]) pre-treatment IOP of the presumed glaucoma patients was 30.3(9.5) mmHg and 90% of them had an abnormal optic disk according to the GDx (the Number > 29).(Heeg et al. 2005, Wesselink et al. 2012)

During the follow-up period, perimetry was performed at a frequency of one test per year. Clinicians were allowed to increase the frequency of testing, for example in case of suspected progression. This was a subjective decision; no formal tools or rules were given (observational study design).

Data analysis

For this study, one eye per patient was included. If both eyes met the above-described criteria, one eye was chosen randomly. Parametric and nonparametric descriptive statistics were performed to describe the study population.

To determine the influence of the factors technician experience, time of the day and season, and of the reliability indices, we first performed a linear regression analysis of MD against time for all included series of visual fields and calculated the MD residuals (that is, the distances between the measured MD values and the corresponding regression line). Subsequently, we selected one random visual field of each patient and performed a multiple linear regression analysis with the MD residual of the included

fields as dependent variable and technician experience, time of day, season and phase of follow-up (see below) as independent variables. We performed these analyses both with and without adjustment for the percentages of FL and FP and FN answers. In this way, we were able to (1) determine the influence of these indices on the MD and (2) explore if potential effects of the other factors were caused by a direct influence or by an indirect influence, through the reliability. The random selection of one visual field per patient and subsequent multiple linear regression analysis was repeated 50 times. After applying this resampling technique, we presented the results as the mean effect estimates averaged over the 50 resamplings with corresponding 95% confidence intervals (CIs), for all studied independent variables. As a secondary analysis, the analyses were repeated with the pattern standard deviation (PSD) residuals as dependent variable. A *P*-value of 0.05 or less was considered statistically significant (see Discussion section).

Categorical variables were recoded into dummy variables. For technician experience, the technicians were stratified into three categories: inexperienced, moderately experienced and highly experienced. Twenty-eight technicians were involved in the study; for the stratification according to experience we ranked these technicians according to the number of visual fields they had performed. Subsequently, we divided the visual fields into three equally-sized groups. This resulted in a group of 20 inexperienced technicians who performed a median of 22 tests (range 1 to 82 tests) during the entire follow-up period, a group of five moderately experienced technicians who performed a median of 120 tests (range 83 to 147 tests) and a group of three highly experienced technicians who performed a median of 171 tests (range 161 to 193 tests) as part of the study. For time of day, the tests were stratified into four categories: performed before 10AM, between 10AM and noon, between noon and 2PM and after 2PM. For season, the tests were also stratified into four categories, of three months each, based on the annual variation of retinal sensitivity as found by Sweeney et al.(Sweeney et al. 1960) The summer was classified as June, July and August, the autumn as September, October and November, and so on. The reliability indices were treated as continuous variables measured relative to their mean value in the concerning patient (for example, 10% FP answers were coded as +5% in a patient who had on average 5% FP answers and as -5% in a patient who had on average 15% FP answers). Finally, we added an additional variable, being the phase of follow-up. This variable reflects if a visual field belongs to the first, middle or last part of the follow-up. This variable aims to adjust for possible deviations from the presumed linear relationship between MD and time. These deviations might occur for example due to prolonged learning or disease acceleration and might confound some of the studied relationships (see Discussion).

The MD analyses were performed for the entire study population, with and without adjustment for the reliability indices. They were also performed after stratification according to glaucoma severity. Here, two strata were employed, being a baseline MD of -6 dB or better (early glaucoma) versus below -6 dB (moderate/severe glaucoma). This cut-off value yielded two, roughly equally sized groups. The PSD analyses were only performed for early glaucoma as the PSD is only for early glaucoma monotonically related to disease stage.

RESULTS

One-hundred-sixty eyes of 160 patients, 89 men and 71 women, were included. The mean (SD) age at baseline was 63.6(11.2) years. The baseline MD was skewed negatively with a median value of -7.8 dB (range: -29.4 to -0.0 dB; interquartile range: -15.4 to -2.9 dB). There was a median MD decline of -0.2 dB/year (range: -3.2 to +0.7 dB/year; interquartile range: -0.5 to 0.0 dB/year; P90 -0.8 dB/year). The mean (SD) follow up was 9.4(1.7) years with a mean of 10.8(2.2) visual field tests per patient. In total 1735 tests were performed, of which 107 had to be excluded from further analysis because either no technician was registered, two technicians accompanied the test together as part of training, no fixation was monitored or the FN answers were not available (which sometimes occurs in end-stage glaucoma; fields with FN answers not available had a mean (SD) MD of -27.0(2.2) dB). Hence, 1628 visual field tests were included.

Figure 1 presents the frequency distributions of the FL and the FP and FN answers of all included 1628 visual field tests. The FL showed the highest percentages, followed by the FN and the FP answers. The percentages of FL and FP answers were significantly higher in early glaucoma compared to moderate/severe glaucoma (Mann-Whitney test $P<0.001$). The percentage of FN answers did not differ between early and moderate/severe glaucoma ($P=0.62$). There was no significant correlation between the three reliability indices (Spearman's rank correlation coefficients <0.30 for all three comparisons).

Table 1 presents the results of the multiple linear regression analysis for the entire study population, for all independent variables including the reliability indices. Table 2 shows the same analysis after the removal of the reliability indices. The effects of the various independent variables were essentially the same in both tables, indicating that these variables influence the MD directly rather than that they affect the reliability – with a subsequent effect on the MD. Figure 2 gives the results after stratification according to glaucoma stage.

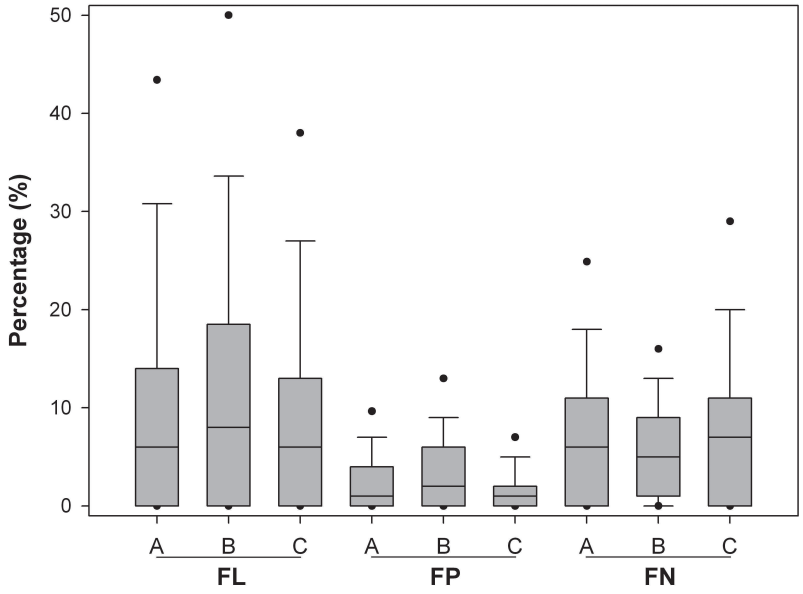


Figure 1. Frequency distributions of the reliability indices FL, FP and FN answer for all patients (A) and after stratification in early (B; mean deviation at baseline -6 or better) and moderate/severe (C) glaucoma. *Box plots* with *error bars* and *black dots* represent the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles.

Reliability indices

As can be seen in Table 1, the percentage of FP answers had the biggest influence on the MD; a one percent increase in FP answers yielded an MD increase of approximately 0.1 dB (that is, the MD of a visual field with 10% FP answers is overestimated by approximately 1 dB in comparison with a visual field with 0% FP answers). The influence of the percentages of FL and FN answers on the MD was also statistically significant, but the effect estimates were obviously lower than that of the percentage of FP answers. Only in early glaucoma, a clinically relevant effect of the percentage of FN answers was found. Here, the MD of a visual field with 10% FN answers is underestimated by approximately 0.5 dB in comparison with a visual field with 0% FN answers (Figure 2).

Table 1. Influence of technician experience, time of day, season, follow-up and test reliability on the mean deviation (dB), in the entire study population

	Beta	95% confidence interval	*P*-value
FL residuals	0.007	0.005 to 0.009	<0.001
FP residuals	0.086	0.078 to 0.095	<0.001
FN residuals	0.006	0.001 to 0.011	0.037
Technician:			
Highly experienced	reference		
Medium experienced	-0.024	-0.108 to 0.060	0.571
Inexperienced	-0.176	-0.261 to -0.091	<0.001
Time of day:			
Before 10:00 am	reference		
10:00 – 11:59 am	-0.033	-0.115 to 0.050	0.433
12:00 – 13:59 pm	-0.197	-0.286 to -0.109	<0.001
After 14:00 pm	-0.062	-0.152 to 0.028	0.174
Season:			
Spring	reference		
Summer	-0.132	-0.225 to -0.038	0.006
Autumn	-0.189	-0.265 to -0.113	<0.001
Winter	-0.052	-0.146 to 0.041	0.266
Follow-up:			
Middle part	reference		
First part	-0.240	-0.307 to -0.173	<0.001
Third part	-0.196	-0.275 to -0.117	<0.001

Table 2. Influence of technician experience, time of day, season and follow-up on the mean deviation (dB), in the entire study population

	Beta	95% confidence interval	*P*-value
Technician:			
Highly experienced	reference		
Medium experienced	-0.063	-0.146 to 0.020	0.135
Inexperienced	-0.245	-0.333 to -0.157	<0.001
Time of day:			
Before 10:00 am	reference		
10:00 – 11:59 am	-0.021	-0.100 to 0.058	0.589
12:00 – 13:59 pm	-0.158	-0.250 to -0.066	0.001
After 14:00 pm	-0.069	-0.155 to 0.017	0.112
Season:			
Spring	reference		
Summer	-0.091	-0.182 to 0.001	0.053
Autumn	-0.152	-0.234 to -0.070	0.001
Winter	-0.072	-0.168 to 0.025	0.144
Follow-up:			
Middle part	reference		
First part	-0.226	-0.295 to -0.157	<0.001
Third part	-0.103	-0.191 to -0.016	0.021

Technician experience

Technician experience appeared to be important, independently of glaucoma stage (Table 1; Figure 2). Guidance by inexperienced technicians yielded MD values that were approximately 0.2 dB lower in comparison with their highly experienced colleagues (P<0.001); for intermediately experienced technicians this difference was negligible.

Time of day

Time of day had a significant influence on the MD. Approximately 0.2 dB lower MD values were found directly after lunch (P<0.001; Table 1). After stratification according to glaucoma stage, patients with early glaucoma performed significantly better in the early morning, with an approximately 0.4 dB better MD compared to the rest of the day. In moderate/severe glaucoma, the effect of time of day was less pronounced (Figure 2).

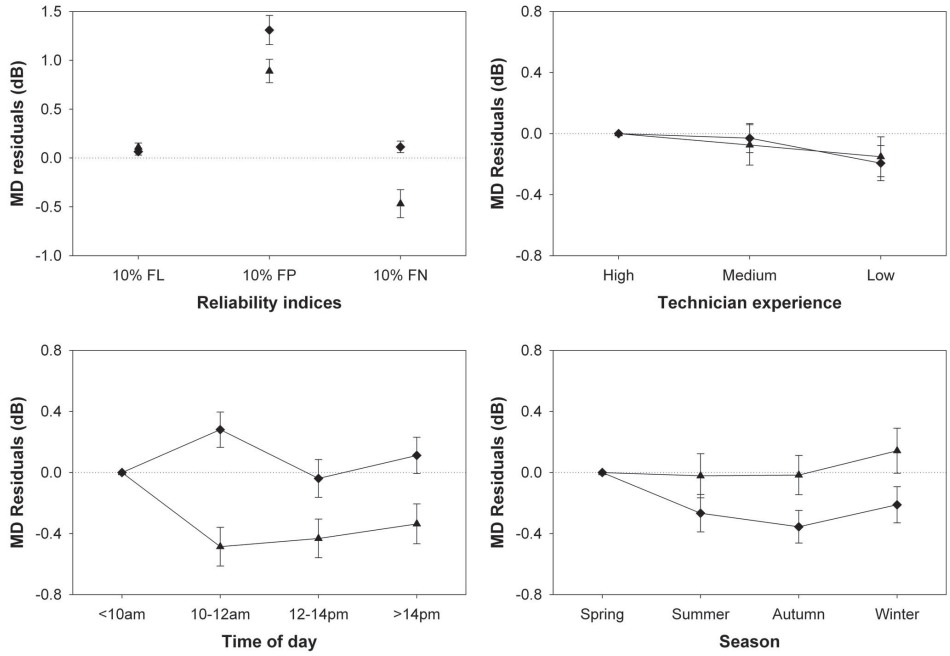


Figure 2. Influence of the test reliability (top left), technicians experience (top right), time of day (bottom left) and season (bottom right) on the mean deviation (MD) after stratification according to glaucoma stage, being early (triangles; mean deviation at baseline -6 dB or better) versus moderate/severe glaucoma (diamonds). Results presented as average beta's with 95% confidence intervals (dB).

Season

Inter-seasonal differences also appeared to play a significant role in visual field testing. Approximately 0.2 dB lower MD values were found in the summer/autumn compared to winter/spring (P<0.001; Table 1). After stratification according to glaucoma stage, patients with early glaucoma appeared to have the highest sensitivity in the winter and patients with moderate/severe glaucoma in the spring (Figure 2).

Follow-up

As can be seen in Table 1, the MD was significantly lower both in the beginning and at the end of the follow-up, compared to the middle part of the follow-up. This indicates a systematic deviation from the assumed linear decay.

Pattern standard deviation

Figure 3 shows the results of the multiple linear regression analysis with the PSD residuals as dependent variable, for patients with early glaucoma. Significant effects were found for the reliability indices (FP answers), time of day, and season. The effects were roughly opposite and generally smaller in comparison with the corresponding analysis with the MD residuals as dependent variable (Figure 2; triangles).

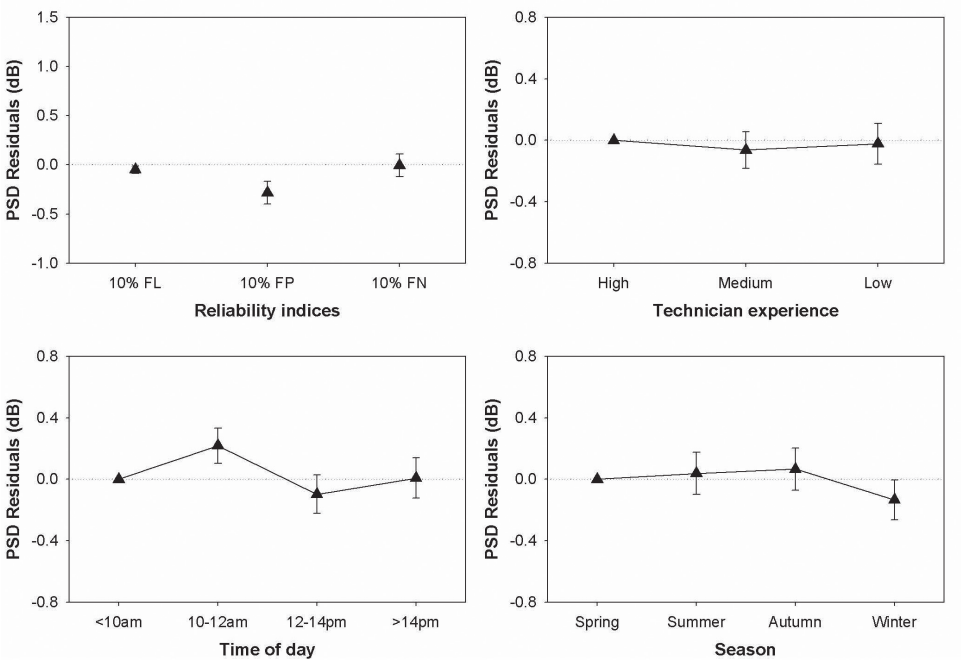


Figure 3. Influence of the test reliability (top left), technicians experience (top right), time of day (bottom left) and season (bottom right) on the pattern standard deviation (PSD) for patients with early glaucoma (mean deviation at baseline -6 dB or better). Results presented as average beta's with 95% confidence intervals (dB).

DISCUSSION

Technician experience, time of day and season have a clinically relevant influence on the MD of SAP test results - together they may cause differences between tests of typically 0.5 dB. Of the three reliability indices, an excess of FP answers is the only serious threat to the test result; the MD is overestimated by 1 dB per 10% of FP answers.

Reliability indices

Lee et al. studied, in a clinical setting, the influence of the reliability indices on the mean sensitivity variable (MS) of the Octopus perimeter (Program G1 201; Interzeag, Schlieren, Switzerland).(Lee et al. 1994) They used multiple linear regression analysis with the inter-test difference of the MS as the dependent variable and the difference of FP or FN catch trials as independent variable (reliable visual field compared to unreliable visual field). They found an increase in MS of 1.5 dB for every 10% FP answers, while the MS decreased by 1.2 dB for every 10% FN answers. Their results are essentially in agreement with our findings, and this suggests that these findings are universal for SAP rather than specific to a single perimeter/strategy. Bengtsson examined the associations between the reliability indices and the reproducibility of the MD with multiple linear regression analysis, using the HFA SITA standard strategy in a clinical setting.(Bengtsson 2000) After disease stage, the percentage of FN answers had the highest, but non-significant, association. This result could be explained by the correlations between FN answers and disease stage and disease stage and reproducibility, which is in agreement with others studies and partially with our results.(Bengtsson 2000, Birt et al. 1997, Katz et al. 1991) In our study, the percentage of FN answers was only of some clinical significance in early glaucoma, while in moderate/severe glaucoma its influence was negligible (Figure 2). Bengtsson did not find a significant influence of the percentage of FP answers. Two studies investigated the influence of FP answers by artificially adding random answers during full threshold(Cascairo et al. 1991) and SITA standard testing. (Newkirk et al. 2006) In healthy subjects, the addition of 33% FP answers resulted in an increase in the Humphrey STATPAC mean defect of 2.9 dB and an increase in the MD of 0.3 dB, respectively. In glaucoma patients, Newkirk et al. found an increase in MD of 2.4 dB.(Newkirk et al. 2006) This agrees with our findings. Our findings also indicate that a FP answers cut-off point of 15%, as advocated for the SITA strategies,(Heijl, Patella & Bengtsson 2012) may be not strict enough.

Technician experience

The role of technicians cannot be underestimated. Proper instructions,(Kutzko et al. 2000) correct refraction(Anderson et al. 2001, Mutlukan 1994) and ensuring optimal conditions contribute to reliable test results, while supervision plays only a minor role. (Johnson et al. 1993, Kramer et al. 2012, Van Coevorden et al. 1999) As far as we know, no publications exist that examined the association between the degree of experience of technicians and the test result. Tables 1 and 2 and Figure 2 show that the test result was negatively influenced by inexperienced technicians, independently of glaucoma stage. Although the average influence of technician experience may be limited, it might be the case that the influence of individual technicians on the test results is

substantially larger. Theoretically, this could be analyzed by putting all technicians individually as a dummy in the technician experience variable. However, there are not enough tests/observations per technician for such a detailed analysis. Thus, all we can conclude is that – on average – inexperienced technicians perform worse. In addition, inexperienced technicians seemed to have performed a substantial number of tests, given that they also performed tests in patients not included in the study (see Methods section). In reality, those classified as inexperienced technicians are working in other parts of our department and visit the glaucoma service only incidentally whereas those classified as experienced technicians have the glaucoma service as their default shop floor.

Time of day

Circadian rhythm plays a major role in the daily life of humans. These rhythms are endogenously generated by the circadian pacemaker, located in the hypothalamus. Adjustments to the circadian pacemaker are made by exposure to environmental light through intrinsically photosensitive retinal ganglion cells containing melanopsin. (Hanifin & Brainard 2007, Markwellet al. 2010) Effects of time of day are difficult to interpret as they are influenced by inter-individual differences in circadian period length, circadian phase, sleep duration and the duration of prior wakefulness, vulnerability to sleep loss, age, and personality.(Blatter & Cajochen 2007, Yoon et al. 1999) Several researchers examined circadian rhythms in visual thresholds.(Bassi & Powers 1986, O'Keefe & Baker 1987, Roenneberg et al. 1992, Tassi et al. 2000, Tuunainen et al. 2001) Although the confidence intervals were wide, retinal sensitivity seemed to be the lowest in the early morning. We found the lowest sensitivity directly after lunch and – in a subgroup of patients with early glaucoma – the highest sensitivity in the early morning. Therefore, another explanation could be that, in perimetry, the influence of cognitive performance dominates that of retinal sensitivity. Cognitive performance seems to be best in the early morning at an age comparable to that of glaucoma patients, although this is influenced by many factors.(Blatter & Cajochen 2007, Yoon et al. 1999) Lower performance in the afternoon is frequently observed and referred to as the "post lunch dip".(Monk 2005) Especially in the elderly who suffer more often from a short and fragmented nocturnal sleep pattern, the severe post lunch dip is counteracted by daytime afternoon naps.(Buysse et al. 1992) The lack of possibility to take an afternoon nap in patients if they have to visit the ophthalmology department may worsen performance during the afternoon.(Takahashi et al. 1998) In patients with moderate/severe glaucoma, the influence of time of day appeared to be less pronounced. This might be explained by a threshold effect, performance is always

bad, or by glaucoma itself which may disturb circadian rhythms by lesioning the non-image forming light-sensitive system.(Agorastoset al. 2011, Drouyer et al. 2008, Feigl et al. 2011, Jean-Louis et al. 2008, Lanzani et al. 2012)

Season

There is little information published on the influence of season on visual thresholds. Sweeney et al.(Bassi & Powers 1986, Sweeney et al. 1960) published in 1960 a paper showing that, in healthy subjects, scotopic sensitivities were lowest in summer and gradually increased until spring - roughly following daily exposure to sunlight or prior light history.(Hébert et al. 2002, Jasser et al. 2006, Smith et al. 2004) Bassi and Powers reported, in their study on circadian effects in healthy subjects, that subjects were slightly more sensitive during the winter.(Bassi & Powers 1986) We found the highest sensitivity in winter and spring – in agreement with the findings of Sweeney et al. (despite the fact that perimetry is not a scotopic task) and Bassi and Powers. (Sweeneyet al. 1960) Very recently, a seasonal influence on the MD was reported in patients with ocular hypertension. Typically 0.1 dB higher MD values were found in the winter compared to summer.(Gardiner et al. 2013) This is in agreement with our findings in patients with early glaucoma.

Pattern standard deviation

As a secondary analysis, we repeated our analyses with the PSD residuals instead of the MD residuals as dependent variable. These analyses were limited to patients with early glaucoma, as only for early glaucoma the PSD is monotonically related to disease stage. The significant effects on the PSD residuals we found (Figure 3) were roughly opposite to and generally smaller than the corresponding effects on the MD residuals (Figure 2; triangles). The opposite direction suggests – at least in early glaucoma - that the effects of the studied factors are larger in the diseased parts of the visual field than in the healthy parts. Obviously, this may change in more advanced disease - as indicated in Figure 2. Clinically, the smaller effects on the PSD residuals (Figure 3) in comparison with the corresponding effects on the MD residuals (Figure 2; triangles) tentatively indicate that the PSD is a more robust global index than the MD in early glaucoma.

Other issues

To weigh the clinical relevance of our findings, the effects found in this study should be compared to the rate of change of the MD due to glaucoma and to the overall variability. As mentioned in the first paragraph results section, the median MD decline in our study population was -0.2 dB/year. This indicates that the effects found in this study (together

typically 0.5 dB) are in the order of magnitude of the MD loss due to glaucoma after 2 to 3 years. Effects of typically 0.5 dB will generally not compromise the interpretation of a single test result. When a series of test results are analyzed as part of progression detection, however, the effects may play a significant role. After all, the effects are non negligible when compared to the mean (SD) overall variability, expressed as the square root of the residual mean square of the MD (dB), which was 1.1(0.7) dB in our study population (Chapter II).(Junoy Montolio et al. 2012) In the case of event detection, additional variability may result in mixing-up stability with suspected progression, or – after suspected progression has been observed – in mixing-up falsification with confirmation. In the case of trend analysis, additional variability will result in a longer period before a slope significantly different from zero can be detected,(Chauhan et al. 2008) and a less precise estimate of the actual slope can be made for a given follow-up.(Jansonius 2010) Another clinically useful message is that only the percentage of false-positive answers has to be taken into account.

In our clinical setting, we adopted the SITA fast strategy because it was considered a time-saving improvement of the SITA standard strategy at the time we designed the study (1999). Later, it became clear that the strategies performed slightly different. For example, more variability is found in SITA fast due to a higher error related factor (ERF) at the end of the test related to a shorter data acquisition time.(Artes et al. 2002, Jansonius 2010) Assuming that both SITA strategies reveal an unbiased estimate of the MD (albeit SITA fast with a higher variability), factors that influence the MD by influencing the physiological visual sensitivity should yield similar effects for both strategies. Here, it is interesting to note that Gardiner et al. found roughly similar seasonal effects, using the full threshold strategy, as we did with SITA fast (discussed above; Season subsection).(Gardiner et al. 2013) Factors that influence the MD through inattention or fatigue might have a more pronounced effect in longer strategies, like SITA standard. The effects of the reliability indices seem to be universal for SAP rather than specific to a single perimeter/strategy (discussed above; Reliability indices subsection).

