Validation of the new fast perimetric strategy GATE (German Adaptive Thresholding Estimation) for static perimetry

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Luithardt, geb. Schmid, Annette Franziska

2016

Dekan:	Professor Dr. I. B. Autenrieth
1. Berichterstatter:	Professor Dr. U. Schiefer
2. Berichterstatter:	Professor Dr. P. Martus

Ich widme diese Arbeit meinem Ehemann und meiner Familie.

Content

Abbreviations

30A	name of a polar test point arrangement within 30° of eccentricity
84NO	name of a polar test point arrangement within 84° of eccentricity
AIDS	acquired immune deficiency syndrome
AION	anterior ischemic optic neuropathy
cLoA	corrected limits of agreement
CMOS	complementary metal oxide silicon
CRP	C-reactive protein
dB	decibel
dBs	standardized decibel scale
DLS	differential luminance sensitivity
dpt	dioptre
ESR	erythrocyte sedimentation rate
EyeSuite	EyeSuite Perimetry, a perimetry software offered by the Haag-Streit company, Köniz, Switzerland
FT	full threshold
GATE	German adaptive thresholding estimation
GATEe	commercially available GATE strategy as integrated in EyeSuite Perimetry
GATEe1075	examination on serial device 1075, max. stimulus ≈ 3183 cd/m² (10,000 asb), using GATEe
GATEe894	examination on serial device 894, max. stimulus ≈ 1273 cd/m² (4000 asb), using GATEe
GATE-i	German adaptive thresholding estimation – initial examination
GATEp	the prototype version of the GATE strategy as developed in Tuebingen
HIV	human immunodeficiency virus
ICC	intraclass correlation coefficient
IOP	intraocular pressure
LOA	limits of agreement
MD	mean defect

MS	mean sensitivity
NAION	non-arteritic anterior ischemic optic neuropathy
NQA	number of questions asked
RAPD	relative afferent pupil defect
RNFL	retinal nerve fibre layer
RP	retinitis pigmentosa
SITA	Swedish interactive thresholding algorithm
TL	test location
VF	visual field

1. Introduction

1.1 Preliminary remarks on perimetry

1.1.1 The visual field

The monocular visual field refers to the sum total of visual perception for an eye fixed on a stationary object of regard with the head and body held fixed in position. [1]
Under photopic conditions, the site with the highest concentration of cones within the retina is called the fovea. This is the site with the highest sensitivity of the visual field. The peripheral visual field extends to approximately 100° on the temporal side, to approximately 70° on the inferior side and to approximately 50° on the superior and the nasal side. [1] Sensitivity decreases with the declining concentration of cones towards the periphery. Thus, in a 3-dimensional illustration a visual field can be depicted as an “island of vision” with a central peak - the fovea - that is sinking with locally varying steepness towards the periphery into the “sea of blindness”. [2]

Areas of lower sensitivity than expected – i.e. visual field defects - are called scotomas. These defects are referred to as relative or absolute scotomas depending on how profound the visual loss is. If there is no more perception, even if the stimulus is presented with maximum luminance or if the local differential luminance sensitivity (DLS) level is reduced by more than 20 dB compared to the age-related normal DLS level, this defect is called an *absolute* scotoma; like for example the blind spot. In case of a *relative* scotoma there can still be perception, but the stimulus has to be presented with a higher luminance level than the local age-adjusted normal value.

Visual fields are always dependent on a functioning of all elements of the visual pathway, starting even with the refractive media, followed by the retinal photoreceptors and ending with the neuronal elements of the visual cortex. By this, they are representing the afferent functions of the visual system. It is

therefore possible to draw conclusions from special patterns of visual field defects in order to predict where the topographic location of the damage to the visual pathway is supposed to be. In addition, the assessment of local DLS values allows for analysis of change (deterioration or improvement over time).

The current gold standard for the assessment of overall visual function in clinical practice and in research environment is visual acuity testing and perimetry.

1.1.2 Perimetry

Perimetry is the psychophysical examination and measurement of visual function at defined topographic locations of the visual field. It is a non-invasive examination providing topo-diagnostic and etiological pathogenic information about the visual field of the examined subject. Stimuli are presented in a dome-shaped projection area and have to be recognized and confirmed by the proband. Test points may vary in size, luminance and location, but the background luminance is kept constant throughout the examination in order to provide a constant state of adaption

Perimetry measures differential luminance sensitivities. The logarithmic measurement unit is decibel (dB). The DLS is defined as the threshold of perception of a test point in relation to the background luminance of the perimeter. This indicates a quantification of the contrast perception capability. A threshold is given if the probability of perception is 50% at a given location of the visual field according to the psychometric function.