The use of linear regression analysis is a common approach in glaucoma progression research.(Artes et al. 2002, Bengtsson & Heijl 2008, Bhandari et al. 1997) However, if a systematic deviation from a linear deterioration would exist, false-positive associations might pop-up in our analyses. This would occur for variables that are associated with the follow-up time point (for example, if the glaucoma service opening hours would change from AM to PM during the study). To adjust for this possible confounding, we added a variable named follow-up. We found a lower sensitivity during the first and third tertiles of the follow-up, indicating a systematic deviation from a linear

deterioration. This might be explained by prolonged learning(Gardiner et al. 2008, Wild et al. 1989) and/or the fact that glaucoma seems to have an accelerating character when using the MD as the outcome measure.(Wesselink et al. 2012) As a consequence, short series of visual fields might underestimate the rate of progression. We explored this possibility by comparing the first three years of follow-up to the entire follow-up in a subset of 104 patients with at least nine years of follow-up (mean [SD] follow-up 10.4[0.8] years). Table 3 shows the results. As can be seen in this table, the most obvious difference between the first three years and the entire follow-up is not a difference in the median rate of progression but rather a greater variability related to the shorter follow-up duration.(Jansonius 2010, Nouri-Mahdavi et al. 2011) This suggests that both the phenomenon of "positive slopes" and the phenomenon of "rapid progressors" seem to be at least partially related to too short series of visual fields.

Table 3. Rate of progression of the mean deviation (dB/year) as a function of follow-up duration (based on a subset of 104 patients with at least nine years of follow-up)

	First 3 years of follow-up	Entire follow-up period
P10	0.71	0.16
P25	0.24	-0.04
median	-0.17	-0.16
P75	-0.68	-0.40
P90	-1.29	-0.63

Another factor that might influence the MD is the development of cataract (or after cataract) or a cataract extraction (or capsulotomy) during follow-up. A change in MD slope due to a gradual development of cataract will not influence the residuals. Also, a deviation from a linear decay due to a faster development of cataract or a cataract extraction will not influence our findings. First, we adjusted also for these deviations with the variable named follow-up. Second, it is reasonable to assume that media opacity changes are not associated with the characteristics of the sampling (technician choice, time of day and season). Thus – albeit the development of cataract or a cataract extraction during follow-up may increase MD variability – they do not influence our findings.

Tests performed by a single patient cannot be considered independent observations. We addressed this issue by the use of a resampling method, by randomly taking one test result of each patient followed by the multiple linear regression analysis, and repeating

this 50 times. In multiple linear regression analysis, multiple hypothesis testing is already taken into account. It is not unequivocal if correction for multiple hypothesis testing would be needed with the resampling method applied in this study. However, the effects reported as being significant in this study all had P<0.001, indicating that they would remain significant even with very conservative corrections for multiple hypothesis testing. We also addressed the dependency issue by taking the reliability indices relative to their mean value in the concerning patient (see Methods section). Finally, effects can only be found if the independent variables display sufficient variability (for example, if all patients would visit the glaucoma service on a fixed time of day and year, no effects of time of day and year will be found). In a scientifically ideal situation, appointments would be allocated at random. As this is not the case in an observational study, the observed effects might be smaller than the true effects. In our department, especially time of day is prone to limited variability. Of all patients, 65% had at least one test in all four time of day strata, 31% had no tests in one stratum, 4% had no tests in two strata, and 0% had all tests in a single stratum. For season these percentages were 63, 34, 3 and 1%, respectively. These percentages suggest that there was sufficient within-subject variability for time of day and season to investigate their influences on perimetry test results.

Besides confounding, interactions between independent variables could also play a role. An interaction could be expected for example for time of day and season (there is little environmental light present in the early morning during the winter compared to summer). Interactions can be explored by stratification. If we stratified our analyses according to season instead of to stage (the latter was performed in Figure 2), we found essentially similar effects of time of day for all four seasons, indicating no obvious interaction between time of day and season. The power (number of included subjects) prohibited more thorough interaction analyses.

In conclusion, in a clinical setting where individual patients are tested at the same time of day while technicians pay attention to patient instruction and reassurance, monitor fixation and address apparent inattention, clinicians should only pay attention to the percentage of FP answers. Visual fields with more than 5% FP answers should be used for quantitative analyses only with caution. Compared to the MD, the PSD appeared to be less influenced by the studied factors and might thus be the preferred global index in early glaucoma.

REFERENCES

Agorastos, Agorastos A & Huber CG (2011): The role of melatonin in glaucoma: implications concerning pathophysiological relevance and therapeutic potential. J Pineal Res. 1–7.

Anderson RS, McDowell DR & Ennis FA (2001): Effect of localized defocus on detection thresholds for different sized targets in the fovea and periphery. Acta Ophthalmol Scand. 79: 60–63.

Artes PH, Iwase A, Ohno Y, Kitazawa Y & Chauhan BC (2002): Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. Invest Ophthalmol Vis Sci. 43: 2654–2659.

Bassi CJ & Powers MK (1986): Daily fluctuations in the detectability of dim lights by humans. Physiol Behav. 871–877.

Bengtsson B (2000): Reliability of computerized perimetric threshold tests as assessed by reliability indices and threshold reproducibility in patients with suspect and manifest glaucoma. Acta Ophthalmol Scand. 78: 519–522.

Bengtsson B & Heijl A (2000): False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? Am J Ophthalmol. 130: 689.

Bengtsson B & Heijl A (2008): A visual field index for calculation of glaucoma rate of progression. Am J Ophthalmol. 145: 343–353.

Bhandari A, Crabb DP, Poinoosawmy D, Fitzke FW, Hitchings RA & Noureddin BN (1997): Effect of Surgery on Visual Field Progression in Normal-tension Glaucoma. Ophthalmology. 104: 1131–1137.

Birt CM, Shin DH, Samudrala V, Hughes BA, Kim C & Lee D (1997): Analysis of Reliability indices from humphrey visual field tests in an urban glaucoma population. Ophthalmology. 104: 1126–1130.

Blatter K & Cajochen C (2007): Circadian rhythms in cognitive performance: methodological constraints, protocols, theoretical underpinnings. Physiol Behav. 90: 196–208.

Blumenthal EZ, Sample PA, Berry CC, Lee AC, Girkin CA, Zangwill L, Caprioli J & Weinreb RN (2003): Evaluating several sources of variability for standard and SWAP visual fields in glaucoma patients, suspects, and normals. Ophthalmology. 110: 1895–1902.

Buysse DJ, Browman KE, Monk TH, Reynolds CF, Fasiczka AL & Kupfer DJ (1992): Napping and 24-Hour sleep/wake patterns in healthy elderly and young adults. J American Geriatr Soc. 779–786.

Cascairo MA, Stewart WC & Sutherland SE (1991): Influence of missed catch trials on the visual field in normal subjects. Graefes Arch Clin Exp Ophthalmol. 229: 437–441.

Chauhan BC, Garway-Heath DF, Goni FJ, Rossetti L, Bengtsson B, Viswanathan AC & Heijl A (2008): Practical recommendations for measuring rates of visual field change in glaucoma. Br J Ophthalmol. 92: 569–573.

Drouyer E, Dkhissi-Benyahya O, Chiquet C, WoldeMussie E, Ruiz G, Wheeler LA, Denis P & Cooper HM (2008): Glaucoma alters the circadian timing system. PLoS One 3: e3931.

Feigl B, Mattes D, Thomas R & Zele AJ (2011): Intrinsically photosensitive (melanopsin) retinal ganglion cell function in glaucoma. Invest Ophthalmol Vis Sci. 52: 4362–4367.

Gardiner SK, Demirel S, Gordon MO & Kass MA (2013): Seasonal changes in visual field sensitivity and intraocular pressure in the ocular hypertension treatment study. Ophthalmology. 120: 724–730.

Gardiner SK, Demirel S & Johnson CA (2008): Is there evidence for continued learning over multiple years in perimetry? Optom and Vis Sci. 85: 1043–1048.

Hanifin JP & Brainard GC (2007): Photoreception for circadian, neuroendocrine, and neurobehavioral regulation. J Physiol Anthropol. 26: 87–94.

Hébert M, Martin SK, Lee C & Eastman CI (2002): The effects of prior light history on the suppression of melatonin by light in humans. J Pineal Res. 33: 198-203.

Heeg GP, Blanksma LJ, Hardus PLLJ & Jansonius NM (2005): The Groningen Longitudinal Glaucoma Study. I. Baseline sensitivity and specificity of the frequency doubling perimeter and the GDx nerve fibre analyser. Acta Ophthalmol Scand. 83: 46–52.

Heijl A, Patella VM & Bengtsson B (2012): The Field Analyzer Primer: Effective Perimetry. Carl Zeiss Incorporated.

Jansonius NM (2010): On the accuracy of measuring rates of visual field change in glaucoma. Br J Ophthalmol. 94: 1404–1405.

Jasser SA, Hanifin JP, Rollag MD & Brainard GC (2006): Dim light adaptation attenuates acute melatonin suppression in humans. J Biol Rhythms. 21: 394–404.

Jean-Louis G, Zizi F, Lazzaro DR & Wolintz AH (2008): Circadian rhythm dysfunction in glaucoma: A hypothesis. J Circadian Rhythms. 6: 1.

Johnson LN, Aminlari A & Sassani JW (1993): Effect of intermittent versus continuous patient monitoring on reliability indices during automated perimetry. Ophthalmology. 100: 76–84.

Junoy Montolio FG, Wesselink C & Jansonius NM (2012): Persistence, spatial distribution and implications for progression detection of blind parts of the visual field in glaucoma: a clinical cohort study. PLoS One. 7: e41211.

Katz J, Sommer A & Witt K (1991): Reliability of visual field results over repeated testing. Ophthalmology. 98: 70–75.

Kramer BC, Musch DC, Niziol LM & Weizer JS (2012): Reliability of simultaneous visual field testing. Ophthalmology. 119: 304–307.

Kutzko KE, Brito CF & Wall M (2000): Effect of instructions on conventional automated perimetry. Invest Ophthalmol Vis Sci. 41: 2006–2013.

Lanzani MF, de Zavalía N, Fontana H, Sarmiento MIK, Golombek D & Rosenstein RE (2012): Alterations of locomotor activity rhythm and sleep parameters in patients with advanced glaucoma. Chronobiol Int. 29: 911-919.

Lee M, Zulauf M & Caprioli J (1994): The influence of patient reliability on visual field outcome. Am J Ophthalmol. 117: 756–761.

Markwell EL, Feigl B & Zele AJ (2010): Intrinsically photosensitive melanopsin retinal ganglion cell contributions to the pupillary light reflex and circadian rhythm. Clin Exp Optom. 93: 137–149.

Monk TH (2005): The post-lunch dip in performance. Clin Sports Med. 24: e15–23, xi–xii.

Mutlukan E (1994): The effect of refractive blur on the detection sensitivity to light offsets in the central visual field. Acta Ophthalmol. 72: 189–194.

Newkirk MR, Gardiner SK, Demirel S & Johnson CA (2006): Assessment of false positives with the Humphrey Field Analyzer II perimeter with the SITA Algorithm. Invest Ophthalmol Vis Sci. 47: 4632–4637.

Nouri-Mahdavi K, Zarei R & Caprioli J (2011): Influence of visual field testing frequency on detection of glaucoma progression with trend analyses. Arch Ophthalmol. 129: 1521–1527.

O'Keefe LP & Baker HD (1987): Diurnal changes in human psychophysical luminance sensitivity. Physiol & Behav. 41: 193–200.

Roenneberg T, Lotze M & Steinbuchel N (1992): Diurnal variation in human visual sensitivity determined by incremental thresholds. Clin Vis Sci. 7: 83–91.

Smith KA, Schoen MW & Czeisler CA (2004): Adaptation of human pineal melatonin suppression by recent photic history. J Clin Endocrinol & Metab. 89: 3610–3614.

Sweeney EJ, Kinney JA & Ryan A (1960): Seasonal changes in scotopic sensitivity. J Opt Soc Am. 50: 237–240.

Takahashi M, Fukuda H & Arito H (1998): Brief naps during post-lunch rest: effects on alertness, performance, and autonomic balance. Eur J Appl Physiol Occup Physiol. 78: 93–98.

Tassi P, Pellerin N, Moessinger M, Eschenlauer R & Muzet A (2000): Variation of visual detection over the 24-hour period in humans. Chronobiol Int. 17: 795–805.

Tuunainen A, Kripke DF, Cress AC & Youngstedt SD (2001): Retinal circadian rhythms in humans. Chronobiol Int 18: 957–971.

Van Coevorden RE, Mills RP, Chen YY & Barnebey HS (1999): Continuous visual field test supervision may not always be necessary. Ophthalmology. 106: 178–181.

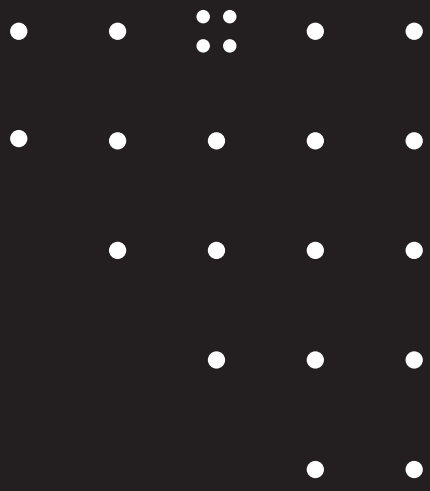
Wesselink C, Heeg GP, Jansonius NM (2009): Glaucoma Monitoring in a Clinical Setting: glaucoma progression analysis vs nonparametric progression analysis in the Groningen Longitudinal Glaucoma Study. Arch Ophthalmol. 127: 270-274.

Wesselink C, Marcus MW & Jansonius NM (2012): Risk factors for visual field progression in the groningen longitudinal glaucoma study: a comparison of different statistical approaches. J Glaucoma. 21: 579–585.

Wild JM, Dengler-Harles M, Searle AET, O'Neill EC & Crews SJ (1989): The influence of the learning effect on automated perimetry in patients with suspected glaucoma. Acta Ophthalmol. 67: 537–545.

Yoon C, May CP & Hasher L (1999). Aging, circadian arousal patterns, and cognition. Cognition, Aging, and Self-Reports. 117–117.

IV



Influence of multifocal intraocular lenses on standard automated perimetry test results

Nancy Aychoua*, Francisco G. Junoy Montolio*, Nomdo M. Jansonius

JAMA Ophthalmology. 2013;131(4):481-5

** = contributed equally to this manuscript*

ABSTRACT

Purpose: To evaluate the influence of a multifocal intraocular lens (MFIOL) on standard automated perimetry (SAP) size III and size V test results

Methods: In this cross-sectional case-control study sixteen eyes of 16 patients with a diffractive MFIOL (median age, 64 years), 18 phakic eyes of 18 healthy individuals serving as controls (median age, 62 years), and 12 eyes of 12 patients with a monofocal IOL (median age, 64 years) were included. All participants underwent (1) SAP using a 30-2 grid and the Swedish Interactive Threshold Algorithm standard strategy with stimulus size III and (2) a full threshold test with stimulus size V. Comparisons between groups were corrected for age and pupil size.

Results: For SAP size III, the average difference in mean deviation (MD) between patients in the MFIOL group and phakic controls was -2.40 dB (P<0.001) and between patients in the monofocal IOL group and phakic controls was -0.32 dB (P=0.52). For SAP size V, the corresponding differences in mean sensitivity (MS) were -1.61 dB (P=0.002) and -0.80 (P=0.09), respectively. The differences were essentially independent of eccentricity for both SAP size III and SAP size V.

Conclusion: Patients with a diffractive MFIOL have a clinically relevant reduction of the visual sensitivity as assessed with SAP size III and size V. The reduction seems to be related to the multifocal design of the IOL rather than to pseudophakia. The reduction interferes with the assessment of common eye diseases such as glaucoma and comes on top of the decline of visual sensitivity due to normal aging or age-related eye diseases, thus potentially accelerating visual impairment.

INTRODUCTION

Cataract surgery is one of the most established interventions in the realm of ophthalmology and is considered to be a safe procedure with good visual outcome. In pursuit of optimal postoperative lens performance, generations of innovative lens designs have led to the introduction of diffractive multifocal intraocular lenses (MFIOLs).(Hansen et al. 1990) These lenses allow the luxury of distance and near spectacle independence and are based on the optical principle of creating 2 focal planes, one for distance vision and the other for near vision. Several studies reported that these new-degeneration diffractive MFIOLs provide satisfactory vision for distance and near conditions and enhance the quality of life.(Cillino et al. 2008, Nowak & Jacobi 1990, Vries et al. 2008)

Nonetheless, MFIOLs appear to induce unwanted visual phenomena, including glare, flare, streaks, and halos.(de Vries et al. 2011, Leyland & Pringle 2006, Leyland & Zinicola 2003) Reduced contrast sensitivity has also been reported in various studies as an important drawback of MFIOLs.(Allen et al. 1996, Haaskjold et al. 1998, Leyland & Pringle 2006) Contrast sensitivity usually declines with age and may be further affected pathologically by different eye diseases.(Nio et al. 2000) Glaucoma, for example, is an insidious chronic progressive eye disease that affects the optic nerve with subsequent unnoticed retinal sensitivity loss. This loss in sensitivity is clinically measured by means of standard automated perimetry (SAP), which measures differential light sensitivity thresholds at various locations across the visual field and is a fundamental tool in the diagnosis and follow-up of glaucoma. Taking the various causes of reduced contrast sensitivity, as well as the possibility of MFIOL and glaucoma coexistence, into consideration, it is therefore justifiable to theorize that MFIOLs may hinder the assessment of glaucoma.(Kumar et al. 2007)

The influence of MFIOLs on perimetry has been addressed using matrix frequency-doubling perimetry, Goldmann perimetry, the Esterman binocular visual field test, and the Octopus 101 (see "Discussion" section).(Bi et al. 2008, Bojikian et al. 2009, Kang & Lee 1994, Stanojcic et al. 2010)

The aim of the present study was to determine the influence of MFIOLs on SAP test results. For this purpose, we performed a cross-sectional case-control study comparing patients with MFIOLs, patients with monofocal IOLs, and individuals with phakic eyes serving as controls. By comparing 3 groups, it should be possible to differentiate between effects due to pseudophakia and effects due to the multifocal design. In addition to SAP with the default size III stimulus, the participants performed a test with

stimulus size V.(Wall et al. 2010, Wilensky et al. 1986) Our rationale for this additional test was to explore the relationship between stimulus size and the effect of MFIOL implantation on the visual sensitivity as assessed with perimetry.

METHODS

Study population

The present study had a cross-sectional case-control design and included patients with MFIOLs (MFIOL group), patients with monofocal IOLs (monofocal IOL group), and healthy individuals with phakic eyes (phakic controls). This study conformed with the tenets of the Declaration of Helsinki and was approved by the medical ethics committee of the University Medical Center of Groningen, Groningen, the Netherlands. All participants gave written informed consent prior to participation.

Patients with MFIOLs and monofocal IOLs were recruited from the cataract databases of the departments of ophthalmology of the University Medical Center Groningen and of the Nij Smellinghe Hospital Drachten, both in the Netherlands. Healthy volunteers were recruited through advertisement. All participants underwent a complete eye examination, including best corrected visual acuity testing, near vision testing (Jaeger reading chart), slit lamp biomicroscopy, intraocular pressure measurement with noncontact tonometry (TCT80; Topcon Medical Systems Inc), and fundus examination with an ultra widefield retinal imaging device (200TX ultra-widefield retinal image; Optos). The pupil diameter was measured by means of the Auto Pupil function of the Humphrey Field Analyzer (see the "Perimetry" subsection).

Inclusion criteria for this study were age 18 to 75 years and, for the MFIOL and monofocal IOL groups, a postoperative period of at least 3 months. Exclusion criteria were an overall astigmatism exceeding 2.5 diopter (D), a spherical equivalent refractive error above 5 or below −5 D, a best-corrected visual acuity above 0.0 logMAR (in individuals 50 years) or 0.1 logMAR (50 years), an intraocular pressure above 21 mmHg, a family history of glaucoma, a vertical cup-disc ratio exceeding 0.5 or any other fundus abnormality, significant lens opacities or after cataract on slit lamp examination, and a history of eye trauma or surgery other than cataract surgery or any other eye disease. All participants were inexperienced with regard to perimetry. If a participant was eligible with both eyes, 1 randomly chosen eye was included.

IOL characteristics

The MFIOLs in this study comprised exclusively diffractive MFIOLs. Two types of diffractive MFIOLs were used in our study population (Tecnis ZM900; Abbott Medical Optics Inc [2 eyes]; and the Zeiss 809M, AT LISA; Carl Zeiss Meditec Inc [14 eyes]). The Tecnis MFIOL is a silicone 3-piece aspheric diffractive lens. The power of the add is 4.00 D with a 50/50 distance/near light distribution. The Zeiss MFIOL is an acrylic single-piece aspheric diffractive lens. The power of the add is 3.75 D with a 65/35 light distribution. All the patients with monofocal IOLs had a monofocal Tecnis (ZA9003; Abbott Medical Optics Inc).

Perimetry

Perimetry was performed with the Humphrey Field Analyzer (Carl Zeiss Meditec Inc). First, all participants performed a shortened visual field test, consisting of 15 test locations distributed over a 30-2 grid (Figure 1) using the 4-2-2 staircase strategy with size III stimulus (0.43 mm diameter; 4 mm2). This test was conducted so that participants could become accustomed to the perimeter but would not be tired before the onset of testing. Subsequently, participants performed a 30-2 Swedish Interactive Threshold Algorithm (SITA) standard test with stimulus size III. Finally, participants performed another shortened visual field test with 15 test locations twice, now with stimulus size V (1.72 mm diameter; 64 mm2). Regarding the 15 test locations, 7 were within 10° eccentricity (Figure 1). We used a shortened test because no SITA program is available for size V and we aimed to avoid fatigue effects resulting from the lengthy full-threshold testing. The SITA standard size III test and the second shortened test with stimulus size V were included in the analysis.

The recommendations of the Humphrey Field Analyzer's manufacturer for using corrective lenses were followed. No corrective lenses were used in patients with MFIOLs unless the near-vision test showed a value below Jaeger 2, which could be improved with corrective lenses at the recommended testing distance of 33 cm. A test result was considered unreliable if false-positive classifications exceeded 10%, the technician reported poor fixation, or lens rim artifacts were observed. If this was the case, the test was repeated after additional explanation. At least 5 minutes of rest was scheduled between the different tests to lower the influence of fatigue.

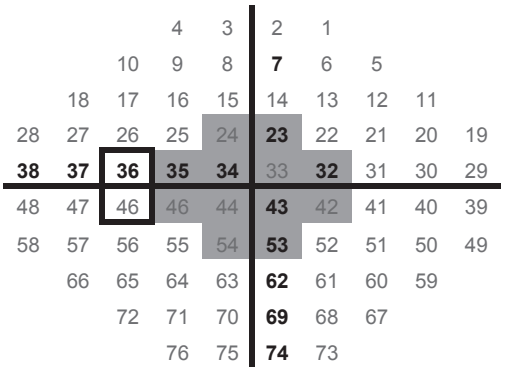


Figure 1. Subset of test locations used in the shortened tests (numbers in bold). Including the fovea (not shown), there were 7 locations within 10 degrees eccentricity (left eye representation)

Statistical analysis

The main outcome measures of the visual field tests were the mean deviation (MD) for size III and the mean sensitivity (MS) for size V. The MD is a measure commonly used in clinical practice; the MS can be considered a proxy of the MD. The MD is an age-adjusted measure (a study participant is compared with age-matched peers) provided by the software of the Humphrey Field Analyzer. For size V, we used a customized grid (see the "Perimetry" section); therefore, no MD value was provided by the device. For that reason we calculated the MS by averaging the recorded raw sensitivities of the included test locations. We also used the MS to explore the effect of eccentricity for both size III and size V by calculating the MS for a subset of test locations within and outside 10° eccentricity (Figure 1). The blind spot and the fovea were excluded from the MS calculation.

Differences in characteristics between the MFIOL group, monofocal IOL group, and phakic controls were analyzed with 1-way analysis of variance for continuous variables and the 2 test for proportions. Because of our recruitment approach (advertisement), the median age of our originally recruited 45 phakic controls (49 years) was considerably lower than that of the MFIOL group (64 years). To exclude any residual age-related confounding, we performed the analysis on an age-matched subset of phakic controls. This subset was formed by excluding controls, starting with the youngest, until the mean age of the controls equaled that of the MFIOL group. The MD and MS intergroup differences were assessed using multiple linear regression analysis including either the MD or the MS as the dependent variable and age, pupil

size, and group as independent variables. For group differences, we used 2 dummy variables for the MFIOL group and monofocal IOL group, with the phakic controls as the reference.

All analyses were performed using commercial software (SPSS, version 18.0.3; SPSS Inc). Statistical significance was set at $P \leq 0.05$.

RESULTS

The Table presents the participants' characteristics. Sixteen patients with MFIOLs, 18 phakic controls, and 12 patients with monofocal IOLs were enrolled in this study. No eyes showed relevant pathologic characteristics. There were significant univariate intergroup differences in pupil diameter, best-corrected visual acuity, MD, MS, and pattern standard deviation. There were no significant intergroup differences in age, sex, and intraocular pressure. As for near vision, all participants attained Jaeger 2 at 33 cm, with or without correction. In the MFIOL group, 4 of 16 patients needed additional correction (0.75, 1.00, 2.00, and 2.50 D).

Figure 2 depicts the differences in unadjusted MD between the MFIOL group, the monofocal IOL group, and the phakic controls when conducting visual field testing with SAP size III. Adjusted for age and pupil size, the MD was, on average, 2.40 dB lower in the MFIOL group than in the phakic controls (P=0.001) and 0.32 dB lower in the monofocal IOL group than in the phakic controls (P=0.52).

Figure 3 illustrates the intergroup differences in un adjusted MS for size V. Adjusted for age and pupil size, the MS was, on average, 1.61 dB lower in the MFIOL group than in the phakic controls (P=0.002) and 0.80 dB lower in the monofocal IOL group than in the phakic controls (P=0.09).

For the subset of test locations within 10° eccentricity, the age- and pupil size–adjusted difference in MS between the MFIOL group and the phakic controls was −2.28 dB for size III (P=0.001) and −1.87 dB for size V (P=0.003). For the monofocal IOL group vs the phakic controls, these differences were 0.15 dB (P=0.77) and −0.39 dB (P=0.49), respectively. For the subset of test locations outside 10° eccentricity, the age and pupil size–adjusted differences in MS between the MFIOL group and phakic controls were −2.49 dB ($P \leq 0.001$) for size III and −1.27 dB (P=0.01) for size V. For the monofocal IOL group vs the phakic controls, these differences were −0.32 dB (P=0.59) and −1.29 dB (P=0.01), respectively. The age- and pupil size–adjusted foveal sensitivity of the SAP

Table 1. Participant characteristics

	MFIOL patients	Phakic controls	Monofocal IOL patients	*P*-value
Number of eyes	16	18	12	
Age (years; median [range])	64 (47 to 74)	62 (53 to 74)	64 (48 to 71)	0.82
Gender (male/female)	6/10	9/9	7/5	0.64
Pupil diameter (mm; median [range])	3.9 (2.9 to 5.2)	5.2 (3.8 to 7.3)	5.2 (3.8 to 6.2)	0.001
IOP (mmHg; median [range])	14 (8 to 19)	14 (8 to 21)	11 (8 to 16)	0.09
BCVA (logMAR; median [range])	0.0 (0.1 to -0.1)	0.0 (0.0 to -0.2)	-0.1 (0.0 to -0.2	0.003
Size III MD (dB; median [range])	-3.0 (-5.4 to -0.7)	-0.5 (-2.1 to +2.0)	-1.0 (-1.8 to +0.3)	<0.001
Size III PSD (dB; median [range])	2.2 (1.5 to 4.2)	1.8 (1.3 to 3.5)	2.0 (1.5 to 2.6)	0.012
Size V MS (dB; median [range])	33.2 (30.9 to 35.8)	34.9 (33.0 to 37.2)	34.2 (32.9 to 35.5)	<0.001
Size III MS within 10° eccentricity	29.5 (26 to 31)	31.0 (29 to 33)	31.5 (30 to 32)	<0.001
Size III MS outside 10° eccentricity	25.5 (22 to 28)	28.0 (26 to 31)	27.5 (26 to 29)	<0.001
Size V MS within 10° eccentricity	34.0 (30 to 37)	36.0 (33 to 38)	35.5 (34 to 37)	0.001
Size V MS outside 10° eccentricity	32.0 (30 to 36)	34.0 (32 to 36)	33.0 (32 to 34)	0.002

MFIOL = multifocal intraocular lens; IOP = intraocular pressure; BCVA = best corrected visual acuity; MD = mean deviation; PSD = pattern standard deviation; MS = mean sensitivity.

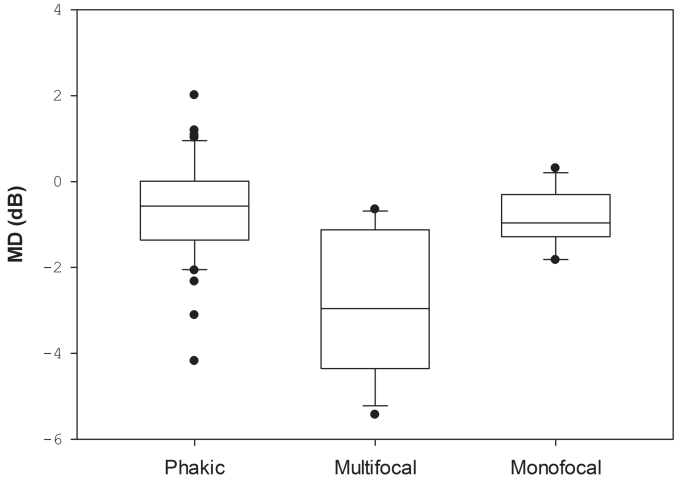


Figure 2. Effect of lens status on the mean deviation (MD) of standard automated perimetry with size III stimulus. Box plots show median, interquartile range, and each outlier. IOL indicates intraocular lens; MFIOL, multifocal IOL.

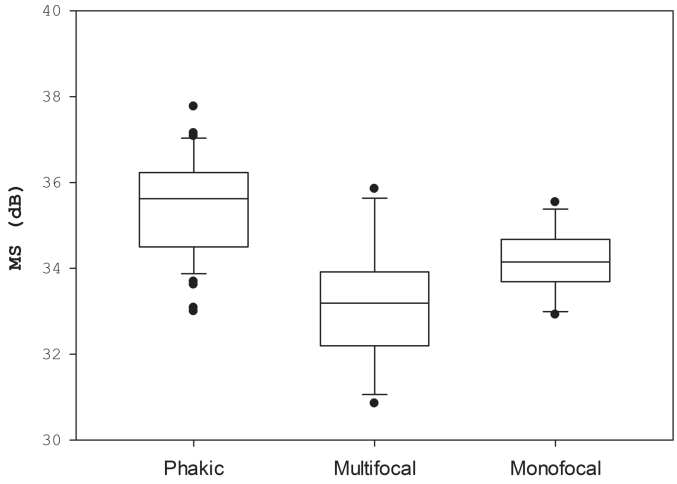


Figure 3. Effect of lens status on the mean sensitivity (MS) of size V perimetry. Box plots show median, interquartile range, and each outlier. IOL indicates intraocular lens; MFIOL, multifocal IOL.

test results was a mean of 2.05 dB lower in the MFIOL group than in the phakic controls (P=0.006) and 0.30 dB greater in the monofocal IOL group than in the phakic controls (P=0.65). The age- and pupil size–adjusted pattern standard deviation of the SAP test results did not differ significantly between the MFIOL group and the phakic controls (P=0.07) or between the monofocal IOL group and the phakic controls (P=0.70). These findings indicate that the effect of a MFIOL on visual sensitivity is essentially similar across the entire visual field; this agrees with a subjective assessment of the visual fields. Approximately 50% of the MFIOL visual fields showed normal total and pattern deviation probability plots; the other 50% showed a general reduction of sensitivity picture: diffuse abnormalities in the total deviation probability plot combined with an intact pattern deviation probability plot. Eleven of the 16 patients with MFIOLs had an MD value below normal, at the P<0.05 level according to the Humphrey Field Analyzer database.

DISCUSSION

Multifocal intraocular lenses reduce the visual sensitivity in SAP, by approximately 2 dB. This reduction is roughly similar for size III and size V and regardless of eccentricity. The reduction seems to be related to the multifocal design of the IOLs rather than to pseudophakia.

To our knowledge, no previous study has evaluated the effect of diffractive MFIOLs on SAP compared with healthy controls. However, various studies have evaluated the influence of MFIOLs on other perimetric tests in comparison with monofocal IOLs. Bojikian et al. reported that diffractive MFIOLs have no influence on the MD of the Matrix frequency doubling perimeter compared with monofocal IOLs.(Bojikian et al. 2009) Their results do not contradict our findings because frequency doubling perimetry uses a stimulus with a very low spatial frequency and is, compared to SAP, less sensitive to optical blur.(Artes et al. 2003, Maddess et al. 1999) Interestingly, Kang and Lee found a significant difference between MFIOLs and monofocal IOLs using Goldmann kinetic perimetry.(Kang & Lee 1994) Stanojcic et al. assessed the difference in binocular visual fields in patients who underwent bilateral cataract surgery with either diffractive MFIOLs or monofocal IOLs by means of the Esterman binocular visual field test.(Stanojcic et al. 2010) They reported no significant difference between their 2 groups. However, it is not possible to compare our results with those of Stanojcic et al. because the Esterman test is based on suprathreshold testing, whereas we tested visual sensitivity at threshold. Bi et al. compared a group of patients who received the AcrySof ReSTOR MFIOL (SA60D3; Alcon Laboratories Inc) with a control group receiving

the AcrySof Natural (SN60AT) monofocal IOL following cataract surgery.(Bi et al. 2008) They performed comparisons in visual acuity, depth of focus, corneal astigmatism, contrast sensitivity, glare sensitivity, visual fields, and spherical aberration. Perimetry was performed using the Octopus 101. The investigators found no significant difference in the MD between groups. This result, which apparently conflicted with our study, might be explained by differences in threshold algorithm, which was not specified in their study.

As for the influence that monofocal IOLs may exhibit on SAP, Mutlu et al. found a significant negative effect of monofocal IOLs on the MD, namely, 1.34 dB lower compared with age-matched phakic subjects.(Mutlu et al. 2009) At first sight, this suggests that about half the negative effect on the MD we found in patients with MFIOLs could be an IOL effect not specific of the multifocal design. However, we found essentially no differences between the monofocal IOL group and the phakic controls. A possible explanation for this apparent contradiction is that the monofocal IOLs as assessed by Mutlu et al. were spherical, whereas the IOLs used in our study were aspheric. There are significant differences in modulation transfer between aspheric and spherical IOLs; the modulation transfer of an aspheric IOL is approximately equal to that of the human lens, whereas it is higher than that of a spherical IOL.(Jansonius 2010) A 1.34 dB difference in MD between eyes with a spherical monofocal IOL and phakic eyes is in line with the differences in modulation transfer as reported in the literature.(Jansonius 2010) This tentatively suggests that the multifocal design degrades the MD by about 2 dB, as we found in our study. Two studies evaluated the effect of cataract extraction on the visual fields of patients with glaucoma after monofocal IOL implantation. (Carrillo 2005, Smith 1997) Both studies reported modest to negligible improvement in MD postoperatively. Obviously, the effects of the removal of the cataract and the implantation of a monofocal IOL on the MD are entangled here.

Intergroup differences in pupil size were encountered in our study, and this makes pupil size a potential confounder.(Artigas et al. 2007, Lindenmuth et al. 1989) To avoid confounding, our multiple linear regression analyses were adjusted for pupil size. In these analyses, pupil size was not significant. Thus, although our study included eyes with a significant intergroup difference in pupil size, this difference does not explain our findings.

Our study included 2 different types of diffractive MFIOLs: the Tecnis (ZM900; Abbott Medical Optics Inc; 2 eyes) and the Zeiss 809M, AT LISA; Carl Zeiss Meditec Inc, 14 eyes). Despite the fact that there is a difference in distance/near light distribution between these 2 MFIOL types (50/50 for Tecnis and 65/35 for Zeiss), the MD values of

the 2 participants with a Tecnis lens (–1.44 and –3.05 dB) fall within the interquartile range of the MFIOL group as a whole. This tentatively suggests that the reduced visual sensitivity is a generic property of diffractive MFIOLs rather than specific for one type. Obviously, the sample size was too small for a decent subgroup analysis.

In conclusion, the results of this study suggest that eyes with MFIOLs will show reductions in visual sensitivity of approximately 2 dB. Therefore, we recommend a new perimetric baseline in patients with MFIOLs with (suspect) glaucoma and preferably in all patients with MFIOLs to guarantee a correct interpretation of any future abnormality. Moreover, patients should be preoperatively informed about MFIOL-related loss of visual sensitivity that may show a further decline with normal aging or age-related eye diseases. As a consequence, the originally highly appreciated spectacle independence might be regretted later.

REFERENCES

Allen ED, Burton RL, Webber SK, Haaskjold E, Sandvig K, Jyrkkiö H, Leite E, Nyström A & Wollensak J. (1996): Comparison of a diffractive bifocal and a monofocal intraocular lens. J Cataract & Refract Surg. 22: 446-451.

Artes PH, Nicolela MT, McCormick TA, LeBlanc RP & Chauhan BC (2003): Effects of blur and repeated testing on sensitivity estimates with frequency doubling perimetry. Invest Ophthalmol Vis Sci. 44: 646–652.

Artigas JM, Menezo JL, Peris C, Felipe A & Díaz-Llopis M (2007): Image quality with multifocal intraocular lenses and the effect of pupil size. J Cataract & Refract Surg. 33: 2111–2117.

Bi H, Cui Y, Ma X, Cai W, Wang G, Ji P & Xie X (2008): Early clinical evaluation of AcrySof ReSTOR multifocal intraocular lens for treatment of cataract. Ophthalmologica. 222: 11–16.

Bojikian KD, Vita JB, Forno CPD, Neto AT & de Moura CR (2009): Does the apodized diffractive intraocular lens Acrysof ReSTOR NaturalTM interfere with FDT Matrix perimetry results? Arq Bras Oftalmol. 72: 755–759.

Carrillo MM (2005): Effect of cataract extraction on the visual fields of patients with glaucoma. Arch Ophthal. 123: 929-932.

Cillino S, Casuccio A, Di Pace F, Morreale R, Pillitteri F, Cillino G & Lodato G (2008): One-year outcomes with new-generation multifocal intraocular lenses. Ophthalmology. 115: 1508–1516.

de Vries NE, Webers CAB, Touwslager WRH, Bauer NJC, de Brabander J, Berendschot TT & Nuijts RMMA (2011): Dissatisfaction after implantation of multifocal intraocular lenses. J Cataract Refract Surg. 37: 859–865.

Haaskjold E, Allen ED, Burton RL, Webber SK, Sandvig KU, Jyrkkiö H, Leite E, Liekfeld A, Philipson B, Nyström A & Wollensak J. (1998): Contrast sensitivity after implantation of diffractive bifocal and monofocal intraocular lenses. J Cataract Refract Surg. 24: 653–658.

Hansen TE, Corydon L, Krag S & Thim K (1990): New multifocal intraocular lens design. J Cataract Refract Surg. 16: 38–41.

Jansonius NM (2010): Spherical aberration and other higher-order aberrations in the human eye: from summary wave-front analysis data to optical variables relevant to visual perception. J Opt Soc Am A Opt Image Sci Vis. 27: 941-950.

Kang SG & Lee JH (1994): The change of visual acuity and visual field by diminished illumination in eyes with multifocal intraocular lens. Korean J Ophthalmol. 8: 72-76.

Kumar BV, Phillips RP & Prasad S (2007): Multifocal intraocular lenses in the setting of glaucoma. Curr Opin Ophthalmol. 18: 62–66.

Leyland M & Pringle E (2006): Multifocal versus monofocal intraocular lenses after cataract extraction. Cochrane Database Syst Rev. 12: CD003169

Leyland M & Zinicola E (2003): Multifocal versus monofocal intraocular lenses in cataract surgery. Ophthalmology. 110: 1789–1798.

Lindenmuth KA, Skuta GL, Rabbani R & Musch DC (1989): Effects of pupillary constriction on automated perimetry in normal eyes. Ophthalmology. 96: 1298–1301.

Maddess T, Goldberg I, Dobinson J, Wine S, Welsh AH & James AC (1999): Testing for glaucoma with the spatial frequency doubling illusion. Vision Res. 39: 4258–4273.

Mutlu FM, Akay F & Bayer A (2009): Effect of pseudophakia on standard perimetry parameters. Curr Eye Res. 34: 711–716.

Nio YK, Jansonius NM, Fidler V, Geraghty E, Norrby S & Kooijman AC (2000): Age-related changes of defocus-specific contrast sensitivity in healthy subjects. Ophthalmic Physiol Opt. 20: 323–334.

Nowak MR & Jacobi KW (1990): Diffractive multifocal intraocular lenses. A prospective clinical study. Klin Monbl Augenheilkd. 196: 43–47.

Smith SD (1997): Effect of cataract extraction on the results of automated perimetry in glaucoma. Arch Ophthalmol. 115: 1515-1519.

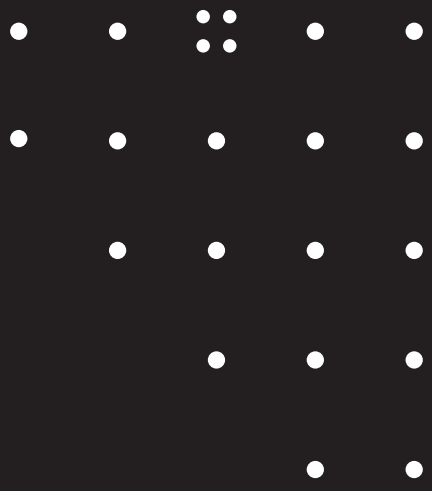
Stanojcic N, Wilkins M, Bunce C & Ionides A (2010): Visual fields in patients with multifocal intraocular lens implants and monovision: an exploratory study. Eye. 24: 1645–1651.

Vries NE de, Webers CAB, Montés-Micó R, Tahzib NG, Cheng YY, de Brabander J, Hendrikse F, Nuijts RMMA. (2008): Long-term follow-up of a multifocal apodized diffractive intraocular lens after cataract surgery. J Cat Refract Surg. 34: 1476-1482.

Wall M, Woodward KR, Doyle CK & Zamba G (2010): The effective dynamic ranges of standard automated perimetry sizes III and V and motion and matrix perimetry. Arch Ophthalmol. 128: 570–576.

Wilensky JT, Mermelstein JR & Siegel HG (1986): The use of different-sized stimuli in automated perimetry. Am J Ophthalmol. 101: 710–713.

V



Influence of glaucoma surgery on visual function:

A clinical cohort study and meta-analysis

Francisco G. Junoy Montolio, Rogier P.H.M. Müskens, Nomdo M. Jansonius

2018 Acta Ophthalmologica

ABSTRACT

Purpose: To determine the cost (loss of visual function associated with the procedure) and benefit (long-term preservation of the visual field) of glaucoma surgery.

Methods: We included 100 patients who underwent glaucoma surgery (Baerveldt glaucoma implant [BGI], n=61; trabeculectomy [TE], n=39). Preoperatively, the median (interquartile range [IQR]) standard automated perimetry mean deviation (MD) was -12(-16 to -6) dB. We analysed the change in visual acuity (BCVA) and MD due to the procedure and, in a subset with at least 5 years of perimetric follow-up both pre- and postoperatively (n=20), the change in rate of progression (ROP; time rate of change of MD). For the surgery induced change in ROP we also performed a meta-analysis including the current and previously published studies. From the surgery induced decrease in MD and change in ROP, we calculated the average postoperative duration needed for the benefit to surpass the cost.

Results: Mean (standard deviation) MD decline was 1.3(2.7) and 1.0(2.3) dB for BGI (P<0.001) and TE (P=0.009), respectively; no significant surgery induced changes in BCVA were found (P=0.08 and P=0.12, respectively). In our study, surgery was associated with a non-significant deceleration of the ROP (from -0.37[0.52] to -0.15[0.48] dB/year; P=0.23). The meta-analysis, based on eight studies, showed an overall surgery-induced change in ROP of 0.44(95% confidence interval 0.25 to 0.64; P<0.001) dB/year.

Conclusions: Glaucoma surgery significantly reduces the progression velocity in glaucoma. On average, the benefit of glaucoma surgery surpasses the cost after approximately 1.5 years.

INTRODUCTION

Glaucoma is one of the leading causes of avoidable blindness, characterized by a degeneration of the optic nerve and corresponding visual field loss. To slow down this chronic, progressive eye disease, treatment is focused on the lowering of the intraocular pressure (IOP) by restricting the production of aqueous or by increasing its outflow. The initial approach to achieve a lower IOP is the daily use of IOP-lowering drugs and/or one or more laser treatments. Surgery is generally considered after impermissible progression of visual field loss despite maximal tolerable medical and laser treatment. Glaucoma surgery is not without risk and may in itself compromise visual function. It is a clinical challenge to weigh the risk and benefit in individual patients.

Within glaucoma surgery, the creation of a fistula (filtering surgery) or implanting a drainage system have earned their spurs. Although invented in the same time period, filtering surgery (trabeculectomy; TE) was the first surgical choice for decades. Nowadays, the popularity of glaucoma drainage devices is growing).(Arora et al. 2015, Gedde et al. 2012, Islamaj et al. 2018) The success of surgery might come with a cost, however, as patients often complain of a decrease in visual function after surgery. This observation gives the clinician the sense of a drawback of surgery. However, the literature showed all three possible outcomes: an improvement in visual function, a decline, and no effect.(Balekudaru et al. 2014, Caprioli et al. 2016, Hagiwara et al. 2000, Sehi et al. 2010) The crucial question is, if there is a surgery-induced loss, if it counterbalances the loss avoided due to the IOP lowering. The IOP lowering aims to yield a lower rate of progression (ROP), that is, a less negative time rate of change of the standard automated perimetry (SAP) mean deviation (MD) – thus delaying blindness.

The aim of this study was to determine the cost and benefit of glaucoma surgery regarding visual function. For this purpose we analysed the change in visual acuity and SAP MD due to the procedure (the cost) and the presumed decline in rate of progression (ROP), that is, the time rate of change of the MD (the benefit). These analyses were performed primarily in the cohort of the Groningen Longitudinal Glaucoma Study (GLGS).(Heeg et al. 2005) For the surgery-induced change in ROP we also performed a meta-analysis including the current and previously published studies. From the surgery-induced decrease in MD and change in ROP, we calculated the average postoperative duration needed for the benefit to surpass the cost.

METHODS

Study population

The present study is part of the GLGS, a prospective observational cohort study performed in a clinical setting. The GLGS started in 2000 and is still ongoing. The ethics board of the University Medical Center Groningen (UMCG) approved that for the current study no informed consent had to be obtained because the study comprised a retrospective anonymous analysis of ophthalmic examination and visual field data collected during regular glaucoma care. The study followed the tenets of the Declaration of Helsinki.

The subpopulation selected for the present study consisted of patients who were treated with a Baerveldt glaucoma implant (BGI) or TE and who had at least (1) two (not including the first visual field, which was discarded because of learning effects) reliable visual fields (measured with SAP [Humphrey Field Analyzer 30-2 SITA fast; Carl Zeiss, Jena, Germany]) both pre- and post-intervention and (2) a follow-up of at least one year after the intervention, in the operated eye. If both eyes met the inclusion criteria, a random eye was chosen (based on even or uneven case number in the database). For the analysis concerning the ROP, a perimetric follow-up of at least five years was required both pre- and post-intervention.(Jansonius 2010, Junoy Montolio et al. 2012) Visual fields had to be reliable. A test result was considered unreliable if false positives exceeded 10% or if both false negatives and fixation losses exceeded 10% and 20%, respectively. We pooled false negatives and fixation losses because they were reported to have a much smaller influence on the MD(Junoy Montolio et al. 2012) and especially the false negative are not informative in glaucoma.(Bengtsson & Heijl 2000) Both primary and secondary glaucoma was allowed. We included both phakic and pseudophakic patients; patients who underwent a cataract extraction simultaneously with the glaucoma surgery were excluded. We also excluded patients who underwent a second glaucoma operation during follow-up.

Surgical procedure

Indication for surgery and method (BGI or TE) was made by either of the two authors who also performed the operations (RM and NJ). In general, in our hospital surgery is delayed until we observe progression too rapid for the age of the patient,(Wesselink et al. 2011) with maximum tolerable medical treatment and after considering or performing laser surgery. Early surgery is considered in cases with high baseline IOP and limited IOP lowering on non-surgical treatment, especially in young patients with

already moderate or severe glaucoma at the time of diagnosis. TE is the first choice in primary glaucoma in phakic eyes with a clear lens; during the last decade we moved from TE to BGI in pseudophakic eyes. BGI is the first choice in secondary glaucoma.

Baerveldt glaucoma implant

A BGI with 350 mm^2 plate was implanted in the superior-temporal quadrant. The tube was closed with a Vicryl 7-0 suture to restrict short term excessive drainage before encapsulation. The tube was placed in the anterior chamber through a peripheral iridotomy made with the 23G needle used to create a scleral entry for the tube, aiming at a position of the tube as close to the iris as possible to prevent endothelial cell loss. (Tan et al. 2017) For this, the anterior chamber was temporarily deepened with an air bubble and the needle entered the eye 3 mm posterior to the limbus, in parallel with the iris plane. The tube entrance was covered with a patch of donor sclera. IOP-lowering drugs were continued unchanged until tube opening (typically six weeks after the intervention) and then tapered depending on the IOP. As of 2009, some patients got a drainage suture for early IOP reduction.(Rietveld et al. 2009, van Hoefen Wijsard et al. 2018) Antibiotic drops were given for two weeks (chloramphenicol 0.4% three times a day); steroids for 10 weeks (dexamethasone 0.1% three times a day for eight weeks followed by two times a day for two weeks).