The psychometric function – also known as the probability-of-seeing-curve - describes the probability of response (in %) dependent on the level of stimulus attenuation (in dB). The steeper the curve at the point of change from unequivocal perception to no perception, the smaller is the statistical variance for repeated measurements. This has immediate implications for test-retest

variability. In other cases the curve shows a shallower course with greater statistical variance representing greater test-retest variability in these places.

This explains one of the major disadvantages of perimetry: its often high test-retest variability. In addition, low reproducibility results in a reduced sensitivity for changes over time. [3] Furthermore, as in any psychophysical test procedure, perimetry bears the risk of fatigue effect; also a learning effect may be observed. [4]

In spite of this, perimetry has a great impact on the evaluation of special eye diseases and their treatment, especially in glaucoma diagnostics and surveillance.

Further indications for a perimetric examination are an impaired vision of unknown origin, a relative afferent pupil defect (RAPD), clarification and surveillance of suspected visual pathway lesions and the need of an expert opinion and formal certification of seeing capability. The main purpose of perimetry, however, is the detection and surveillance of scotomas. [5]

There are two basic principles of perimetry: kinetic and static perimetry.

1.1.2.1 Kinetic isopter perimetry

In kinetic perimetry test points that are constant in size and intensity are introduced by movement. By this, the borders between fields of normal vision and scotomas are examined. Since it is a very interactive method, results are highly dependent on the examiner, but offer a high efficiency and flexibility.
In order to achieve reproducible results, it is helpful to follow some rules for all examinations: Movement should be introduced from the non-seeing into the seeing part of the visual field with a constant angular velocity (approximately 2-5°/s). Three to four passes, each with different stimulus brightness and/or stimulus size should be executed. [6] By that, the examination delivers lines displaying the same DLS, the so-called isopters, which are similar to contour

lines of topographic maps. Manual or semi-automated examinations may be executed. The latter offers a higher degree of independency from the examiner at the cost of reduced flexibility.
Kinetic perimetry is a good tool for the characterization of advanced visual field loss and is of high relevance for the evaluation of limitations of vision in patients' everyday life and for expert opinion. [5]

1.1.2.2 Static automated grid perimetry (used in this study)

In static automated grid perimetry stationary stimuli are presented in random order at varying intensity levels at different test locations that are arranged in a defined grid pattern. The depth of a defect can be evaluated by comparing the achieved DLS to the age-corrected normative values of the hill of vision.

Automated static visual field examinations have some major advantages compared to manual or semi-automated static perimetry. The most obvious advantage is the automated process of stimulus presentation allowing for randomization and increasing the reproducibility of examination. Different grids and strategies may be applied depending on the patient's ocular pathology and capability to perform the perimetric examination. In more recent perimeters infrared cameras allow the supervision of fixation, position and even vigilance of the test subject throughout the whole examination. In spite of all these great advantages, perimetry has stayed a long and tiring task with excessive demands of the patient making it a rather unpopular examination for patients [7]. Results are highly dependent on vigilance, cooperation, motivation and understanding of the patient and even though the examination itself is automated, technician experience still has a significant influence on the mean defect (MD). [8–10] With shorter test durations the strain of performing perimetry should be eased with helpful effects on fatigue, vigilance, cooperation and motivation. Static automated grid perimetry is a very important tool for glaucoma diagnostics and trend surveillance. [11]

1.1.2.3 Test point arrangements (grids)

By using automatic perimeters with a mirror-projection system for stimulus projection, almost every location of the visual field can potentially be tested technically. But since it is not possible to test every perceptive point of the retina for practical reasons, test points and their arrangement have to be chosen wisely depending on the goal of the perimetric examination. Rectangular grids with an equidistant test point arrangement offer the advantage to detect defects within mathematically well-defined Cartesian coordinates. [11] These grids are easy to manage and used widely. However, they do not realistically represent the arrangement of photoreceptors, which are arranged in a circular order around the center and increase in density towards the foveola.
In this study a concentric test point arrangement that respects the horizontal and vertical meridian was applied. By that, the actual arrangement of cones in the retina is represented more adequately. This is especially important, because conventional perimetry examines under photopic conditions (10 cd/m²), thereby exclusively addressing the cone system. Two different concentric grids were used in this study, one restricted to the central 30° (grid 30A) and one covering the whole almost 90° of the visual field (grid 84NO). Concentric grids often implement an increase of stimulus density towards the visual field center. [5]