Trabeculectomy

A TE with fornix-based conjunctival flap and limbus-based scleral flap was performed superiorly with the application of mitomycin C (0.2 mg/ml during 2-4 minutes; applied before making the scleral flap). The scleral flap was sutured with two or three Nylon 10-0 sutures, with the possibility for laser suture lysis afterwards. Conjunctiva was closed with two or three Vicryl 7-0 sutures. Antibiotic drops were given for two weeks (chloramphenicol 0.4% three times a day); steroids for 12 weeks (dexamethasone 0.1% six times a day for six weeks followed by four times a day for four weeks and finally two times a day for two weeks; oral prednisone 30 mg for 10 days was added if the inflammation was above average in the early postoperative phase).

Statistical analysis

As most of the characteristics of the study population were not normally distributed (according to the Shapiro-Wilk test), we used nonparametric descriptive statistics (median with interquartile range [IQR]) to describe the study population. Groups were compared using a Mann-Whitney test; for paired data we used a Wilcoxon test. Proportions were compared with a chi-square test. For pre-intervention and post-intervention visual acuity, IOP, and the number of IOP-lowering medications, we took

the last value before the intervention and a value measured as close as possible to 1 year after the operation; for MD we took the average value of the last two tests before the intervention and the first two tests after the intervention. Inevitably, a perimetric gap exists because of a waiting list before the intervention and delayed testing after the intervention (until full recovery or a situation deemed stable was reached). Hence, the observed MD decline is a composite of a surgery induced decline and a decline due to intrinsic disease progression during this perimetric gap (see Discussion section). ROP was determined using linear regression. Changes (post-intervention minus pre-intervention) were evaluated with a paired t-test or Wilcoxon test, depending on the distribution of the paired changes. Similarly, we used Pearson or Spearman correlation coefficients depending on the distribution of the concerning variables. A *P* value of 0.05 or less was considered statistically significant.

Meta-analysis

For the meta-analysis, we included glaucoma surgery studies regarding either BGI or TE that reported on preoperative and postoperative ROP of MD based on a mean follow-up of at least 3 years with SAP. Studies had to report a paired comparison of the change in ROP with standard error (SE), confidence interval (CI), standard deviation (SD) or *P* value. We searched for literature in PubMed with search string "(glaucom*) AND ((shunt OR tube) OR (trabeculect*) OR surgery) AND ((Visual field OR HFA OR perimetr*)) AND (progression OR (Rate of progression))" and in Embase. Additionally, we searched through the reference list of the identified articles. The search was performed in June 2017.

We calculated the mean of the change in ROP using a random-effects model. (DerSimonian & Laird 1986) If the SE of the change (or CI or SD) was not provided, it was calculated from the reported *P* value; if no exact *P* value was given, we used the upper limit (for example, 0.05 in case of $P<0.05$), yielding a conservative estimate of the SE.(Fu et al. 2013) We used a random-effects model because we expected significant heterogeneity amongst the included studies (indication of surgery, type of surgery, cataract extraction during follow-up, disease stage, postoperative care, ethnicity, etc.). Heterogeneity was evaluated using the I^2 statistic.(Higgins et al. 2003) I^2 is the percentage of the total variation across the studies, that is due to heterogeneity. Values up to 25%, 25-49%, 50-74%, and 75% and above are considered no, low, moderate, and high heterogeneity, respectively.(Higgins & Thompson 2002, Higgins et al. 2003) We performed a sensitivity analysis to evaluate the contribution of each individual study to the heterogeneity by sequentially leaving out one study and re-analysing the

pooled estimate for the remaining studies.(Tobias 1999) Publication bias was assessed with Egger's regression asymmetry test(Egger et al. 1997) and Begg's adjusted rank correlation test.(Begg & Mazumdar 1994)

RESULTS

Table 1 presents the general characteristics of the study population. As can be seen in this Table, 61 patients received a BGI and 39 a TE. For only 20 (11 BGI and 9 TE) patients, we had a sufficiently long follow-up duration (at least 5 years both pre- and postoperatively) to compare ROP pre- versus postoperatively, and for that reason, we pooled BGI and TE for this analysis. BGI and TE patients were comparable in age at surgery and gender. Preoperatively, the BGI patients had a lower BCVA (P=0.043), a slightly higher IOP (P=0.027), and a similar MD (P=0.92), compared to the TE patients. Pre- compared to postoperatively, both the BGI and the TE patients had a significantly lower IOP with less medication (P<0.001).

Table 1. General characteristics of the study population

	BGI (n=61)	TE (n=39)	*P*-value†
Age (year; median [IQR])	70 (57 to 75)	68 (59 to 73)	0.60
Gender (% female)	53	51	0.91
Visual acuity pre-op (logMAR, median [IQR])	0.14 (0.02 to 0.28)	0.10 (0.00 to 0.17)	0.043
IOP pre-op (mmHg; median [IQR])	21 (17 to 25)	18 (15 to 21)	0.027
IOP post-op (mmHg; median [IQR])	13 (11 to 16)	13 (9 to 15)	0.40
Pre-op number of IOP-lowering medications (median [IQR])	3 (2 to 4)	3 (2 to 4)	0.68
Post-op number of IOP-lowering medications (median [IQR])	2 (1 to 3)	0 (0 to 0)	<0.001
Pre-op mean deviation (dB; median [IQR])	-12 (-16 to -6)	-11 (-16 to -7)	0.92
Pseudophakic (%)*	54.1	10.3	<0.001
Secondary glaucoma (%)	23.0	5.1	0.035

*BGI = Baerveldt glaucoma implant; TE = trabeculectomy; IQR = interquartile range; IOP = intraocular pressure; dB = decibels; * = before surgery; † = Mann-Whitney test for medians, chi-square test for proportions.*

Figure 1 shows the change in BCVA as a function of preoperative visual field loss. No significant correlations were found (P=0.92 and P=0.55 for BGI and TE, respectively). A non-significant mean (SD) increase in logMAR BCVA (i.e., a decrease in decimal BCVA) of 0.03(0.14) was found for BGI (P=0.08) and of 0.04(0.14) for TE (P=0.12).

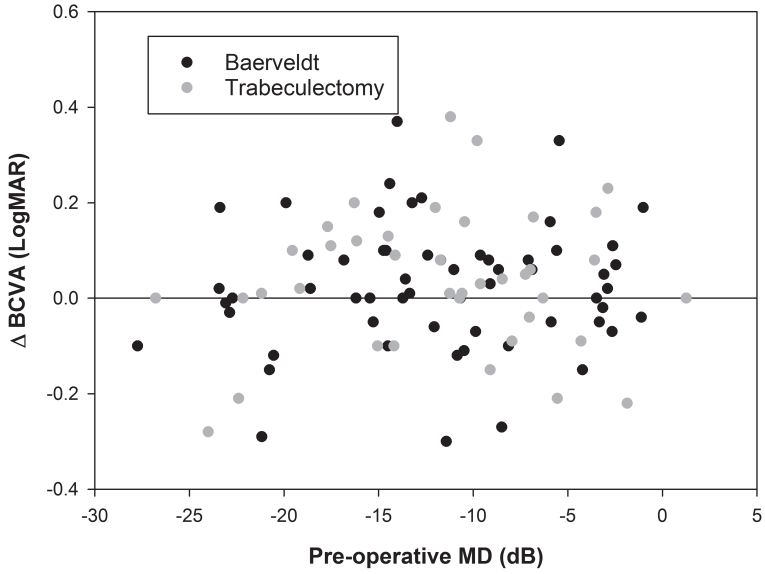


Figure 1. Surgery-induced change in best-corrected visual acuity (BCVA) as a function of preoperative visual field loss.

Figure 2 illustrates the surgery-induced visual field change as a function of preoperative visual field loss. Again, no significant correlations were found (P=0.19 and P=0.57 for BGI and TE, respectively). A significant mean(SD) MD decline of 1.3 (2.7) dB was found for BGI (P<0.001) and of 1.0(2.3) dB for TE (P=0.009). Of the 61 patients with a BGI, 18 had received a drainage suture for early IOP reduction. MD decline was 1.4 dB (P=0.004) without drainage suture and 1.1 dB (P=0.025) with drainage suture. The decline did not differ between the subgroups (P=0.73). The midpoint of the last two preoperative visual fields was 6 months before surgery; the midpoint of the first two postoperative visual fields was 1.5 years after surgery.

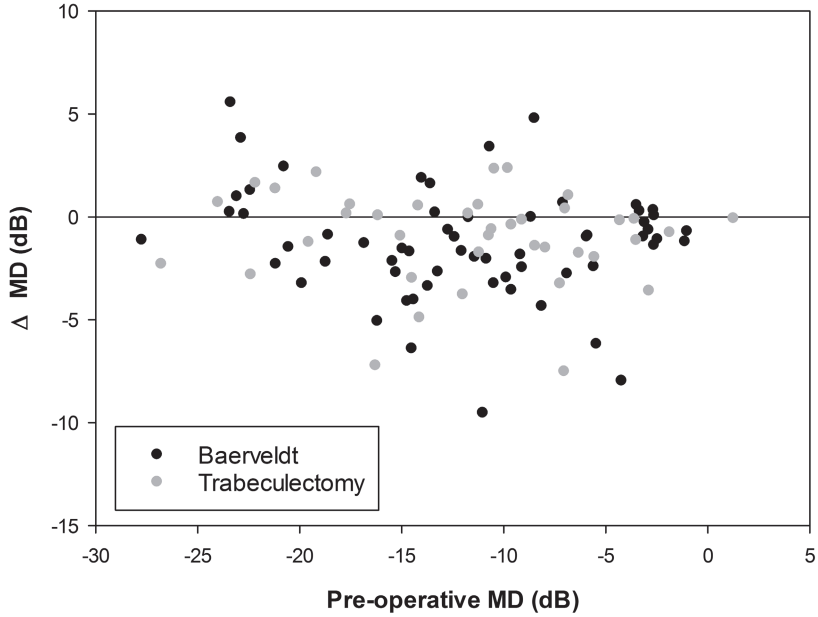


Figure 2. Surgery-induced change in visual field mean deviation (MD) as a function of preoperative visual field loss.

For the longitudinal analysis, the median (IQR) follow-up was 7.0(5.7 to 8.8) years preoperatively and 7.6(6.3 to 9.3) years postoperatively. The mean(SD) preoperative rate of progression was -0.37(0.52) dB/year. After surgery this was -0.15(0.48) dB/year; the paired difference (SE) was 0.22(0.18) dB/year (*P*=0.23). Of the 20 patients with a longitudinal follow-up, 4 (3 BGI, 1 TE) underwent a cataract extraction before the glaucoma surgery and 4 (2 BGI, 2 TE) after the surgery. Exclusion of these patients would result in a mean pre-op ROP of -0.37 dB/year and a mean post-op ROP of -0.24 dB/year (to be compared to -0.37 and -0.15 dB/year in Table 2). Hence, the bias seems limited.

Complications for patients receiving BGI (n=61) included ptosis (n=2), persistent (i.e., orthoptic consultation was warranted) diplopia (n=1), and additional surgery (cyclodiode laser therapy; n=1). TE (n=39) associated complications were persistent bleb leakage (that is, bleb leakage requiring surgical repair; n=2), failure (i.e., no visible bleb in combination with a higher than desired IOP; n=2), hypotony with choroidal detachment closer than two disc diameters from the optic disc (n=1), and hypotony associated maculopathy (n=1).(Abbas et al. 2018)

Table 2. Characteristics of the studies included in the meta-analysis

Study	Number of Patients	Mean follow-up (pre/post; yr)	pre-op ROP (mean [SD]; dB/yr)	post-op ROP (mean [SD]; dB/yr)	Change in ROP [mean (SE); dB/year]*	*P*-value
Folgar et al. (2010)	28	3.6/3.5	-1.48 (1.4)	-0.43 (0.8)	1.05 (0.41)	0.01
Bhardwaj et al. (2013; pre-op nonprogressors)	8	5.8/4.5*	0.1 (0.8)	-0.2 (0.4)	-0.30 (0.35)	0.40
Bhardwaj et al. (2013; pre-op progressors)	9	5.8/4.5*	-1.0 (0.9)	-0.2 (0.38)	0.80 (0.34)	0.02
Bertrand et al. (2014)	52	3.8/3.9	-0.36 (0.79)	-0.16 (0.58)	0.20 (0.14)	0.15
Mataki et al. (2014)	34	4.6/5.7	-0.70 (0.52)	-0.25 (0.50)	0.45 (0.11)	<0.001
Caprioli et al. (2016)	74	5.1/5.4	-0.70 (1.1)	-0.10 (0.8)	0.60 (0.18)	<0.001
Iverson et al. (2016)	9	7.6/5.4	-1.05 (0.66)	-0.25 (0.86)	0.80 (0.41)	0.05
Oie, Ishida & Yamamoto (2017)	17	5.9/15.6	-0.86 (0.51)	-0.19 (0.2)	0.67 (0.20)	<0.001
Junoy Montolio et al.	20	7.0/7.6	-0.37 (0.52)	-0.15 (0.48)	0.22 (0.18)	0.23

ROP = rate of progression; SD = standard deviations; SE = standard error.

** Estimated SE in case of non-exact P-value.*

Table 2 shows the characteristics of the studies included in the meta-analysis. Eight studies were included, of which one (Bhardwaj et al.) reported two subgroups.(Bhardwaj et al. 2013) All but two (Folgar et al.; this study)(Folgar et al. 2010) included only TE. Figure 3 shows the preoperative and postoperative ROP with SE. As can be seen in this figure, there was a wide variety in ROP preoperatively but not postoperatively (see Discussion section). The weighted mean (95% CI) surgery-induced change in ROP was 0.44(0.25 to 0.64; P<0.001) dB/year, indicating that glaucoma surgery results in a significant deceleration of ROP. I^2 was 47%, that is, the studies showed a low heterogeneity. Sensitivity analysis showed that one of the subgroups of the study of Bhardwaj et al. substantially influenced the pooled estimate.(Bhardwaj et al. 2013) After excluding this study, the overall effect size was 0.45(95% CI 0.32 to 0.57; *P*<0.001) dB/year with a remaining I^2 34%. There was no evidence of publication bias (Egger's test: *P*=0.48; Begg's test: *P*=0.30).

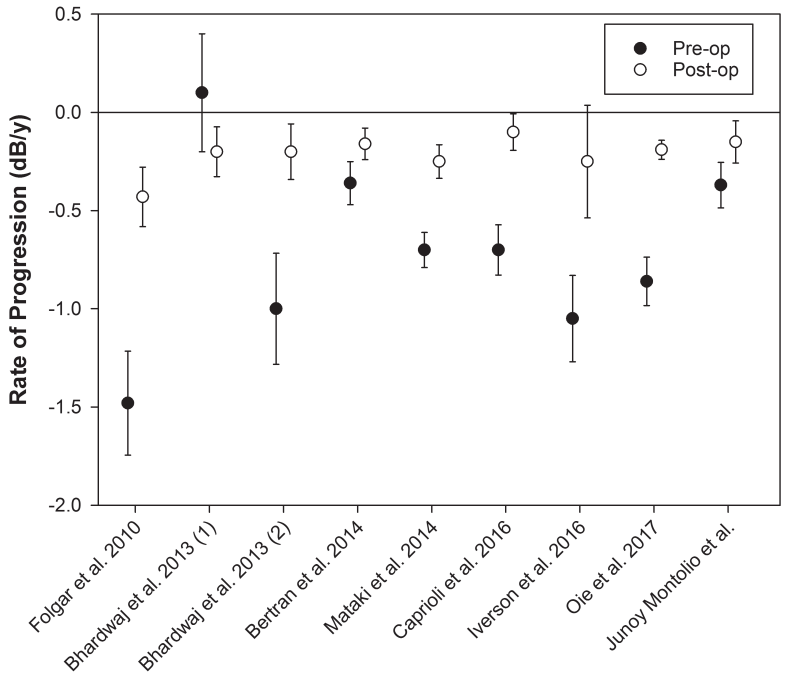


Figure 3. Forest plot of effect sizes of the surgery-induced change in rate of visual field progression.

DISCUSSION

In our study, both TE and BGI were associated with a small but significant MD loss; no surgery-induced changes in BCVA were found. Surgery was associated with a non-significant deceleration of the ROP. In the meta-analysis, based on eight studies, the overall effect size indicated that glaucoma surgery significantly reduces the progression velocity in glaucoma.

Several studies reported an unchanged BCVA after glaucoma surgery, for both TE(Balekudaru et al. 2014, Bertrand et al. 2014, Bevin et al. 2008, Bhardwaj et al. 2013, Iverson et al. 2016) and BGI,(Namavari et al. 2016) and in a recent randomized clinical trial (RCT) comparing primary BGI with TE.(Islamaj et al. 2018) This is in agreement with our results. Only a few studies described BCVA loss after surgery.(Christakis et al. 2017, Gedde et al. 2007, Stead & King 2011) Stead & King(Stead & King 2011) found, in patients with very severe glaucoma (mean[SD] MD -25.3[3.5] dB) who underwent a TE, that nine of 104 patients lost two lines of Snellen or more - deemed to be related to glaucoma - at the end of the follow-up (i.e., on average three years postoperatively). In the multicentre RCT of Gedde et al.(Gedde et al. 2007) comparing BGI with TE, the BCVA decreased significantly, on average (SD) from 0.43(0.54) and 0.37(0.38) logMAR preoperatively to 0.61(0.75) and 0.49(0.56) after one year for the BGI and TE patients, respectively. The authors described that the main reasons for the loss of BCVA were cataract and other, not glaucoma-related. Christakis et al.(Christakis et al. 2017) found in their multicenter RCT comparing Ahmed drainage devices with BGI a mean(SD) decrease in BCVA from 1.1(1.0) to 1.3(1.2) logMAR in BGI patients after one year. They described that 34% of the patients lost two or more lines of Snellen, of which 18% was caused by glaucoma, 15% by cataract, and 68% by macular disease or other causes. Harju et al.(Harju et al. 2018) studied the long-term results of deep sclerectomy. Four of 37 patients had a VA loss of two lines or more; in two cases this was attributed to cataract formation. The main difference between these four studies and our study is the much higher preoperative BCVA in our study (Table 1). This suggests that a low preoperative visual acuity is an indicator of further surgery-induced loss, a plausible hypothesis from a clinical point of view. The limited variability in our preoperative BCVA preludes a detailed check of this hypothesis in our data. Development of cataract could be an issue, as progression of cataract is found within 1 year after surgery, particularly in eyes with post-operative complications.(Gedde et al. 2007) However, this seemed not to be the case in our data. Subgroup analysis comparing patients who were pseudophakic before surgery with patients who were phakic did not reveal any difference in visual acuity change, neither in the BGI group nor in the TE group.

Only a few earlier studies reported on a surgery-induced MD decline. Most studies found no influence on the MD;(Balekudaru et al. 2014, Islamaj et al. 2018, Sehi et al. 2010, Tavares et al. 2006, Wright et al. 2015, Yamazaki & Hayamizu 2012) one study found a deterioration of the MD after a TE.(Hagiwara et al. 2000) In the latter study, the postoperative MD was assessed at the end of the follow-up (mean follow-up duration was 4.75 years). In most of the studies that found no effect, the authors compared the first visual field after the operation to the last visual field before the operation; in two studies they performed two visual fields at the same time point. We reported a mean surgery-induced MD decline of 1.3 and 1.0 dB for BGI and TE, respectively. This decline was based on the last two tests before the intervention and the first two tests after the intervention. This method was chosen to reduce variability and to compensate for a possible regression to the mean effect: a poor VF could trigger the clinician to initiate surgery and considering only this field could thus mask surgery-induced damage. By using four visual fields, however, a part of the observed decline will actually be caused by progression that would have occurred anyway – the fields were not clustered (as would have been done in a trial) but made as part routine clinical care. This is illustrated in Figure 4. If we assume an observed surgery-induced MD decline of 1.2 dB, an interval of two years between the time point halfway the last two tests before the intervention and the time point halfway the first two tests after the intervention (0.5 years before the intervention and 1.5 years after the intervention; based on our 61+39 interventions), and a ROP of -0.63 dB/year before the intervention and -0.19 dB/year after the intervention (inverse-variance estimates based on the meta-analysis), then 50% (0.5*0.63+1.5+0.19 = 0.6 of 1.2 dB) of the observed decline could be attributed to the perioperative perimetric gap and the other 50% to the surgery itself.

In our study, glaucoma surgery yielded a non-significant deceleration of the ROP. The non-significance could be related to the small sample size and/or remaining variability in the ROP estimate in individual patients - despite a relatively long follow-up duration, or may actually denote no effect. The small sample size was related to the exclusion of combined surgery (glaucoma surgery and cataract extraction simultaneously; see next paragraph) and the required long follow-up duration; in our tertiary referral center, most glaucoma surgeries are performed on patients with secondary glaucoma(de Vries et al. 2016) and on referred patients. Both groups often lack an uninterrupted perimetric follow-up. We confined the current study to the regular visitors of our department, for whom we collect observational data prospectively as of 2000 - as part of the GLGS. Interestingly, all identified published studies that reported on ROP before and after glaucoma surgery pointed in the same - beneficial - direction, but the change in ROP

was not always statistically significant. The overall effect size of the meta-analysis, however, was highly significant, indicating the value of meta-analysis as a statistical tool to provide a solid ground for performing surgery.

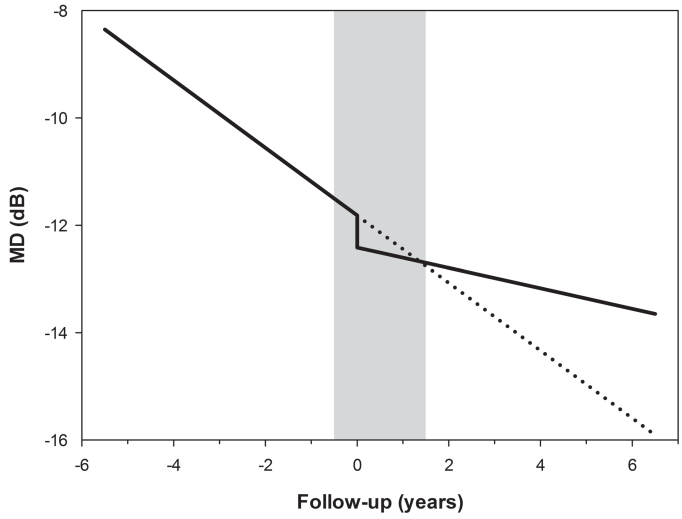


Figure 4. Schematic time course of the visual field mean deviation (MD) with (continuous line) and without (dashed line) surgery. The surgery itself is performed at time point 0 years. The surgery-induced drop in MD is counterbalanced by the slower rate of progression, on average 1.5 years after surgery (crossing of continuous and dashed line). The grey bar depicts the period in which no perimetry is performed (from 6 months before to 1.5 years after surgery; see Results section). The observed drop in MD, that is, the difference in MD between the beginning and the end of the grey bar, is about twice as large as the actual drop in MD, because of autonomous disease progression in the period without perimetry.

We excluded patients who underwent a cataract extraction simultaneously with the glaucoma surgery but we did not exclude those who underwent a cataract extraction during the longitudinal follow-up. Both inclusion and exclusion of patients who underwent a cataract extraction during the longitudinal follow-up could induce bias. As shown in the Results section, this bias seems limited in this population.

Before the operation, the mean ROP varied largely between the studies. This might reflect, amongst others, differences in policy regarding indicating surgery. After the operation, the mean ROP was close to -0.2 dB/year in most of the studies (Table 2 and Figure 3). Interestingly, a value of -0.2 dB/year is the typical mean value as reported in several glaucoma cohorts.(De Moraes et al. 2012, Heijl 2002, Wesselink et al. 2012) This

indicates that glaucoma surgery is indeed able to bring rapid progressors back into the normal ROP range. Despite the preoperative variability in ROP, the studies in the meta-analysis showed a low heterogeneity. The majority of the heterogeneity could be attributed to the subgroup of the study of Bhardwaj et al. in which patients were included without preoperative progression.(Bhardwaj et al. 2013)

Limitations of this study were the observational design of the GLGS and the limited sample size in especially the longitudinal part of the current study. Reasons for the limited sample size were given above; we addressed this by performing a meta-analysis. By combining many studies that showed not always significant results on their own, we were able to determine the highly significant benefit of glaucoma surgery. The observational nature of the GLGS implies that the patients included in the current study may form a biased sample. In case of a real disaster, post-op visual fields may be lacking, resulting in exclusion and bias towards good outcomes. On the other hand, visual fields may also be taken less frequently in cases with a well-regulated IOP (as they are less urgent for clinical decision making), yielding bias towards poor outcomes. Clearly, an observational study, and even not a meta-analysis of observational studies, can beat a large RCT. However, RCTs performed in the field of glaucoma surgery thus far reported only on IOP and visual field change defined as a post-op progression event;(Anderson et al. 2003, Ederer et al. 1994, Gedde et al. 2012, Musch et al. 1999) no RCT data regarding change in ROP are available. We required a follow-up of at least 5 years pre- and postoperatively for our own longitudinal analysis (see Methods section); we had to weaken this to 3 years for the meta-analysis (Table 2). With 3 years, ROP assessment in individual patients is still very noisy, as was argued by modelling(Jansonius 2010) and shown in patient data (Table 3 of the study of Junoy Montolio et al. shows that a short follow-up causes a spurious widening of the ROP distribution).(Junoy Montolio et al. 2012)

Glaucoma surgery could be an ungrateful intervention to perform as most patients have little complaints preoperatively, while experiencing the perioperative and postoperative hassle and possible visual acuity and visual field loss. Therefore, it is very important to understand what patients relinquish to prevent future blindness. Figure 4 illustrates the cost and benefit of glaucoma surgery, compiled from our cost data (surgery-induced MD decline) and the ROP results from the studies included in the meta-analysis. It shows that, on average 1.5 years after the operation (depicted as time point 0 years in Figure 4), patients that were operated are better off than patients who declined surgery (crossing of continuous and dashed line in Figure 4). Note that the end of the perimetric gap (right side of grey bar) and the crossing of the lines coincide coincidentally; this may give the clinician, when restarting perimetry after surgery,

the spurious impression that the surgery did not have any effect. Figure 4 may also be used to inform the patient properly, and it illustrates that the decision to operate is not always an easy one, but should involve the weighing of, amongst others, current damage, ROP, IOP, and life expectancy.(Wesselink et al. 2011)

In conclusion, the cost of glaucoma surgery is the loss of visual function associated with the procedure; the benefit is the long-term preservation of the visual field due to the reduced rate of progression. Both are significant. On average, the benefit of glaucoma surgery surpasses the cost after approximately 1.5 years.

REFERENCES

Abbas A, Agrawal P & King AJ (2018): Exploring literature-based definitions of hypotony following glaucoma filtration surgery and the impact on clinical outcomes. Acta Ophthalmol. 96: e285–e289.

Anderson DR, Drance SM & Schulzer M (2003): Factors that predict the benefit of lowering intraocular pressure in normal tension glaucoma. Am J Ophthalmol. 136: 820–829.

Arora KS, Robin AL, Corcoran KJ, Corcoran SL & Ramulu PY (2015): Use of various glaucoma surgeries and procedures in medicare beneficiaries from 1994 to 2012. Ophthalmol. 122: 1615–1624.

Balekudaru S, George R, Panday M, Singh M, Neog A & Lingam V (2014): Prospective evaluation of early visual loss following glaucoma-filtering surgery in eyes with split fixation. J Glaucoma. 23: 211–218.

Begg CB & Mazumdar M (1994): Operating characteristics of a rank correlation test for publication bias. Biometrics. 50: 1088–1101.

Bengtsson B & Heijl A (2000): False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? Am J Ophthalmol. 130: 689.

Bertrand V, Fieuws S, Stalmans I & Zeyen T (2014): Rates of visual field loss before and after trabeculectomy. Acta Ophthalmol. 92: 116–120.

Bevin TH, Molteno ACB & Herbison P (2008): Otago Glaucoma Surgery Outcome Study: long-term results of 841 trabeculectomies. Clin Experiment Ophthalmol. 36: 731–737.

Bhardwaj N, Niles PI, Greenfield DS, Hymowitz M, Sehi M, Feuer WJ & Budenz DL (2013): The impact of surgical intraocular pressure reduction on visual function using various criteria to define visual field progression. J Glaucoma. 22: 632–637.

Caprioli J, de Leon JM, Azarbod P, Chen A, Morales E, Nouri-Mahdave K, Coleman A, Yu F & Afifi A. (2016): Trabeculectomy can improve long-term visual function in glaucoma. Ophthalmology. 123: 117–128.

Christakis PG, Zhang D, Budenz DL, Barton K, Tsai JC, Ahmed IIK & ABC-AVB Study Groups (2017): Five-year pooled data analysis of the Ahmed Baerveldt Comparison Study and the Ahmed Versus Baerveldt Study. Am J Ophthalmol. 176: 118–126.

De Moraes CG, Demirel S, Gardiner SK, Liebmann JM, Cioffie GA, Ritch R, Gordon MO, Kass MA & Ocular Hypertension Treatment Study Group. (2012): Effect of treatment on the rate of visual field change in the ocular hypertension treatment study observation group. Invest Ophthalmol Vis Sci. 53: 1704–1709.

DerSimonian R & Laird N (1986): Meta-analysis in clinical trials. Control Clin Trials. 7: 177–188.

de Vries MM, Müskens RPHM, Renardel de Lavalette VW, Hooymans JMM & Jansonius NM (2016): Glaucoma drainage device surgery after vitreoretinal surgery: incidence and risk factors. Acta Ophthalmol. 94: 135–139.

Ederer F, Gaasterland DE, Sullivan EK & AGIS Investigators (1994): The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients. Control Clin Trials. 15: 299–325.

Egger M, Davey Smith G, Schneider M & Minder C (1997): Bias in meta-analysis detected by a simple, graphical test. BMJ. 315: 629–634.

Folgar FA, de Moraes CGV, Prata TS, Teng CC, Tello C, Ritch R & Liebmann JM (2010): Glaucoma surgery decreases the rates of localized and global visual field progression. Am J Ophthalmol. 149: 258–264.e2.

Fu R, Vandermeer BW, Shamliyan TA, O'Neil ME, Yazdi F, Fox SH & Morton SC (2013): Handling continuous outcomes in quantitative synthesis. Methods Guide for Effectiveness and Comparative Effectiveness Reviews.

Gedde SJ, Schiffman JC, Feuer WJ, Herndon LW, Brandt JD & Budenz DL (2007): Treatment outcomes in the tube versus trabeculectomy study after one year of follow-up. Am J Ophthalmol. 143: 9–22.

Gedde SJ, Schiffman JC, Feuer WJ, Herndon LW, Brandt JD, Budenz DL & Tube versus Trabeculectomy Study Group (2012): Treatment outcomes in the Tube Versus Trabeculectomy (TVT) study after five years of follow-up. Am J Ophthalmol. 153: 789–803.e2.

Hagiwara Y, Yamamoto T & Kitazawa Y (2000): The effect of mitomycin C trabeculectomy on the progression of visual field defect in normal-tension glaucoma. Graefes Arch Clin Exp Ophthalmol. 238: 232–236.

Harju M, Suominen S, Allinen P & Vesti E (2018): Long-term results of deep sclerectomy in normal-tension glaucoma. Acta Ophthalmol. 96: 154–160.

Heeg GP, Blanksma LJ, Hardus PLLJ & Jansonius NM (2005): The Groningen Longitudinal Glaucoma Study. I. Baseline sensitivity and specificity of the frequency doubling perimeter and the GDx nerve fibre analyser. Acta Ophthalmol Scand. 83: 46–52.

Heijl A (2002): Reduction of intraocular pressure and glaucoma progression: Results from the Early Manifest Glaucoma Trial. Arch Ophthalmol. 120: 1268-1279.

Higgins JPT & Thompson SG (2002): Quantifying heterogeneity in a meta-analysis. Stat Med 21: 1539–1558.

Higgins JPT, Thompson SG, Deeks JJ & Altman DG (2003): Measuring inconsistency in meta-analyses. BMJ 327: 557–560.

Islamaj E, Wubbels RJ & de Waard PWT (2018): Primary Baerveldt versus Trabeculectomy Study after one-year follow-up. Acta Ophthalmol. 96: e740–e746.

Iverson SM, Schultz SK, Shi W, Feuer WJ & Greenfield DS (2016): Effectiveness of single-digit IOP targets on decreasing global and localized visual field progression after filtration surgery in eyes with progressive normal-tension glaucoma. J Glaucoma. 25: 408–414.

Jansonius NM (2010): On the accuracy of measuring rates of visual field change in glaucoma. Br J Ophthalmol. 94: 1404–1405.

Junoy Montolio FG, Wesselink C & Jansonius NM (2012): Persistence, spatial distribution and implications for progression detection of blind parts of the visual field in glaucoma: a clinical cohort study. PLoS One. 7: e41211.

Musch DC, Lichter PR, Guire KE & Standardi CL (1999): The Collaborative Initial Glaucoma Treatment Study: study design, methods, and baseline characteristics of enrolled patients. Ophthalmology. 106: 653–662.

Namavari A, Hyde RA, Wang D, Vajaranant TS & Aref AA (2016): Primary Baerveldt shunt implantation: outcomes and complications. Ophthalmol Ther. 5: 253–262.

Rietveld E, Jansonius NM, & Müskens RPHM (2009): Immediate pressure reduction with baerveldt glaucoma implants. Invest Ophthalmol Vis Sci. 50: 445–445.

Sehi M, Grewal DS, Goodkin ML & Greenfield DS (2010): Reversal of retinal ganglion cell dysfunction after surgical reduction of intraocular pressure. Ophthalmology. 117: 2329–2336.

Stead RE & King AJ (2011): Outcome of trabeculectomy with mitomycin C in patients with advanced glaucoma. Br J Ophthalmol. 95: 960–965.

Tan AN, Webers CAB, Berendschot TTJM, de Brabander J, de Witte PM, Nuijts RMMA, Schouten JSAG & Beckers HJM (2017): Corneal endothelial cell loss after Baerveldt glaucoma drainage device implantation in the anterior chamber. Acta Ophthalmol. 95: 91–96.

Tavares IM, Melo LAS, Prata JA, Galhardo R, Paranhos A & Mello PAA (2006): No changes in anatomical and functional glaucoma evaluation after trabeculectomy. Graefes Arch Clin Exp Ophthalmol. 244: 545–550.

Tobias A (1999): Assessing the influence of a single study in the meta-analysis estimate. Stata Tech Bull. 47: 15–17.

van Hoefen Wijsard M, Haan M, Rietveld E & van Rijn LJ (2018): Donor sclera versus bovine pericardium as patch graft material in glaucoma implant surgery and the impact of a drainage suture. Acta Ophthalmol. 96: 692–698.

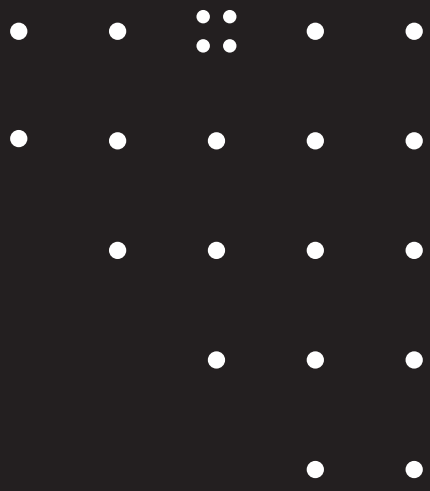
Wesselink C, Marcus MW & Jansonius NM (2012): Risk factors for visual field progression in the Groningen Longitudinal Glaucoma Study: a comparison of different statistical approaches. J Glaucoma. 21: 579–585.

Wesselink C, Stoutenbeek R & Jansonius NM (2011): Incorporating life expectancy in glaucoma care. Eye. 25: 1575–1580.

Wright TM, Goharian I, Gardiner SK, Sehi M & Greenfield DS (2015): Short-term enhancement of visual field sensitivity in glaucomatous eyes following surgical intraocular pressure reduction. Am J Ophthalmol. 159: 378–85.e1.

Yamazaki Y & Hayamizu F (2012): Superior segmental optic nerve hypoplasia accompanied by progressive normal-tension glaucoma. Clin Ophthalmol. 6: 1713–1716.

VI



Variability in perimetry:

Continuous light increment perimetry versus the staircase procedure

Francisco G. Junoy Montolio & Nomdo M. Jansonius

ABSTRACT

Purpose: To determine the influence of the psychophysical method on the variability of static perimetry: continuous light increment perimetry (CLIP) versus the staircase procedure (SP).

Methods: Twenty-eight eyes of twenty-eight glaucoma patients of varying disease stages were included in a cross-sectional comparative study. All patients performed two CLIP tests and two SP tests using the Twinfield perimeter. We compared the mean sensitivity (MS) within and between the two psychophysical methods.

Results: Mean sensitivity coefficient of repeatability was 3.2 dB for CLIP and 2.4 dB for SP (P=0.08). The limits of agreement for SP compared to CLIP were between -4.0 and +2.6 dB with a lower sensitivity for SP (mean difference -0.7 dB; P=0.03). Test-time was significantly shorter in CLIP compared to SP (mean[SD]: 04:42[01:10] versus 09:14[00:46] min:sec; P<0.001).

Conclusion: Test-retest variability in perimetry is not explained by the psychophysical method used, which strengthens the idea that glaucoma itself is the major factor contributing to variability.

INTRODUCTION

Visual field (VF) testing, or perimetry, is essential to assess the effectiveness of glaucoma management. Glaucoma patients are familiar with a half-yearly to yearly cadence of testing. This frequency is increased to validate test results, when test results are unreliable or disease progression is suspected. Particularly disease progression has to be confirmed or falsified to distinguish genuine progression from test-retest variability. Until recently, glaucoma itself was the best known factor causing this variability.(Heijl et al. 1989) However, additional factors (e.g., time of day, season, technician experience) have been found to influence perimetry by adding noise to the main outcome measure (i.e., the Mean Deviation [MD]).(Junoy Montolio et al. 2012, Gardiner et al. 2013) Despite this knowledge, these factors explain only a fraction of the variability found. This leaves the question unanswered if the perimetric variability as seen in glaucoma patients indeed is an intrinsic disease property, or that we are still overlooking other, yet unknown factors.

A factor that may cause noise in perimetry is the applied psychophysical method. Perimetry essentially is a systematic determination of the sensitivity to a visual stimulus at different locations in space, and this sensitivity can be assessed in various ways (psychophysical methods). The classical method is the staircase procedure (SP). In this method, a stimulus is presented for a short period of time (typically 200 ms), and, depending on the response of the subject (seen or not seen), the next stimulus has a higher or lower intensity. This is repeated until the threshold can be determined (typically after one to three crossings). About a decade ago, a new strategy was developed to decrease test time and to increase patient compliance: the Continuous Light Increment Perimetry (CLIP) strategy (Twinfield perimeter; Oculus Inc., Wetzlar, Germany). After calculation of the reaction time in four threshold trials, a continuous increase in light intensity is given for each test location. Starting with a 5 dB dimmer intensity than expected to be the threshold based on age, central threshold, and neighboring values if already collected, the intensity intensifies (continuously) with 1 dB per reaction time (typically 100 ms), over 8 dB in total. Thereafter, the intensity increases with 2 dB per reaction time over 6 dB, followed by 4 dB per reaction time until the stimulus is seen or the maximum intensity is reached (318 cd/m^2). According to several studies, CLIP would perform better than the classical SP and its variants like the Swedish Interactive Thresholding Algorithm (SITA).(Wabbels et al. 2001) CLIP showed promising results in glaucoma patients as well.(Wabbels et al. 2005, Capris et al. 2008)

The comparison between psychophysical methods performed by Wabbels et al. and Capris et al. were partially performed in different brands of perimeters (Humphrey Field Analyzer [HFA] vs Twinfield perimeter), which makes inter-test comparisons difficult.(Wabbels et al. 2005, Capris et al. 2008) In addition, the objectives in these studies were to compare the inter-test visual field indices (i.e. mean sensitivity) rather than comparison of intra-test variability.

The aim of this study was to determine the influence of the psychophysical method on the variability of static perimetry in glaucoma by comparing CLIP with SP in glaucoma patients. As a secondary aim we also compared the agreement between both methods and test time. For this purpose, patients were tested twice per strategy, while a suprathreshold strategy was performed before and after each session to introduce patients to the new perimeter and to monitor any learning effects and the influence of fatigue. Stimulus size V was chosen because it has been reported to have less variability than the commonly used size III stimulus,(Wall et al. 1997, Wall et al. 2013) thus making it more likely to uncover differences in variability caused by the psychophysical method.

METHODS

Study population

The present study was performed in the Department of Ophthalmology of the University Medical Center Groningen in the Netherlands. Glaucoma patients who were scheduled for a regular appointment at our outpatient department in August 2010 received an invitation for this institutional review board–approved cross-sectional comparative study (METc2010.183). Only patients with known glaucoma were included, after providing informed consent. Glaucoma was defined as a characteristic optic nerve head appearance and a VF defect compatible with glaucoma and without any other explanation. Patients had to have experience with standard automated perimetry (SAP). The selected patients performed, prior to their regular appointment, six consecutive VF tests with the Twinfield perimeter. The study followed the tenets of the Declaration of Helsinki.

Perimetry

Perimetry was performed using the Oculus Twinfield 24-2 CLIP and SP strategy. The CLIP strategy is described above (see INTRODUCION), and in more detail by others,(Wabbels et al. 2001) while the SP is comparable with the HFA version (see INTRODUCTION).

All tests included a 54-point grid covering the central VF with a radius of 24 degrees and were performed using a stimulus size V. Each patient subsequently performed six consecutive VFs with a five-minute pause in between: two suprathreshold (ST), two CLIP and two SP tests, divided into two sections. The first section always started with a suprathreshold test (ST test) while the second section always ended with a ST test. In between, two CLIP and two SP tests which were performed in random order, giving four possible sequences. The first ST test was added so patients could get used to the Twinfield perimeter and greater stimulus size; the second ST test to investigate possible fatigue effects.

Data analysis

Patients were stratified in three groups, determined by the MD of the latest clinical VF test (HFA, SITA fast), being up to -6 dB (early glaucoma), from -6 to -12 dB (moderate), and beyond -12 dB (severe). When both eyes were eligible for the study, one eye was randomly selected. To describe the study population, parametric or nonparametric descriptive statistics was performed, depending on the concerning distributions.

Pointwise comparisons of raw thresholds were presented in scatterplots, showing the results for the first versus the second test, for both strategies. We plotted <0 dB thresholds as -1 dB to make a distinction between perimetrically blind thresholds (<0 dB) and a near blind threshold (0 dB; i.e., only the maximum light intensity is seen by the patient). Subsequently, the intra- and inter-strategy variability were analysed using the mean sensitivity (MS; i.e., the average of the pointwise sensitivities across the VF). First, the comparisons (first versus second CLIP test, first versus second SP test, and SP versus CLIP test) were illustrated in Bland-Altman plots.(Bland & Altman 1986) Secondly, analyses were quantified using the coefficient of repeatability (CoR) and limits of agreement (LoA), being 1.96 times the standard deviation of differences and mean difference +/- 1.96 times the standard deviation of differences, respectively. Thirdly, we compared the repeatability of the MS between the strategies by applying a paired t-test to the absolute differences between the first and second MS of each strategy. Calculations were performed in SPSS statistics 18.0.3 (SPSS Inc. Chicago, IL, USA) and WinPepi (11.65; http://www.brixtonhealth.com/pepi4windows.html).

RESULTS

Twenty-nine out of 36 invited patients (81%) agreed to participate in this study; data from one patient had to be excluded because of wrong perimeter settings during the test. Table 1 presents the baseline characteristics of the remaining 28 participants. Early, moderate, and severe glaucoma consisted of 10, 7, and 11 patients, respectively.

Table 1. Baseline characteristics for early, moderate and severe glaucoma (mean with standard deviation between brackets unless stated otherwise)

Baseline	Early	Moderate	Severe	Total
Number of patients	10	7	11	28
Age (years)	70.2 (7.7)	67.6 (8.3)	66.8 (13.3)	68.3 (10.0)
Gender (% male)	40.0	71.4	45.5	48.3
Right eye (%)	60.0	71.4	36.4	51.7
Mean Deviation (latest clinical VF test; dB)	-2.8 (1.9)	-9.7 (1.6)	-17.3 (3.6)	-10.3 (6.8)

Table 2 presents the number of black dots for the ST test and the mean sensitivity (MS), reliability indices (fixation losses [FL] and false positives [FP]), and test time for the CLIP and SP strategy, stratified according to glaucoma stage. After Bonferroni correction, there were no significant differences between test one and test two for ST, CLIP, or SP. CLIP in particular showed high percentages of FL, while FP answers were reasonably consistent between strategies. Test time differed significantly between the test strategies (P<0.001; Friedman Test), with CLIP being the shorter test. The ST test showed no evidence of fatigue (P=0.43).

Figure 1 shows scatterplots of the raw threshold values within the strategies. Both strategies showed a reasonably similar trend of low variability for the higher sensitivities and high variability for the lower sensitivities. Figure 2 presents Bland-Altman plots of the MS for the two CLIP tests (A), the two SP tests (B), and for the second CLIP versus the second SP test (C). Coefficients of repeatability were 3.2 and 2.4 dB for CLIP and SP, respectively. The apparent difference in repeatability of the MS between the strategies was not significant (P=0.08). The MS of the second CLIP test was significantly higher compared to the first test (mean [SD] difference 0.7 [1.6] dB; P=0.03), indicating that learning was more important than fatigue. The two SP tests did

Table 2. Results (mean with standard deviation between brackets unless stated otherwise and *P*-value if relevant)

				Total (n=28)		
Subgroup	**Early (n=10)**	**Moderate (n=7)**	**Severe (n=11)**	**FL (%)**	**FP (%)**	**test time (mm:ss)**
Latest clinical test (MD, dB)	-2.7 (1.9)	-9.7 (1.6)	-17.3 (3.6)	-	-	-
ST1 (No. ■)	3.8 (3.7)	6.4 (5.2)	21.6 (9.0)	6.6 (12.5)	2.4 (5.2)	02:39 (00:29)
ST2 (No. ■)	2.4 (3.5)	6.3 (5.8)	21.8 (8.6)	5.9 (11.5)	0.0 (0.0)	02:39 (00:32)
Paired differences ST1-ST2	1.4 (4.1) *P=0.30*	0.14 (2.7) *P*=0.89	-0.18 (2.2) *P*=0.79			
CLIP1 (MS, dB)	22.0 (1.6)	17.5 (2.4)	12.6 (3.6)	11.9 (22.6)	1.5 (4.5)	04:46 (01:07)
CLIP2 (MS, dB)	22.3 (1.8)	19.2 (1.5)	13.0 (2.7)	20.5 (23.9)	2.8 (6.1)	04:42 (01:10)
Paired differences CLIP1-CLIP2 (MS, dB)	-0.29 (1.1) *P*=0.42	-1.70 (1.6) *P*=0.03	-0.44 (1.9) *P*=0.46)			
SP1 (MS, dB)	21.4 (1.8)	17.7 (2.3)	12.5 (3.4)	6.4 (8.3)	3.1 (5.3)	09:13 (00:49)
SP2 (MS, dB)	21.9 (2.0)	18.2 (1.7)	12.2 (2.9)	8.5 (13.0)	2.4 (4.3)	09:14 (00:46)
Paired differences SP1-SP2 (MS, dB)	-0.49 (1.2) *P*=0.22	-0.49 (1.1) *P*=0.27	0.33 (1.3) *P*=0.44			

Abbreviations: *ST = Suprathreshold; CLIP = Continuous Light Increment Perimetry; SP = Staircase Procedure; FL = Fixation losses; FP = False positives.*

not differ significantly (mean [SD] difference 0.2 [1.2] dB; *P*=0.48). Limits of agreement were between -4.0 and +2.6 dB for the second SP test versus the second CLIP test; there was a bias of -0.7 dB with SP showing a significantly lower mean MS (*P*=0.03).

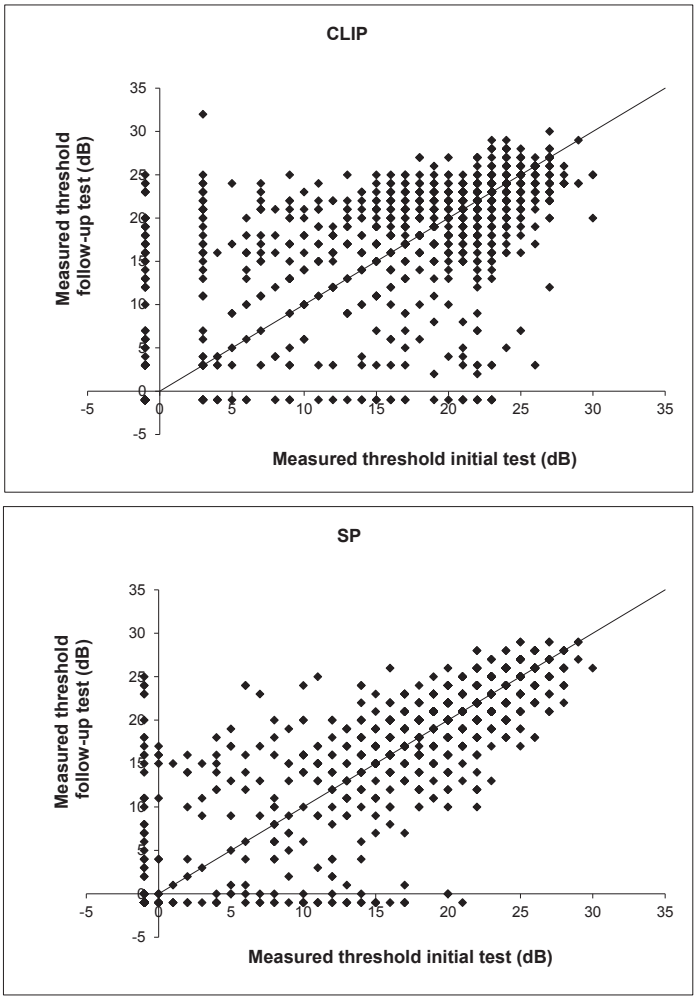


Figure 1. Scatter plots showing intra-strategy, pointwise raw thresholds. Lines show y=x.

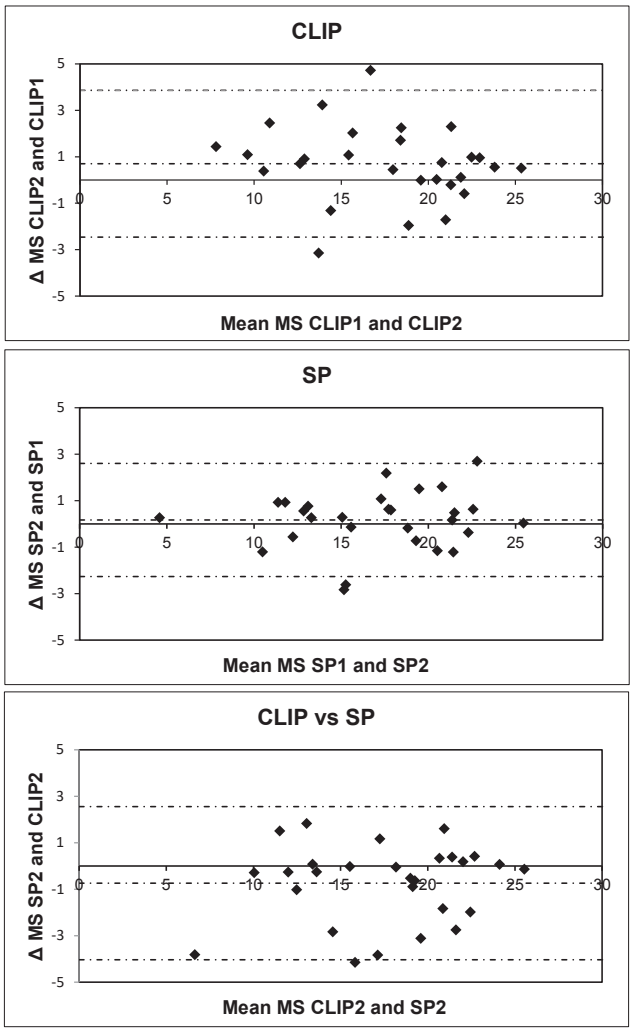


Figure 2. Bland-Altman analyses of the mean sensitivity (MS). Dotted line represent mean and the limits of agreement (mean +/-(1.96*SD)).

DISCUSSION

The use of an alternative psychophysical method for static perimetry, CLIP, did not result in a decrease in variability when compared to the 'gold standard' SP despite a shorter test time. In addition, CLIP showed a clear learning effect.

Wabbels et al. investigated variability in CLIP in ten healthy subjects.(Wabbels et al. 2001) After averaging the SD of each test location of each subject of each VF performed (10 repetitions), they found an averaged SD of 1.18 dB, which was comparable to that of the SITA Standard strategy (1.19 dB), but statistically lower compared to that of the SP strategy (1.44 dB). Our study population did not include healthy subjects. In our population of early glaucoma, as illustrated in Table 2, the variability seems lower in CLIP, but this difference was not significant (P=0.65 for CoR CLIP versus CoR SP in early glaucoma). Capris et al. compared the CLIP with the SITA fast (SITAf) strategy in 21 static perimetrically experienced glaucoma patients.(Capris et al. 2008) Interestingly, they found a higher learning effect in SITAf compared to CLIP (0.67 dB higher MS in the second SITAf test versus 0.39 dB in CLIP). BA-analyses were performed with pointwise sensitivity differences for both strategies and they appeared to be comparable. All these results are in line with our results.

We found a similar variability for both strategies, suggesting that the psychophysical method is not critical, which, in turn, could suggest that variability might be an intrinsic property of glaucoma. However, the strategies differed in at least two aspects that could influence variability: (1) test duration differed significantly between the strategies and (2) CLIP showed, unlike SP, a clear learning effect. A shorter test duration should have resulted in less variability; a learning effect, on the other hand, should have resulted in more variability.

The use of stimulus size V is currently reserved for advanced glaucoma only. Although this stimulus is standard available on the HFA, the advantage of a larger stimulus is lost because of the longer test duration of the FT strategy (compared to a faster SP procedure, SITA, only available for size III). Especially M. Wall investigated size V in SAP and found a smaller test-retest variability, the same sensitivity to detect abnormal test locations, and a greater effective dynamic range.(Wall et al. 1997, Wall et al. 2008, Wall et al. 2009, Wall et al. 2010) Despite stimulus size V is available in every strategy of the Twinfield perimeter, age-corrected thresholds are artificially determined as five decibels higher compared to size III thresholds. Therefore, no age-related deviation calculations could be performed. In addition, the starting intensity does not correspond with the actual hill of vision. As a consequence, test time increases as thresholds are

too close to the estimated threshold and are repeated, or too far and patients have to wait unnecessarily. Finally, in CLIP fixation losses were determined by a short foveal suprathreshold stimulus which was recommended by the manufacturer. As patients suspect a continuous stimulus somewhere in the VF, they often reacted too slowly on the catch trial. This explains the high percentages of 'fixation losses'.

In conclusion, test-retest variability was large and essentially identical for two, very different psychophysical methods. This favors the hypothesis that glaucoma itself is a major factor contributing to variability.

REFERENCES

Bland JM & Altman D (1986): Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1: 307–310.

Capris P, Autuori S, Capris E & Papadia M (2008): Evaluation of threshold estimation and learning effect of two perimetric strategies, SITA Fast and CLIP, in damaged visual fields. Eur J Ophthalmol. 18: 182–190.

Gardiner SK, Demirel S, Gordon MO & Kass MA (2013): Seasonal changes in visual field sensitivity and intraocular pressure in the Ocular Hypertension Treatment Study. Ophthalmology. 120: 724–730.

Heijl A, Lindgren A & Lindgren G (1989): Test-retest variability in glaucomatous visual fields. Am J Ophthalmol. 108: 130–135.

Junoy Montolio FG, Wesselink C, Gordijn M & Jansonius NM (2012): Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day, and season. Invest Ophthalmol Vis Sci. 53: 7010–7017.

Wabbels BK, Burk ROW & Kolling G (2001): CLIP: An improved strategy in automated static perimetry. Perimetry Update. 2000/2001 2001: 177–186.

Wabbels BK, Diehm S & Kolling G (2005): Continuous light increment perimetry compared to full threshold strategy in glaucoma. Eur J Ophthalmol. 15: 722–729.

Wall M, Brito CF, Woodward KR, Doyle CK, Kardon RH & Johnson CA (2008): Total deviation probability plots for stimulus size v perimetry: a comparison with size III stimuli. Arch Ophthalmol. 126: 473–479.

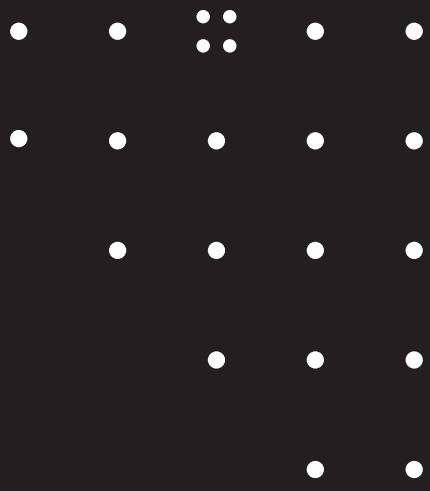
Wall M, Doyle CK, Zamba KD, Artes P & Johnson CA (2013): The repeatability of mean defect with size III and size V standard automated perimetry. Invest Ophthalmol Vis Sci. 54: 1345–1351.

Wall M, Kutzko KE & Chauhan BC (1997): Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. Invest Ophthalmol Vis Sci. 38: 426–435.

Wall M, Woodward KR, Doyle CK & Artes PH (2009): Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. Invest Ophthalmol Vis Sci. 50: 974–979.

Wall M, Woodward KR, Doyle CK & Zamba G (2010): The effective dynamic ranges of standard automated perimetry sizes III and V and motion and matrix perimetry. Arch Ophthalmol. 128: 570–576.

VII



Lateral inhibition in the human visual system in patients with glaucoma and healthy subjects:

A case-control study

Francisco G. Junoy Montolio, Wilma Meems, Marieke S.A. Janssens, Lucas Stam, Nomdo M. Jansonius

PLoS One. 2016;11(3):e0151006

ABSTRACT

Purpose: In glaucoma, the density of retinal ganglion cells is reduced. It is largely unknown how this influences retinal information processing. An increase in spatial summation and a decrease in contrast gain control and contrast adaptation have been reported. A decrease in lateral inhibition might also arise. This could result in a larger than expected response to some stimuli, which could mask ganglion cell loss on functional testing (structure-function discrepancy). The aim of this study was to compare lateral inhibition between glaucoma patients and healthy subjects; we used a case-control design.

Methods: Cases (n=18) were selected to have advanced visual field loss in combination with a normal visual acuity. Controls (n=50) were not allowed to have symptoms or signs of any eye disease. Lateral inhibition was measured psychophysically on a computer screen, with (1) a modified illusory movement experiment and (2) a contrast sensitivity (CS) test. Illusory movement was quantified by nulling it with a real movement; measure of lateral inhibition was the amount of illusory movement. CS was measured at 1 and 4 cycles per degree (cpd); measure of lateral inhibition was the difference between log CS at 4 and 1 cpd. Both measures were compared between cases and controls; analyses were adjusted for age and gender.

Results: There was no difference between cases and controls for these two measures of lateral inhibition (P=0.58 for illusory movement; P=0.20 for CS). The movement threshold was higher in cases than in controls (P=0.008) and log CS was lower, at both 1(-0.20; P=0.008) and 4(-0.28; P=0.001) cpd.

Conclusion: Our results indicate that spatially antagonistic mechanisms are not specifically affected in glaucoma, at least not in the intact center of a severely damaged visual field. This suggests that the structure-function discrepancy in glaucoma is not related to a decrease in lateral inhibition.

INTRODUCTION

Glaucoma is a chronic and progressive eye disease characterized by loss of retinal ganglion cells (RGCs) and subsequent visual field loss. It is largely unknown how the loss of RGCs influences information processing within the retina. An increase in Ricco's area has been described as well as changes in contrast gain control and contrast adaptation (see Discussion section).(Redmond et al. 2010) A decrease in lateral inhibition might also arise.(Sunga & Enoch 1970) This could result in a larger than expected response to some stimuli, which could mask RGC loss on functional glaucoma testing. A decrease in lateral inhibition may thus play a role in the presumed observation that structural loss precedes functional loss in glaucoma (structure-function discrepancy).

The aim of this study was to compare lateral inhibition in the visual system between glaucoma patients and healthy subjects. For this purpose we developed a new psychophysical method and applied this method to glaucoma patients and controls. We also performed contrast sensitivity (CS) measurements at 1 and 4 cycles per degree (cpd). The difference between the CS at 4 and 1 cpd is presumed to be a measure of lateral inhibition as well (the inhibition makes the visual system optimally tuned for spatial frequencies around 4 cpd - for photopic vision in the center of the visual field). (Kelly 1964, Levinson 1964)

METHODS

Study population

The present study was a case-control study and comprised 18 glaucoma patients (cases) and 50 healthy subjects (controls). The ethics board of the University Medical Center Groningen (UMCG) approved the study protocol. All participants provided written informed consent. The study followed the tenets of the Declaration of Helsinki.

Glaucoma patients were selected from visitors of the outpatient department of the department of Ophthalmology, UMCG, using the visual field database of the Groningen Longitudinal Glaucoma Study (GLGS), a prospective observational cohort study performed in a clinical setting.(Heeg et al. 2005) The subpopulation selected for the present study comprised open angle glaucoma patients (primary n=16, pigment dispersion n=1, pseudoexfoliation n=1) with (1) a visual field mean deviation (MD) of -12 dB or worse (as measured with standard automated perimetry [Humphrey Field Analyzer 30-2 SITA fast; Carl Zeiss, Jena, Germany]) and (2) a best-corrected visual

acuity (BCVA) of 0.0 logMAR or better (up to 50 years of age) or 0.1 logMAR or better (above 50 years), in at least one eye. If both eyes met the inclusion criteria, the eye with the lowest MD value was chosen.

Healthy subjects were recruited by advertisement. We aimed for subjects between 40 and 70 years of age, at least 15 subjects per decade, with a ratio of approximately 3 controls per case. First, healthy volunteers who responded to the advertisement were asked to complete a questionnaire to screen for any known eye abnormality and a positive family history of glaucoma (exclusion criteria). After this preselection, the subjects completed an eye examination, including a BCVA measurement, a non-contact intraocular pressure (IOP) measurement (TCT80; Topcon Medical Systems, Oakland, USA), and a fundus examination with the Optos ultra-widefield retinal imaging device (200TX; Optos, Marlborough, USA). Exclusion criteria consisted of any known eye abnormality, a positive family history of glaucoma, a BCVA worse than 0.0 logMAR (up to 50 years of age) or 0.1 logMAR (above 50 years), an IOP above 21 mmHg, a vertical cup-to-disc ratio above 0.7(Wolfs et al. 2000) or any other fundus abnormality (as observed by an ophthalmologist [NJ] who evaluated the Optos images and all other available data; in case of doubt, the subject was re-invited and a full eye exam was performed including fundoscopy in mydriasis as well as laser polarimetry of the optic nerve head (GDx ECC; Carl Zeiss, Jena, Germany) and a frequency doubling technology visual field test (FDT C20-1 screening mode; Carl Zeiss, Jena, Germany). A GDx VFI above 35 or any reproducibly abnormal test location at $P<0.01$ on the FDT test result implied exclusion. If both eyes were eligible, the dominant eye was chosen according to the Dolman method.(Jansonius et al. 2014)

Lateral inhibition

Two different psychophysical experiments were performed. Both experiments target spatially antagonistic mechanisms, of which the physiological equivalent is presumed to be lateral inhibition (see Discussion section). The experiments were carried out monocularly, in a sparsely illuminated room (luminance of the wall typically 10 cd/m^2; luminance of the screen (see below) if switched off < 1 cd/m^2). No cycloplegia, mydriasis, or artificial pupil was used. All experiments were performed with optimal correction for the viewing distance. Pupil diameter was measured with a ruler, while the subject was looking at the stimulus with the contralateral eye occluded (as was the case during the experiments).

Illusory movement

The first experiment is based on a psychophysical phenomenon called illusory movement. Illusory movement has been described in detail by Jansonius et al.. (Jansonius et al. 2014) In short, a narrow bar or line (width around 1 arcminute) between two fields of which the luminances are sinusoidally and in counterphase modulated in time (Figure 1A) appears to make an oscillatory movement. It is possible to annihilate this illusory movement with a real movement and thus to analyze this phenomenon quantitatively. The phenomenon can be explained by a model that includes low-pass spatial filtering in the visual system. With some modification, that is, replacing the modulated fields by modulated stripes positioned at a certain distance from the line (Figure 1B), a 180 degree phase shift occurs and this shift is presumed to reflect lateral inhibition.(Kooi & Kuiper 1986; Jansonius & Kuiper 1989) Following Kooi and Kuiper(Kooi & Kuiper 1986) and Jansonius and Kuiper,(Jansonius & Kuiper 1989) we used a line width of 1.2 arcminute and a distance between the line and the border of the stripes of 6 arcminute. Modulation depth was 0.08; modulation frequency 2.5 Hz.

The stimulus was generated on an LCD monitor (Samsung SyncMaster 2243WM). The refresh rate was 75 Hz; the mean luminance of the screen 150 cd/m^2 as measured with a Minolta luminance meter with built-in photometric filter (LS-110; Minolta Camera Co. Ltd., Japan). The refresh rate is far beyond the critical fusion frequency, which is about 40 Hz at this luminance.(de Lange Dzn 1954, Hecht & Verrijp 1933) The stimulus was generated and the data were collected using Octave (version 3.2.4; www.gnu.org/software/octave/) for Linux (Ubuntu 10.10) in combination with the Psychophysics Toolbox (PTB-3).(Brainard 1997, Pelli 1997) The distance between subject and screen was 6 m. The screen size was about 4 degrees (0.4 m at 6 m); the stimulus size was limited by a digital mask to 1 by 1 degree. The luminances of the area outside the mask, the areas between the line and the stripes, and the mean luminances of the modulated stripes were equal (150 cd/m^2).

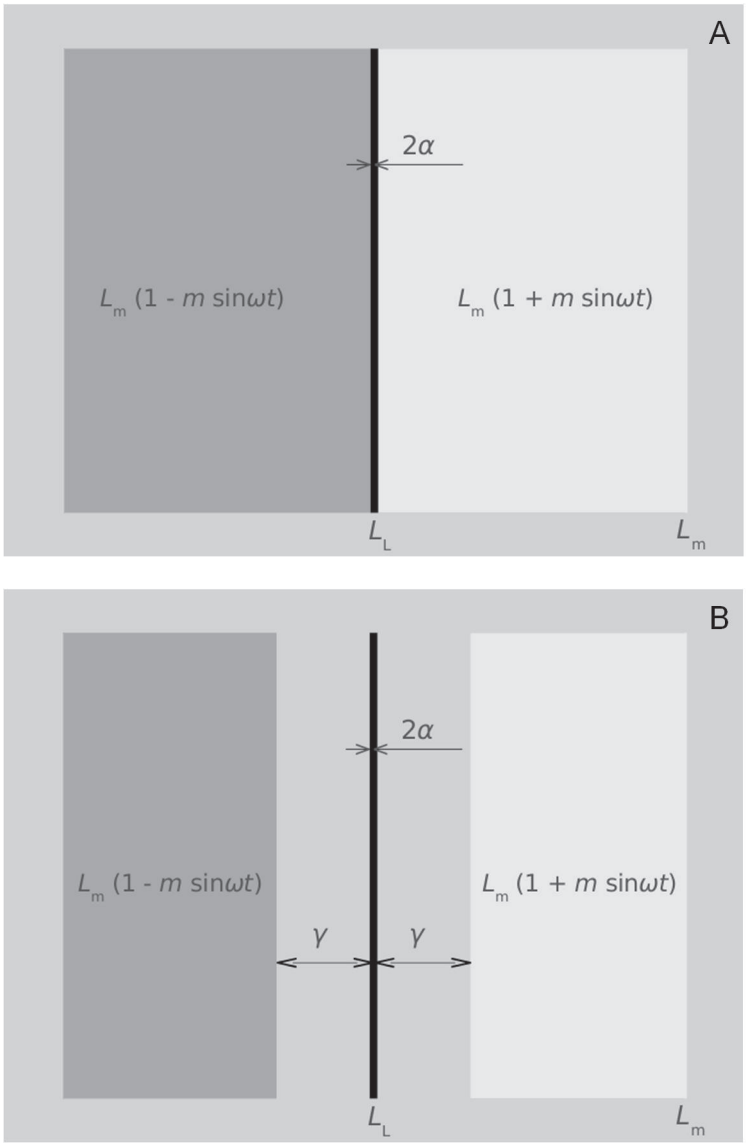


Figure 1A & B. Original (A) and modified (B) illusory movement stimulus. 2α is the line width, L_m the mean luminance, L_L the luminance of the line, m the modulation depth, ω/2π the modulation frequency, and γ the distance between the line and the border of the stripes.

We measured the illusory movement by compensating it with a real movement, using a staircase method. Figure 2 illustrates this method. The stimulus is presented during 4 seconds and the subject has to report yes/no movement observed. In the beginning, no real movement is added (in this situation, no movement is observed because the threshold for seeing movement is larger than the illusory movement itself). Subsequently, real movement is added in steps of 5 arcsecond (amplitude of the oscillatory movement) until movement is observed, subsequently removed in steps of 2 arcsecond until no movement is observed, added in steps of 2 arcsecond until movement is observed, and so on. In this way, six reversals are collected. Another six reversals are collected by making the initial 5 arcsecond steps in the opposite direction. These two series of reversals are collected alternately (Figure 2). From these two series of reversals, we calculated two thresholds: an overcompensation threshold and an undercompensation threshold. For each threshold we calculated two median values from the corresponding reversals (see Figure 2: median of 1,3,5 and median of 2,4,6 for the upper threshold; median of 1',3',5' and median of 2',4',6' for the lower threshold) and averaged the two medians. Two outcome measures were subsequently calculated: (1) the amount of illusory movement (the measure of lateral inhibition), which is the average of the overcompensation threshold and the undercompensation threshold and (2) the movement threshold (the difference threshold between illusory movement and real movement), which is the difference between the overcompensation and undercompensation threshold divided by two.

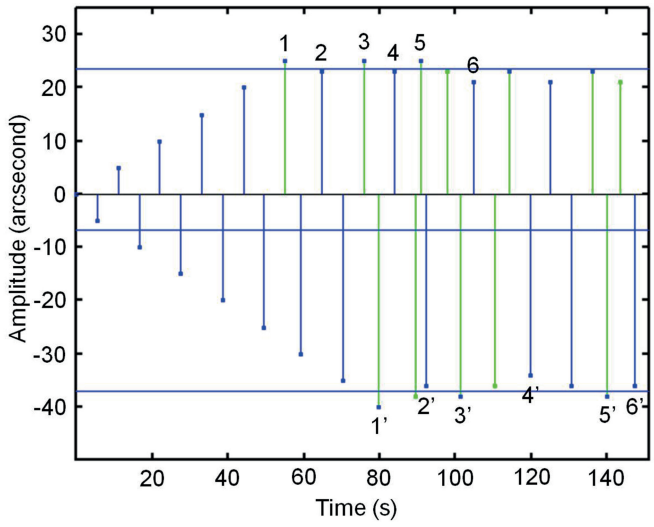


Figure 2. Staircase method applied to the illusory movement experiment.

The experiment was performed three times, preceded by a short try-out. The median test results of these three experiments was the final test result.

Contrast sensitivity

The second experiment consisted of a sine-wave grating CS test, using two spatial frequencies: 1 and 4 cpd. Here, the psychophysical method was a tracking method according to von Békésy.(Von Békésy 1975, Jansonius & Kooijman 1997) At the beginning of the experiment, contrast is negligible and gradually increases. If the subject observes the sine-wave grating, a button is pushed and held, resulting in a gradual decrease in contrast. If the grating disappears, the button is released and the contrast increases again, and so on. Contrast changed with a speed of 0.3 log/s and a contrast threshold was calculated from 12 (6 upper and 6 lower) reversals. For details see Nio et al..(Nio et al. 2005) Contrast sensitivity is the reciprocal of the contrast threshold. Contrast is defined as $(L_{max}\text{-}L_{min})/(L_{max}+L_{min})$, where L_{max} and L_{min} are the maximum and minimum luminance on the screen, respectively. Hardware and software were identical to that of the illusory movement experiment; the mean luminance of the screen was 150 cd/m^2. Testing distance was 4m. The sine-wave gratings, which were oriented vertically, filled the entire screen, resulting in a stimulus size of approximately 6 (horizontally) by 4 (vertically) degrees.

The experiment was performed two times per spatial frequency, preceded by a short try-out. The mean value per spatial frequency was the final CS test result.

Statistical analysis

Nonparametric descriptive statistics (median with interquartile range [IQR]) was used to describe the study population; the corresponding univariable comparisons between cases and controls were made with a Mann-Whitney test. Proportions were compared with the Chi-square test. Yates (continuity) correction was applied if applicable. Multiple linear regression was used to compare the outcome measures between cases and controls adjusted for age and gender. All analyses were performed using R (version 2.11.1; R Foundation for Statistical Computing, Vienna, Austria). A *P*-value of 0.05 or less was considered statistically significant.

RESULTS

Table 1 presents the general characteristics of the study population. Glaucoma patients were older (p<0.001) and more often male (p=0.012) compared to the healthy subjects. Visual acuity was statistically lower (logMAR higher) in the glaucoma patients (median [IQR] logMAR 0.00[0.00 to 0.05]) than in the healthy subjects (-0.08[-0.08 to 0.00]; *P*<0.001). The mean difference was 0.07. The difference became smaller after adjustment for age (0.05; *P*=0.02; a logMAR difference of 0.05 corresponds to half a Snellen line). The glaucoma patients had a median (IQR) visual field MD of -23.5(-26.9 to -17.2) dB. The pupil diameter did not differ between the groups (*P*=0.16; *P*=0.41 when adjusted for age; based on 14 cases and 17 controls for whom pupil diameter data were available).

Table 1. General characteristics of the study population

	Cases (n=18)	Controls (n=50)	*P*-value
Age (year; median [IQR])	70 (67 to 72)	55 (47 to 62)	<0.001
Gender (% female)	28	66	0.012
Pupil diameter (mm; median [IQR])*	4.3 (4.0 to 5.3)	5.0 (4.0 to 6.0)	0.16
Visual acuity (logMAR; median [IQR])	0.00 (0.00 to 0.05	-0.08 (-0.08 to 0.00)	<0.001

** = based on a subset of 14 cases and 17 controls with pupil-diameter data; IQR = interquartile range.*

Figure 3 presents the results of the illusory movement experiment. There was no significant difference between cases and controls in a univariable analysis (*P*=0.61) with a median (IQR) of -3.0(-4.4 to -1.0) and -2.6(-4.3 to -1.0) arcsecond for the cases and controls, respectively. Table 2 shows the results of the multivariable analysis. No significant difference between cases and controls was found when adjusted for age and gender (*P*=0.58).

Figure 4 presents the results of the second experiment, the difference between the CS at 4 and 1 cpd. This experiment showed - in the univariable analysis - a significantly lower lateral inhibition in the glaucoma patients compared to the healthy subjects (median [IQR] 0.27[0.20 to 0.32] versus 0.36[0.29 - 0.44]; *P*=0.006). Table 2 shows the results of the corresponding multivariable analysis. No significant difference between cases and controls was found when adjusted for age and gender (*P*=0.20).

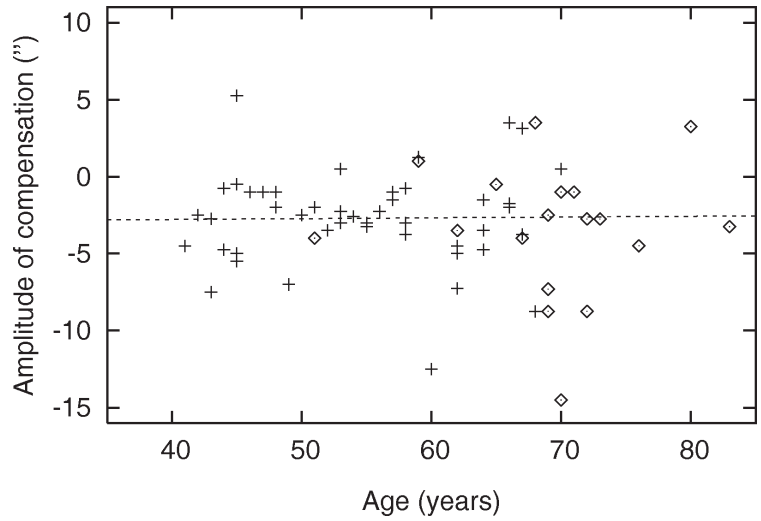


Figure 3. Amount of illusory movement as a function of age, for cases (◊) and controls (+). Regression line refers to the controls.

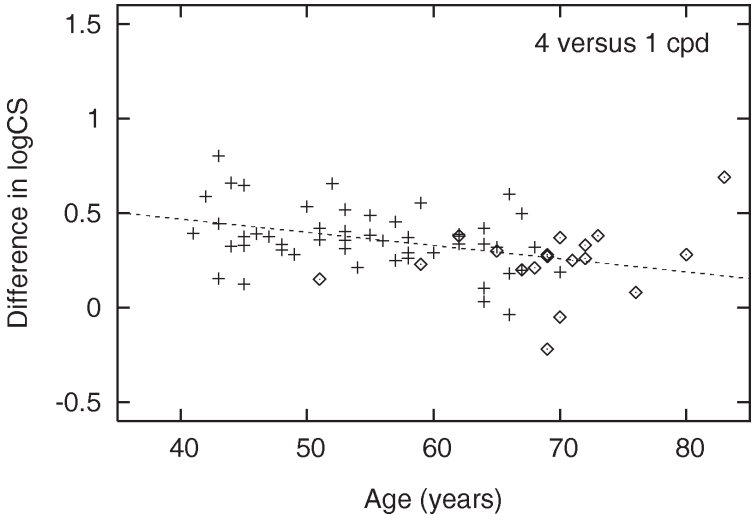


Figure 4. Difference in contrast sensitivity between 4 and 1 cpd as a function of age, for cases (◊) and controls (+). Regression line refers to the controls.

Table 2. Multivariable regression analysis showing the lateral inhibition outcome measures (amount of illusory movement and contrast sensitivity difference between 4 and 1 cpd), the movement threshold, and the contrast sensitivity test results, adjusted for age and gender

		Beta	*P*-value
Illusory movement*	Glaucoma	-0.740	0.58
	Age	0.005	0.93
	Gender	0.093	0.92
Movement threshold*	Glaucoma	13.820	0.008
	Age	0.208	0.33
	Gender	8.037	0.03
CS4-CS1	Glaucoma	-0.079	0.20
	Age	-0.004	0.11
	Gender	-0.045	0.31
CS1	Glaucoma	-0.204	0.008
	Age	0.002	0.60
	Gender	-0.023	0.67
CS4	Glaucoma	-0.283	0.001
	Age	-0.003	0.48
	Gender	-0.068	0.27

** = based on 18 cases and 46 controls (not all participants were able to perform this test); CS1 = contrast sensitivity measured at 1 cpd; CS4 = contrast sensitivity measured at 4 cpd.*

Figure 5 presents the underlying CS measurements at 1 and 4 cpd, for the cases and controls. Cases had a statistically significantly lower CS compared to controls, both at 1 (P=0.008) and at 4 (P<0.001) cpd. These differences remained equally significant when adjusted for age and gender (Table 2).

The movement threshold (that is, the difference threshold between illusory movement and real movement) was larger in the glaucoma patients (median [IQR] 46[31–60] arcsecond) than in the controls (30[23–39] arcsecond; P=0.003). A significant difference was also found in the multivariable analysis (Table 2).

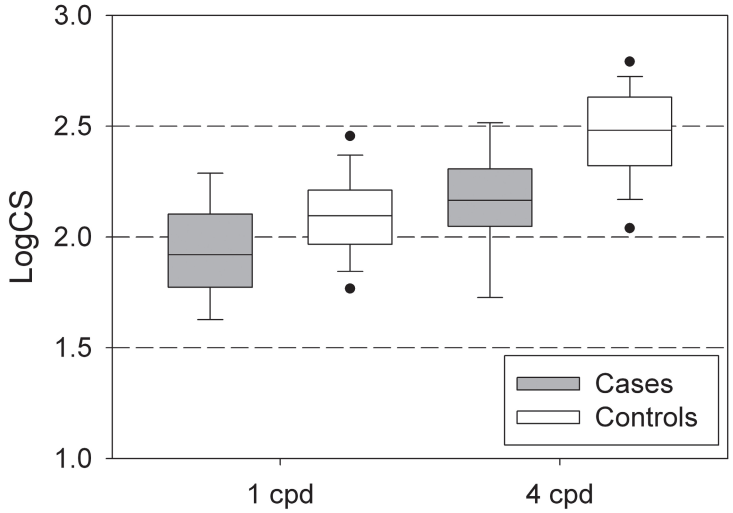


Figure 5. Contrast sensitivity at 1 and 4 cpd, for cases and controls (percentiles: 5, 10, 25, 50, 75, 90, and 95).

DISCUSSION

Spatially antagonistic mechanisms are not specifically affected in glaucoma, at least not in the center of the visual field of patients with severe glaucoma and a normal visual acuity. The CS of the included patients was approximately half as high as the CS of the controls. Their movement threshold was significantly increased.

We found a single description of a decrease in lateral inhibition in glaucoma in the literature.(Sunga & Enoch 1970) Sunga and Enoch used an experimental method as described by Westheimer(Westheimer 1967).(Sunga & Enoch 1970, Westheimer 1967) They performed measurements both within relative scotoma and in apparently normal areas of the visual field (at an eccentricity of at least 5 degrees), in four patients with glaucomatous visual field loss. They found more lateral inhibition in the unaffected parts of the visual field than in the scotoma and hypothesized a retrograde damage of the synapses involved in lateral inhibition. We performed the measurements exclusively in an apparently unaffected area of the visual field. As such, our findings are in agreement with that of Sunga and Enoch. We focused on an unaffected area because our hypothesis (see Introduction section) was that the visual field would appear normal because an actual decrease in sensitivity was masked by a decrease in lateral inhibition. Given the glaucoma stage of the included patients, a significant thinning of the RGC layer is to be expected in the investigated area of the

retina - despite the apparently unaffected function. Twelve of the 18 cases had had an assessment of the combined RGC layer/retinal nerve fiber layer/inner plexiform layer thickness in the macular area (6x6 mm scan centered around the fovea) with optical coherence tomography (OCT) as part of their regular glaucoma care. In 10 (83%) of them the thickness was outside normal limits according to the built-in normative database of the clinical device (Canon OCT HS-100; software version 1.0).

In his review paper on the control of sensitivity in the retina, Werblin described two types of lateral interactions: antagonistic effects in stationary patterns mediated by horizontal cells (lateral inhibition in a narrow sense) and antagonistic effects in changing patterns mediated by amacrine cells.(Werblin 1973) The former is related to predictive coding;(Srinivasan et al. 1982) the latter to contrast gain control(Meister & Berry 1999, Shapley & Victor 1978) and contrast adaptation.(Blakemore & Campbell 1969, Meister & Berry 1999) Our stimuli were presumably targeting the former. For the - static - stimuli used to assess contrast sensitivity this is clear. The modulation used in the illusory movement experiment could theoretically trigger the contrast gain control mechanism as this mechanism already starts below our modulation frequency of 2.5 Hz; the applied modulation depth of 0.08, however, is much lower than used in psychophysical experiments targeting contrast gain control.(Hood et al. 1997, Snippe et al. 2000) Moreover, as the illusory movement was assessed by nulling it with a real movement, contrast gain control should not influence it.(Jansonius et al. 2014) Contrast gain control and contrast adaptation have been shown to be affected in glaucoma.(Dul et al. 2015, McKendrick et al. 2004, Sun et al. 2008) Together with our results this suggests that glaucoma affects the inner retina more than the outer retina. However, there is some electrophysiological evidence for a generalized outer retina involvement in glaucoma(Vaegan et al. 1995) and histological changes in horizontal cells in humans with glaucoma have been described as well.(Janssen et al. 1996) More recently, loss of horizontal cells has been described in mice.(Fernández-Sánchez et al. 2014) This apparently contradicts earlier studies in the same animal.(Fuchs et al. 2012, Moon et al. 2005)

The significant decrease in CS as found in this study is in line with earlier reports,(Lustgarten et al. 1990, Ross et al. 1984, Wood 1992) although less clear effects have been published as well.(Lundh 1985, Shabana et al. 2003) The observed decrease of 0.2 to 0.3 log units corresponds to almost a halving of the CS. Further, the movement threshold displayed a significant increase of approximately 50% (0.2 log units) in the glaucoma patients compared to the healthy subjects. A decrease in motion sensitivity has been described earlier, also in apparently normal areas of the visual field of glaucoma patients and in patients with ocular hypertension.(Shabana et al. 2003)

A normal lateral inhibition in glaucoma is not conflicting with an enlargement of the area of complete summation, also known as Ricco's area, as recently described by Redmond et al..(Redmond et al. 2010) Due to lateral inhibition, mainly the RGCs at the edge of the stimulus contribute to the signal. This signal is reduced in glaucoma because of the loss of RGCs – the increase in Ricco's area, which implies an increase in the circumference (edge), compensates for this. If lateral inhibition would have been reduced, the RGCs in the center of the stimulus would contribute to the signal as well, which could initially mask glaucomatous RGC loss: our – falsified – hypothesis.

A limitation of this study is the difference in age distribution between the cases and the controls. Especially the illusory movement test is a subjective measurement and not easy to perform. For that reason we originally aimed to include subjects between 40 and 70 years of age. However, glaucoma is a disease of the elderly, and, as a consequence, the vast majority of the patients with severe glaucoma in our database was between 60 and 80 years old. This resulted in different age distributions in cases and controls, which hampered a direct comparison of the groups. However, the groups showed a considerable overlap in age and we adjusted all analyses for age. Also, none of the lateral inhibition measures showed a significant age dependency (Table 2). All this indicates that the different age distributions will not have influenced our major findings.

In conclusion, patients with severe glaucoma and a normal visual acuity and healthy controls display similar spatially antagonistic mechanisms in the central part of the visual field. Future research may focus on eccentric visual field areas, areas with a reduced sensitivity, and changing patterns.

REFERENCES

Von Békésy G (1975): Sensory Inhibition. Princeton: Princeton University Press.

Blakemore C & Campbell FW (1969): On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. J Physiol. 203: 237–260.

Brainard DH (1997): The psychophysics toolbox. Spat Vis. 10: 433–436.

Dul M, Ennis R, Radner S, Lee B & Zaidi Q (2015): Retinal adaptation abnormalities in primary open-angle glaucoma. Invest Ophthalmol Vis Sci. 56: 1329–1334.

de Lange DznH (1954): Relationship between critical flicker-frequency and a set of low-frequency characteristics of the eye. J Opt Soc Am. 44: 380-389.

Fernández-Sánchez L, de Sevilla Müller LP, Brecha NC & Cuenca N (2014): Loss of outer retinal neurons and circuitry alterations in the DBA/2J mouse. Invest Ophthalmol Vis Sci. 55: 6059-6072.

Fuchs M, Scholz M, Sendelbeck A, Atorf J, Schlegel C, Enz R & Brandstätter JH (2012): Rod photoreceptor ribbon synapses in DBA/2J mice show progressive age-related structural changes. PLoS One. 7: e44645.

Hecht S & Verrijp CD (1933): Intermittent stimulation by light: III. The relation between intensity and critical fusion frequency for different retinal locations. J Gen Physiol. 17: 251–268.

Heeg GP, Blanksma LJ, Hardus PLLJ & Jansonius NM (2005): The Groningen Longitudinal Glaucoma Study. I. Baseline sensitivity and specificity of the frequency doubling perimeter and the GDx nerve fibre analyser. Acta Ophthalmol Scand. 83: 46–52.

Hood DC, Graham N, von Wiegand TE & Chase VM (1997): Probed-sinewave paradigm: a test of models of light-adaptation dynamics. Vision Research. 37: 1177–1191.

Jansonius M, Stam L, de Jong T & Pijpker BA (2014): Quantitative analysis of illusory movement: spatial filtering and line localization in the human visual system. Perception. 43: 1329–1340.

Jansonius NM & Kooijman AC (1997): The effect of defocus on edge contrast sensitivity. Ophthalmic Physiol Opt. 17: 128–132.

Jansonius NM & Kuiper JW (1989): Quantitative analysis of illusory movement in the human visual system. Neural Mechanisms of Behaviour. 202.

Janssen P, Naskar R, Moore S, Thanos S & Thiel HJ (1996): Evidence for glaucoma-induced horizontal cell alterations in the human retina. Ger J Ophthalmol. 5: 378–385.

Kelly DH (1964): Sine waves and flicker fusion. Doc Ophthalmol. 18: 16–35.

Kooi F & Kuiper JW (1986): Apparent motion induced by contrast modulation. Invest Ophthalmol Vis Sci. ARVO abstract 93.

Levinson J (1964): Nonlinear and spatial effects in the effects in the perception of flicker. Doc Ophthalmol. 18: 36–55.

Lundh BL (1985): Central and peripheral contrast sensitivity for static and dynamic sinusoidal gratings in glaucoma. Acta Ophthalmol. 63: 487–492.

Lustgarten JS, Marx MS, Podos SM, Bodiswollner I, Campeas D & Serle JB (1990): Contrast sensitivity and computerized perimetry in early detection of glaucomatous change. Clin Vis Sci. 5: 407–413.

McKendrick AM, Badcock DR & Morgan WH (2004): Psychophysical measurement of neural adaptation abnormalities in magnocellular and parvocellular pathways in glaucoma. Invest Ophthalmol Vis Sci. 45: 1846–1853.

Meister M & Berry MJ 2nd (1999): The neural code of the retina. Neuron. 22: 435–450.

Moon JI, Kim IB, Gwon JS, Park MH, Kang TH, Lim EJ, Choi KR & Chun MH (2005): Changes in retinal neuronal populations in the DBA/2J mouse. Cell Tissue Res. 320: 51–59.

Nio YK, Jansonius NM, Lamers P, Mager A, Zeinstra J & Kooijman AC (2005): Influence of the rate of contrast change on the quality of contrast sensitivity assessment: a comparison of three psychophysical methods. Ophthalmic Physiol Opt. 25: 18–26.

Pelli DG (1997): The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis. 10: 437–442.

Redmond T, Garway-Heath DF, Zlatkova MB & Anderson RS (2010): Sensitivity loss in early glaucoma can be mapped to an enlargement of the area of complete spatial summation. Invest Ophthalmol Vis Sci. 51: 6540–6548.

Ross JE, Bron AJ & Clarke DD (1984): Contrast sensitivity and visual disability in chronic simple glaucoma. Br J Ophthalmol. 68: 821–827.

Shabana N, Cornilleau Pérès V, Carkeet A & Chew PTK (2003): Motion perception in glaucoma patients: a review. Surv Ophthalmol. 48: 92–106.

Shapley RM & Victor JD (1978): The effect of contrast on the transfer properties of cat retinal ganglion cells. J Physiol. 285: 275–298.

Snippe HP, Poot L & van Hateren JH (2000): A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. Vis Neurosci. 17: 449–462.

Srinivasan MV, Laughlin SB & Dubs A (1982): Predictive coding: a fresh view of inhibition in the retina. Proc R Soc Lond B Biol Sci. 216: 427–459.

Sunga RN & Enoch JM (1970): A static perimetric technique believed to test receptive field properties. 3. Clinical trials. Am J Ophthalmol. 70: 244–272.

Sun H, Swanson WH, Arvidson B & Dul MW (2008): Assessment of contrast gain signature in inferred magnocellular and parvocellular pathways in patients with glaucoma. Vision Res. 48: 2633–2641.

Vaegan, Graham SL, Goldberg I, Buckland L & Hollows FC (1995): Flash and pattern electroretinogram changes with optic atrophy and glaucoma. Exp Eye Res. 60: 697–706.

Werblin FS (1973): The control of sensitivity in the retina. Sci Am. 228: 70–79.

Westheimer G (1967): Spatial interaction in human cone vision. J Physiol. 190: 139–154.

Wolfs RC, Borger PH, Ramrattan RS, Klaver CC, Hulsman CA, Hofman A, Vingerling JR, Hitchings RA & de Jong PT (2000): Changing views on open-angle glaucoma: definitions and prevalences--The Rotterdam Study. Invest Ophthalmol Vis Sci. 41: 3309–3321.

Wood JM (1992): Contrast sensitivity deficits in subjects with glaucoma-like discs using an oscilloscope-based method. Acta Ophthalmol. 70: 610–614.

VIII

Summary and conclusions, clinical implications and future directions

Francisco G. Junoy Montolio

SUMMARY AND CONCLUSIONS

1. Summary of findings

The main objective of this thesis is to improve glaucoma care by optimizing the information that can be obtained from visual field (VF) testing. Therefore a glimpse of light is shed on VF testing to better understand how perimetry works and which factors influence the test results. By this, we hopefully contribute to preventing patients from blindness and preserving quality of life. After introducing the topic of this thesis in **Chapter I,** we focus on a better understanding of VF testing in healthy subjects and/or patients with glaucoma: **Chapter II, III and IV** describe how different intrinsic and extrinsic factors influence VF testing. In **Chapter V**, we compare the cost of glaucoma surgery on visual field function (loss of visual functioning due to the surgery) with the long term benefit (slowing down disease progression). We go back to basics in **Chapter VI** and conclude that an alternative psychophysical method for measuring the sensitivity of the eye does not perform better compared to the current standard, despite a shorter testing time. This favors the hypothesis that glaucoma itself is a major factor contributing to variability. Lastly, in **Chapter VII**, we challenge the hypothesis of decreased lateral inhibition in glaucoma patients (a possible explanation for normal VF test results in patients with a glaucomatous optic nerve). We will explicate the main findings in the following section.

Chapter II - The glaucomatous VF is characterised by healthy parts, parts that are affected but still have function, and parts that are (perimetrically) blind. This chapter shows that particularly the affected areas of the VF can give two totally different values if tested twice (high test-retest variability). In addition, this chapter shows that the blind parts of the VF can be omitted without loss of information - they do not contribute to distinguishing stable from progressive disease. Employment of this knowledge could result in a shorter test time and less fatigue for the patient.

Chapter III - Many different factors can influence test results, particularly in a subjective test like perimetry. This chapter proves and quantifies to what extent factors like technician experience, season, and time of the day influence perimetry results. In addition, of the three reliability indices displayed by the perimeter, false positive answers are the only true indication of unreliability. False negative answers indicate unreliability to a lesser extent and only in early glaucoma, while fixation losses seem unrelated to the test result.

Chapter IV - Spectacle independence is a prospect every presbyopic or ametropic patient would wish for after (clear) lens extraction. The benefit of multifocal intraocular lenses (MFIOLs) should be weighed against the drawback of decreased visual function, as is shown in this chapter. The diffractive property of MFIOLs causes a decrease in sensitivity of 2 dB (if tested with perimetry), which might be a negligible decrease in daily practice. For a patient with well regulated glaucoma however, a 2 dB sensitivity loss gives the patient a ten year leap of VF loss after implanting an MFIOL - added upon an already compromised visual sensitivity. The additional necessity of a new VF baseline is only a minor issue compared to the visual implications. Lastly, in the case of patients with other ophthalmic comorbidities like macular degeneration, any extra loss of retinal sensitivity will have clinical relevance. Therefore, patients with glaucoma or macular degeneration should be discouraged having MFIOL implants.

Chapter V - From the patients perspective, glaucoma surgery can be an ungrateful solution as patients might develop complaints after the intervention. Indeed, this chapter shows a decrease in VF sensitivity after the intervention, without a significant loss of visual acuity. Not performing surgery and accepting a high rate of progression, however, catches up with the surgery-induced damage already after 1.5 years. From that moment on patients with a Baerveldt glaucoma implant or trabeculectomy benefit from surgery compared to patients without intervention.

Chapter VI - Only a few methods are used in standard daily practice to measure the VF. This chapter evaluated an alternative commercially available psychophysical method: the Continuous Light Increment Perimetry (CLIP). This method shows comparable test-retest variability compared to the gold standard, the full threshold strategy, despite a shorter time to perform the test.

Chapter VII - Different psychophysical laws interact with each other to register incoming light and transform it to electric pulses that are sent to the brain. This equilibrium of psychophysical laws, including lateral inhibition, work well in a healthy eye, but might be deflected during glaucoma. The loss of lateral inhibition might contribute to the structure-function discrepancy (i.e., structural loss apparently precedes functional loss). In this chapter, however, no evidence is found that lateral inhibition is decreased in patients with severe glaucoma and an intact central VF.

2. Clinical implications

A large part of this thesis was performed with the clinicians in mind and the findings can be used for everyday practice. The next section will further explicate these findings.

Firstly, the subjective nature of perimetry makes the results susceptible to the influence of several different factors. Any clinician should be aware of these factors and how they affect the perimetrical end point for both the diagnosis and progression detection of glaucoma: Mean Deviation (MD). The first section will explain these factors in more detail and how to cope with them. The second part will explicate other issues that concern perimetry.

- **Glaucoma -** The disease glaucoma itself is an intrinsic factor that contributes to a high test-retest variability, probably related to ganglion cell saturation.(Swanson et al. 2011) Therefore, a clinician should be aware that varying test results might be a sign of glaucoma, rather than a patient's lack of concentration. The increase in test time in severe glaucoma could be another explanation for the increase in test-retest variability, and is a weakness of the SITA strategy (even for the recently published SITA faster (SFR)(Heijl et al. 2019)).

- **Season, time of day and technician experience -** Although these three factors have a small overall influence on perimetry of 0.5 dB, it can influence both event and trend based progression detection. Seasonal and diurnal influences are difficult to take into account in daily practice however, but investment in (experienced) technicians can decrease the variability reasonably. The decreased variability gives the clinician the possibility to distinct progression from stable disease earlier and more reliable.

- **Multifocal intraocular lenses (MFIOLs) -** The decline of retinal sensitivity of 2 dB caused by implantation of a MFIOL has different clinical implications for different patient categories. First, for a patient with cataract as the only ophthalmic condition, the influence might be negligible in daily practice. For a patient with glaucoma, however, a 2 dB sensitivity loss gives the patient a ten year leap of VF sensitivity loss after implanting a MFIOL - added upon an already compromised visual sensitivity. The additional necessity of a new VF baseline is only a minor issue compared to the visual implications. Lastly, in the case of patients with other ophthalmic comorbidities like macular degeneration, any extra loss of retinal sensitivity will have clinical relevance. Therefore, patients with glaucoma or retinal diseases should be discouraged having MFIOL implants.

- **Psychophysical methods -** The staircase method is the strategy usually used for the detection and follow up of glaucoma in clinical practice. The degree of variability is comparable in alternative methods that are commercially available, like CLIP, shown in **Chapter VI**. Mainly because a follow up of 5 years is required to reliably distinguish progression from stable disease state, the current used method remains the best option in glaucoma care for the time being; interchanging between strategies means starting a new trajectory that lasts 5 years before the glaucoma state can be reevaluated and therefore not recommended.(Jansonius 2010)

Besides these factors, other topics have to be assessed when VFs are reviewed by clinicians:

- **Reliability indices -** Of the three indices, only the percentage of false positive answers appeared to be a true and important measure of (un)reliability. More specifically, a high percentage of false positives answers indicates overestimation of the true sensitivity, and the commonly used upper limit of 10% (or sometimes 15%) corresponds to an overestimation of 1 dB. Hence, every clinician should particularly take this factor into account analysing a VF. Finally, partly because of these shortcomings, the new SFR strategy replaced two reliability indices and removed the third (i.e. the catch trial to measure false negative answers).

- **30-2 to 10-2 switch -** In advanced glaucoma, only a small part of the VF contributes to perimetry, containing little but essential clinical information. To regain information, the grid can be changed from a 30-2 (or 24-2) to a 10-2 grid, although this thesis shows that there is no clear MD cut-off point defined to make this transition. A residual VF might present as a classic central island, but a residual temporal island which lies outside the 10-2 grid also occurs, or even a combination of both. Therefore, the clinician should take the residual VF location into account while making the choice to switch to a 10-2 grid.

- **Surgery -** The prevention of blindness is a common goal for both the patient and clinician. When the clinician foresees blindness - a rate of progression that corresponds with blindness before a patient departs from life - the eye pressure needs to be lower than the current pressure. For this purpose, surgery is sometimes required despite the development of surgery induced complaints. Figure 4 (page 181) in **Chapter V** is illustrative and therefore useful to explain the necessity of

surgery to patients; the short term loss of visual field function (MD; drop in MD at time point 0, moment of surgery) and the long term benefit of a lower rate of progression after surgery (crossing of continuous and dashed line).

- **Lateral Inhibition -** The reason for structure-function discrepancy in glaucoma assessments is unknown and can be caused by the variation in measurement of either structure or function. A clinical example is a glaucomatous optic nerve combined with a normal VF or vice versa. Lateral inhibition might have been the missing link to explain this discrepancy, and could have led to the development of new perimetrical tests. However, this thesis shows that lateral inhibition does not play a significant role in this discrepancy. Therefore, it can be disputed that the answer for the above mentioned discrepancy (only) lies in perimetry.

3. Future directions

As stated in this thesis, perimetry is essential in glaucoma care, but has its limitations. There are several issues that should be addressed to improve glaucoma care and help clinicians to treat patients properly.

Several chapters of this thesis show the subjective nature of perimetry. This might give the clinician the impression that perimetry is not very helpful to diagnose glaucoma or to detect glaucoma progression. A correctional algorithm to give an overall MD taking many factors into account would be the ideal solution, but is an utopia in the short term. Every clinician needs to keep in mind the clinical relevant factors that might influence the MD and therefore influence the progression analysis, as mentioned in the previous section. Keeping these factors in mind, perimetry is still a valuable test, clearly illustrated in **Chapter IV and V** showing the effect of MFIOL and glaucoma surgery on the VF.

The first recommendation to optimize VF testing is to give less attention to unreliable parts of the VF; parts with low sensitivity. These parts show high test-retest variability that is part of glaucoma rather than a weakness of perimetry. Turpin et al. recently published the ARREST method, where sensitivities below 17 dB are marked either as 'defect' or 'blind', ignoring the parts which give noise to the MD.(Turpin et al. 2018) With the additional benefit of a shorter test time, another opportunity arises. As parts with low sensitivities give less information for the clinician, these parts can be (re-) measured with a larger stimulus, as proposed by Wall et al. increasing the dynamic range of perimetry.(Wall et al. 1997) Gardiner et al. suggested an alternative approach with the use of different stimuli in different locations, as eccentricity plays a role in variability as well.(Gardiner 2018) In addition, test locations can be added in the residual, and therefore, precious parts of the VF in case of severe glaucoma (i.e. central and/or temporal part of the VF).

The incorporation of intrapatient VF data is the second recommendation. Both eye structure (like refraction, for example) and disease state could be taken into account in positioning test points. Therefore, we should leave the idea of perimetry based on normative data that should fit the whole population and develop custom made VFs. Implementing individual anatomical characteristics is already investigated by Lamparter et al..(Lamparter et al. 2013) They used mapping of retinal locations to the retinal nerve fiber layer and optic nerve head compared to the VF test grid. Although the measurement of structures is still imperfect (amongst other things under the influence of glaucoma stage(Bowd et al. 2017) and refraction(Qiu et al. 2015)), a combined result could give a more specific and sensitive result and might be the missing link

in the structure-function discrepancy.(Yohannan & Boland 2017) The use of previous performed VFs of each patient is another example of incorporation of intrapatient data (see **Chapter II**). The implementation of these improvements require the development of software that is able to handle longitudinal data not only during analysis, but also during data acquisition.

A third possibility to optimize VF testing is a relatively new topic within ophthalmology, namely machine learning (ML): Automated algorithms that extract complex information of different modalities. A recent comprehensive review published by Li et al. describes ML in different ophthalmic fields like diabetic retinopathy and age related macular degeneration.(Li et al. 2020) Also in glaucoma, ML shows promising results in both VF(Yousefi et al. 2020) and OCT data.(Christopher et al. 2018, Ran et al. 2019) However, the degree of variability in VF testing and structural nerve fiber layer needs to be lower before implementation of ML in glaucomatous progression analysis is possible. The establishment of these improvements will lead to higher sensitivity and specificity to detect glaucoma progression in the future.(Thompson et al. 2020)

In conclusion, a modern style glaucoma care needs to be developed, and perimetry has a part in that future. A multimodal strategy should be developed whereby perimetry should evolve and combined with objective measurements. Using these different modalities together with the (existing) modern computer technologies in one device solution every clinician is waiting for.

REFERENCES

Bowd C, Zangwill LM, Weinreb RN, Medeiros FA & Belghith A (2017): Estimating optical coherence tomography structural measurement floors to improve detection of progression in advanced glaucoma. Am J Ophthalmol. 175: 37–44.

Christopher M, Belghith A, Weinreb RN, Bowd C, Goldbaum MH, Saunders LJ, Medeiros FA & Zangwill LM (2018): Retinal nerve fiber layer features identified by unsupervised machine learning on optical coherence tomography scans predict glaucoma progression. Invest Ophthalmol Vis Sci. 59: 2748–2756.

De Moraes CG, Demirel S, Gardiner SK, Liebmann JM, Cioffie GA, Ritch R, Gordon MO, Kass MA & Ocular Hypertension Treatment Group. (2012): Effect of treatment on the rate of visual field change in the Ocular Hypertension Treatment Study Observation Group. Invest Ophthalmol Vis Sci. 53: 1704–1709.

Gardiner SK (2018): Differences in the relation between perimetric sensitivity and variability between locations across the visual field. Invest Ophthalmol Vis Sci. 59: 3667–3674.

Heijl A, Leske MC, Bengtsson B, Hyman L, Bengtsson B, Hussein M & Early Manifest Glaucoma Trial Group (2002): Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. Arch Ophthalmol. 120: 1268–1279.

Heijl A, Patella VM, Chong LX, Iwase A, Leung CK, Tuulonen A, Lee GC, Callan T & Bengtsson B. (2019): A new SITA perimetric threshold testing algorithm: Construction and a multicenter clinical study. Am J Ophthalmol. 198: 154–165.

Jansonius NM (2010): On the accuracy of measuring rates of visual field change in glaucoma. Br J Ophthalmol. 94: 1404–1405.

Lamparter J, Russell RA, Zhu H, Asaoka R, Yamashita T, Ho T & Garway-Heath DF (2013): The influence of intersubject variability in ocular anatomical variables on the mapping of retinal locations to the retinal nerve fiber layer and optic nerve head. Invest Ophthalmol Vis Sci. 54: 6074–6082.

Li J-PO, Liu H, Ting DSJ, Jeon S, Chan RVP, Kim JE, Sim DA, Thomas PBM, Lin H, Chen Y, Sakomoto T, Loewenstein A, Lam DSC, Pasquale LR, Wong TY, Lam LA & Ting DSW. (2020): Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective. Prog Retin Eye Res. 6; 100900.

Qiu K, Schiefer J, Nevalainen J, Schiefer U & Jansonius NM (2015): Influence of the retinal blood vessel topography on the variability of the retinal nerve fiber bundle trajectories in the human retina. Invest Ophthalmol Vis Sci. 56: 6320–6325.

Ran AR, Cheung CY, Wang X, Chen H, Luo L-Y, Chan PP, Wong MOM, Chang RT, Mannil SS, Young AL, Yung H-W, Pang CP, Heng P-A & Tham CC. (2019): Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: a retrospective training and validation deep-learning analysis. Lancet Digit Health. 1**:** e172–e182.

Saunders LJ, Russell RA & Crabb DP (2015): Measurement precision in a series of visual fields acquired by the standard and fast versions of the Swedish interactive thresholding algorithm: analysis of large-scale data from clinics. JAMA Ophthalmol. 133**:** 74–80.

Swanson WH, Sun H, Lee BB & Cao D (2011): Responses of primate retinal ganglion cells to perimetric stimuli. Invest Ophthalmol Vis Sci. 52: 764–771.

Thompson AC, Jammal AA & Medeiros FA (2020): A review of deep learning for screening, diagnosis, and detection of glaucoma progression. Transl Vis Sci Technol. 9: 42.

Turpin A, Morgan WH & McKendrick AM (2018): Improving spatial resolution and test times of visual field testing using ARREST. Transl Vis Sci Technol. 7: 35.

Wall M, Kutzko KE & Chauhan BC (1997): Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. Invest Ophthalmol Vis Sci. 38: 426–435.

Wesselink C, Marcus MW & Jansonius NM (2012): Risk factors for visual field progression in the groningen longitudinal glaucoma study: a comparison of different statistical approaches. J Glaucoma. 21: 579–585.

Yohannan J & Boland MV (2017): The evolving role of the relationship between optic nerve structure and function in glaucoma. Ophthalmology. 124: S66–S70.

Yousefi S, Elze T, Pasquale LR, Saeedi O, Wang M, Shen LQ, Wellik SR, De Moraes CG, Myers JS & Boland MV. (2020): Monitoring glaucomatous functional loss using an artificial intelligence-enabled dashboard. Ophthalmology 127: 1170–1178.

Yousefi S, Goldbaum MH, Varnousfaderani ES, Belghith A, Jung T-P, Medeiros FA, Zangwill LM, Weinreb RN, Liebmann JM, Girkin CA & Bowd C. (2015): Detecting glaucomatous change in visual fields: Analysis with an optimization framework. J Biomed Inform. 58: 96–103.

NEDERLANDSE SAMENVATTING

Glaucoom is een oogaandoening waarbij de oogzenuw is aangedaan met irreversibel gezichtsveldverlies tot gevolg. Met verschillende behandelingen die zijn gericht op het verlagen van de oogdruk is het mogelijk om de progressie van gezichtsveldverlies te voorkomen, danwel te vertragen. Om te beoordelen of glaucoom stabiel is (oftewel de behandeling zorgt voor een voldoende verlaging van de oogdruk) of progressief (oftewel onderbehandeling) heeft het gezichtsveldonderzoek een essentiële rol. Ondanks dat dit onderzoek de gouden standaard is in de glaucoom zorg, is test-hertest variabiliteit een belangrijke tekortkoming waar in dit proefschrift aandacht aan wordt besteed.

Het belangrijkste doel van dit proefschrift is het verbeteren van glaucoomzorg door het optimaliseren van informatie die verkregen wordt uit gezichtsveldonderzoek. In het eerste deel van dit proefschrift heb ik onderzocht hoe gezichtsveldonderzoek werkt bij gezonde proefpersonen en patiënten met glaucoom. In **hoofdstukken II, III en IV** heb ik de verschillende intrinsieke en extrinsieke factoren beschreven die gezichtsveldonderzoek beïnvloeden en variabiliteit veroorzaken, zoals moment van meten (moment van de dag of jaar), degene die het gezichtsveldonderzoek begeleid en staaroperaties. In **hoofdstuk V,** heb ik deze kennis toegepast om de nadelen van glaucoom chirurgie (verlies van visueel functioneren door deze ingrepen) te vergelijken met het langetermijnvoordeel (vertragen van ziekte progressie). Ik ben teruggegaan naar de basis van visueel onderzoek in **hoofdstuk VI** en concludeerde dat een alternatieve psychofysische methode om de retinale sensitiviteit te meten niet beter werkte dan de huidige standaard, ondanks een kortere testtijd. Tot slot, in **hoofdstuk VII** heb ik laterale inhibitie onderzocht, een proces in de retina om contrast tussen beelden te vergroten. Een verlaging van laterale inhibitie zou een antwoord kunnen zijn op de discrepantie tussen een relatief normaal gezichtsveld onderzoek bij patiënten een glaucomateuze oogzenuw. Laterale inhibitie leek echter niet aangedaan te zijn bij patiënten met ernstig glaucoom met een intact centraal gezichtsveld.

RESUMEN EN ESPAÑOL

El glaucoma es una afección ocular en la que el nervio óptico se ve afectado, lo que resulta en una pérdida irreversible del campo visual. Con varios tratamientos destinados a reducir la presión ocular, es posible prevenir o ralentizar la progresión de la pérdida del campo visual. El examen del campo visual juega un papel esencial en la evaluación de si el glaucoma es estable (el tratamiento proporciona una reducción suficiente de la presión intraocular) o progresivo (es decir, tratamiento insuficiente). Aunque este estudio es el estándar de oro en la atención del glaucoma, la variabilidad test-retest es una deficiencia importante que se aborda en esta tesis.

El principal objetivo de esta tesis es mejorar la atención del glaucoma optimizando la información obtenida de los estudios del campo visual. En la primera parte de esta tesis, investigué cómo funcionan las pruebas de campo visual en sujetos sanos y pacientes con glaucoma. En los capítulos II, III y IV describí los diferentes factores intrínsecos y extrínsecos que influyen en el examen del campo visual y causan variabilidad, como el tiempo de medición (hora del día o año), quién supervisa el examen del campo visual y la cirugía de cataratas. En el capítulo V, apliqué este conocimiento para comparar las desventajas de la cirugía de glaucoma (pérdida de la función visual debido a estos procedimientos) con el beneficio a largo plazo (desaceleración de la progresión de la enfermedad). Volví a los conceptos básicos del examen visual en el Capítulo VI y concluí que un método psicofísico alternativo para medir la sensibilidad de la retina no funcionaba mejor que el estándar actual, a pesar de un tiempo de prueba más corto. Finalmente, en el capítulo VII investigué la inhibición lateral, un proceso en la retina para mejorar el contraste entre imágenes. Una reducción de la inhibición lateral podría ser una respuesta a la discrepancia entre un examen del campo visual relativamente normal en pacientes con un nervio óptico glaucomatoso. Sin embargo, la inhibición lateral no pareció verse afectada en pacientes con glaucoma severo con campo visual central intacto.

LIST OF ABBREVIATIONS

AGIS	Advanced Glaucoma Intervention Study
BCVA	best-corrected visual acuity
BGI	Baerveldt glaucoma implant
cd/m2	candela per square metre
CI	confidence interval
CIs	confidence intervals
CLIP	Continuous light increment perimetry
CoR	coefficient of repeatability
cpd	cycles per degree
CS	contrast sensitivity
D	diopters
dB	decibels
ERF	error related factor
FDT	Frequency doubling perimeter
FL	fixation loss
FN	false-negative
FP	false-positive
FT	Full Threshold
GATE	German Adaptive Threshold Estimation
GDx	laser polarimetry
GLGS	Groningen Longitudinal Glaucoma Study
GPA	Glaucoma Progression Analysis
HFA	Humphrey Field Analyzer
IOL	intraocular lens
IOP	intraocular pressure

IQR	interquartile range
MD	mean deviation
MFIOL	multifocal intraocular lens
ML	machine learning
MS	mean sensitivity
ms	millisecond
NPA	Non-parametric Progression Analysis
PACG	primary angle closure glaucoma
POAG	primary open angle glaucoma
PSD	pattern standard deviation
RCT	randomized clinical trial
RGC	retinal ganglion cell
RONI	Regions of no interest
ROP	rate of progression
SACG	secondary angle closure glaucoma
SAP	standard automated perimetry
SD	standard deviation
SE	standard error
SITA	Swedish Interactive Threshold Algorithms
SLT	selective laser trabeculoplasty
SP	staircase procedure
ST	suprathreshold
TE	trabeculectomy
UMCG	University Medical Center Groningen
VF	visual field

CURRICULUM VITAE

Francisco Gerbert Junoy Montolio was born in Emmen, the Netherlands in 1985. He graduated in 2003 from secondary school and started straight away at medical school. First, for a very short period in Amsterdam (VU), and thereafter in Groningen at the University of Groningen. During this study, he had different workrelated experiences like Noorder Dierenpark Emmen and working in a home care organization performing night shifts. By coincidence, he performed a mandatory research project at the department of Ophthalmology in the second year of medical school in 2005. In retrospect, the seed called Ophthalmology was planted then, sprouting several years later in 2008. In that year, the penultimate year of the Master, he performed the internship Ophthalmology in the 'Tjongerschans' hospital in Heerenveen, from where he decided to finalize his Master's degree in medicine at the department of Ophthalmology in the University Medical Center Groningen (UMCG). The academic internship starting in that year was the foundation for this PhD. During this PhD-trajectory, Francisco accomplished his specialization in Ophthalmology (2017) and fellowship vitreo-retinal surgery (2019) under supervision of drs. Huiskamp, drs. Postma and drs. Renardel de Lavalette. He is currently working as a staff member at the UMCG.

LIST OF PUBLICATIONS

Influence of glaucoma surgery on visual function: a clinical cohort study and meta-analysis. **FG Junoy Montolio**, RPHM Müskens, NM Jansonius. Acta Ophthalmologica. 2018 oct 4; https://doi.org/10.111/aos.13920

Lateral Inhibition in the Human Visual System in Patients with Glaucoma and Healthy Subjects: A Case-Control Study. **FG Junoy Montolio**, W Meems, MSA Janssens, L Stam, NM Jansonius. PLoS ONE. 2016;11(3): e0151006

Influence of multifocal intraocular lenses on standard automated perimetry test results. N Aychoua, **FG Junoy Montolio**, NM Jansonius. JAMA Ophthalmol 2013 Apr 1;131(4):481-85.

Persistence, spatial distribution and implications for progression detection of blind parts of the visual field in glaucoma: a clinical cohort study. **FG Junoy Montolio**, C Wesselink, NM Jansonius. PLoS ONE. 2012;7:e41211.

Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day, and season. **FG Junoy Montolio**, C Wesselink, M Gordijn, NM Jansonius. Invest Ophthalmol Vis Sci.2012; 53:7010– 7017.

Myopia as a risk factor for open-angle glaucoma: a systematic review and meta-analysis. MW Marcus, MM de Vries, **FG Junoy Montolio**, NM Jansonius. Ophthalmology, 118 (2011), pp. 1989–1994 e2.

PRESENTATIONS

2019	Noordelijke Oogheelkundige Nascholing (N.O.N., Paterswolde) *oral*
2016	Nederland Oogheelkundig Genootschap (NOG, Maastricht) *oral*
2014	European Association for Vision and Eye Research (EVER, Nice, France) (Granted with the paper abstract prize 2014) *oral*
2013	Nederland Oogheelkundig Genootschap (NOG, Groningen) *oral*
2012	Nederland Oogheelkundig Genootschap (NOG, Groningen) *oral*
2012	ARVO (Fort Lauderdale, Florida) *poster*
2011	ARVO (Fort Lauderdale, Florida) *poster*
2011	Nederland Oogheelkundig Genootschap (NOG, Maastricht) *oral*
2011	ARVO-NED (Rotterdam) *oral*
2010	ARVO-NED (Groningen) *oral*

DANKWOORD

Na het starten van de studie geneeskunde in 2003 met het idee, 'misschien is dat wel iets voor mij', en er tijdens de studie achter komen dat wetenschappelijk onderzoek en het werken in een universitair centrum echt niet iets voor mij is, is het afronden van dit proefschrift een bijzondere moment. Wellicht komt het door een gebrek aan zelfinzicht, maar waarschijnlijk door de collega's van de oogheelkunde die mij veel werkplezier geven dat ik altijd ben gebleven. Daarnaast ook alle mensen buiten oogheelkunde die ervoor zorgen dat er een goede balans is tussen werk en privé.

Professor Jansonius, beste Nomdo. Als de dag van gisteren kan ik me herinneren dat ik de supervisie kamer inliep met de vraag of ik mijn wetenschappelijk stage bij je kon uitvoeren, niet beseffende dat dit proefschrift uiteindelijk het gevolg zou zijn. Je advies dat glaucoom onderzoek tijd nodig heeft om betrouwbare resultaten te geven heb ik misschien iets te letterlijk genomen. Ik ben je dankbaar voor je onuitputtelijke energie en wijsheid en de mogelijkheden die je me hebt gegeven om te doen wat ik leuk vind. Tot slot vind ik het ontzettend knap hoe je als een grootmeester op zoveel borden tegelijk kan schaken.

Professor Hooymans, beste Anneke. Als afdelingshoofd ten tijde van mijn beginperiode bij de oogheelkunde ben ik je dankbaar voor de kansen die ik heb gekregen om op je afdeling te werken, te groeien en uiteindelijk zowel mijn wetenschappelijke als oogheelkundige carrière te beginnen.

Beste Christiaan. Bijzonder hoe onze relatie is gegroeid vanaf het moment dat deze begon als mijn stagebegeleider naar het begeleiden van je promotie door je paranimf te zijn, wat ik een grote eer vond. In je rol als stagebegeleider, onderzoeker, AIOS, oogarts, collega, maar ook op persoonlijk vlak ben je een voorbeeld voor me geweest.

De leden van de beoordelingscommissie, prof. Dr. P. van Dijk, prof. Dr. D. Crabb en prof. Dr. H. Beckers. Ik ben u zeer erkentelijk voor het beoordelen van mijn proefschrift.

Collega's van het Laboratorium voor Experimentele Oogheelkunde. Vooral de eerste twee en een half jaar en daarna in minder intensieve vorm hebben we goed samengewerkt, leuke discussies gehad en gezellige momenten daaromheen gehad, om de leerzame congressen naar bijvoorbeeld Florida niet te vergeten.

Beste AIOS, zonder jullie was het proefschrift waarschijnlijk al eerder klaar geweest. Het was een ontzettend fijne periode waar ik met veel plezier aan terugdenk. Zowel op het werk, samenwerken op de poli, samen studeren en congressen bezoeken, als

daarbuiten maakte mijn AIOS-tijd tot een zeer plezierige. Het is jammer dat die tijd voorbij is… Nogmaals dank: Eylem, The Anh, Danna, Angela, Marieke, Christiaan, Mathijs, Stijn, Sophie, Rosemarie, Nathalie, Ward, Jelle, Ryanne, Margriet, Nienke, Alie, Susanne, Britt, Victor, Martin, Nancy, Sharon, Marianne, Lisette, Lisanne, Doety, Inge, Lotus en Esra.

MMA, TOA's, optometristen, verpleegkundigen en alle andere medewerkers van de oogheelkunde. Dank voor jullie hulp bij het opvragen dossiers of gezichtsvelden en al het andere (onzichtbare) werk wat jullie dagelijks verzetten.

Vrienden van de dierentuin, wat is er veel veranderd sinds die tijd. We hebben niet zoveel contact meer als dat we ooit hadden, maar als we elkaar nodig hebben vinden we elkaar blindelings.

Ploeggenoten van A.S. Exstudiantes, de bal is wat ons verbindt vanaf het moment dat ik afstudeerde. Niets beter om het weekend te starten na een week onderzoek met een potje zaalvoetbal gevolgd door de bitterballenborrel. Gezond voor zowel lichaam als geest. De sportieve ambities spelen inmiddels een minder belangrijke rol, maar we vinden elkaar alsnog samen met onze gezinnen en/of bij een BB-avond.

Mam, je opvoeding heeft gemaakt tot wie ik nu ben, met een onbezorgde jeugd die ervoor zorgde dat ik mijn eigen keuzes kon maken. De onbezorgde tijd is helaas voorbij waardoor dingen voor altijd anders zullen zijn, maar je hebt laten zien hoe sterk je bent. Paps, plotseling was je er niet meer. Zoals Gerbert en mam op elkaar lijken, lijken wij op elkaar ook al blijf ik in de schaduw staan van je kennis en kunde. Ik mis je. Gerbert, broertje. Ik ben ontzettend trots hoe je als ondernemer de wereld (lees Emmen) aan het veroveren bent en in moeilijke tijden hebt staande gehouden.

Familia, agradezco su apoyo incondicional y tu cariño siempre me hace sentir muy feliz en Barcelona. Marjon, Bert en Job, bedankt dat ik mijn carrière bij jullie mocht beginnen. We hebben een kleine en waardevolle familie, waar we zuinig op zijn. Hans en Ineke en rest van de schoonfamilie, bedankt voor jullie warmte en steun in allerlei zaken.

Tot slot mijn paranimfen. Beste Tim, een aantal van onze gemeenschappelijke interesses zorgden ervoor dat het opstarten in de ochtend altijd wat tijd kostte, maar zorgde ook voor een fijne afleiding. Het lukte ons ook altijd goed om Margriet van haar

werk af te houden. Tot slot, waardeer ik het zeer dat je af en toe een bal van mij doorlaat als keeper. Al met al verschillende goede redenen dat je de aangewezen persoon bent om mij bij te staan op deze dag.

Djoeke, wat ontzettend fijn dat je tijdens de verdediging aan mijn zijde wilt staan, zoals ik je altijd graag aan mijn zijde heb, om me moed in te praten. Wie had kunnen bedenken dat we zouden staan waar we nu staan toen we elkaar voor het eerst ontmoetten op ons vakantiebaantje. Zonder jouw kracht en onvoorwaardelijke liefde zou ik dit proefschrift, opleidingen en alle andere uitdagingen nooit hebben kunnen aangaan en afronden. Ik kijk uit naar de toekomst die we verder met elkaar gaan delen, samen met onze twee prachtige dochters, Amarins en Lauri, die de wereld nog mooier maken. Ik hou van jullie!