1.1.2.4 Strategies for automated static perimetry

Suprathreshold tests

Suprathreshold tests offer an efficient and easy-to-perform evaluation of visual field status. They are based on the principle that the initial stimulus intensities are set a little above the expected threshold level. Test points are classified into three defect levels: normal, absolute and relative defect. However, the local DLS values are not assessed quantitatively. Therefore, suprathreshold tests are a good tool for screening examinations, for comparatively high spatial resolution and are useful especially for subjects that are inexperienced or incapable of performing other perimetric strategies. However, due to the lack of quantitative

local DLS values they are of minor importance for perimetric follow-up of chronic diseases (e.g. glaucoma).

Threshold tests

Threshold tests estimate the sensitivity at each test location. For good interpretability they need an adequate spatial resolution of the test point arrangements. Threshold tests allow follow-up testing to detect change in an early stage. So, they are the preferred strategy for surveillance of chronic eye diseases affecting the visual field. [9,12] The local DLS values are usually estimated on the basis of a staircase algorithm.

Full threshold (FT)

The gold standard of threshold testing is the Full Threshold strategy with a 4-2-1-dB staircase algorithm which needs three reversals of responses to terminate the examination at a given location.[13] Threshold tests have longer examination durations than screening tests, but they offer quantitative information. [14]

In order to shorten the examination duration of threshold testing several algorithms have been developed:

FASTPAC

This strategy saves about 40% of test time by accepting less accurate estimates of threshold. It uses 3-dB steps instead of 4-2-dB steps and stops testing already after one single response-reversal. This leads to higher short-term fluctuation, which makes it less accurate and reliable in surveillance and following of defects. [15,16]

TOP (Tendency-oriented Perimetry)

TOP is an ultra-short automated perimetry test, which only tests each stimulus once at each test location and calculates the threshold estimate by taking into account information from adjacent points. By this, it is up to four times faster

than a standard thresholding technique while producing similar results with a good diagnostic ability. TOP, however, tends to soften edges of sharp scotoma and makes them seem shallower and smaller. This may lead to an underestimation of visual field loss. [17–19]

SITA (Swedish Interactive Threshold Algorithm)

When introduced in 1998, [20] this method was extraordinary, because it was able to approximately halve test time by several different ways: Staircase starting values are calculated by visual field modeling, Bayesian posterior probability functions and frequency-of-seeing-curves taking into account surrounding test locations. The 4-2-dB staircase procedure is interrupted at a predetermined level of uncertainty. Furthermore, test pacing allowing adaptation of stimulus presentation time to the patient's reaction time and a method of calculating catch trials instead of testing them [21] lead to further test time reduction. Like for FT, test times increase with growing visual field defects. [22] There are two types of SITA: SITA Standard uses double crossing of threshold with a 4-2-dB-staircase, analogous to the Full Threshold method. SITA Fast, however, uses a 3-dB-staircase with single crossing like FASTPAC. When SITA was validated for normal and glaucomatous eyes [20,23], sensitivities were observed that were 1-2 dB higher than for conventional testing. [23,24] Disadvantages of SITA are the restriction to rectangular grids within 30° eccentricity and its deficient consideration of previous examinations. Furthermore, it has been released for manifest glaucomatous visual field defect only and not all details of the post-processing algorithm have been published.

The dynamic strategy by Weber

This strategy uses step sizes varying between 2 and 10 dB depending on the sensitivity according to physiological data. It showed to be more efficient than a strategy with fixed step sizes. [25] It is not as fast as TOP [26] and has shown a higher short-term fluctuation than the standard Octopus program [27].

CLIP (Continuous light increment perimetry)

By using a modified ramp stimulus with continuously increasing intensity this strategy enhances patient compliance (also in children [28]). The stimulus presentation starts with a subthreshold intensity and stops when recognized by the proband. CLIP saves approximately 38% of test time compared to FT. [29] However, local adaptation varies due to differing presentation durations for each test location.

In 2009 another new fast-thresholding algorithm was launched based on a modified 4-2-dB staircase strategy for automated static perimetry: GATE. This method does not only incorporate population information, but also benefits by including information from previous tests. [30]

1.2 GATE (German Adaptive Thresholding Estimation) strategy

For a faster completion of threshold-estimating static visual field examinations a new fast-thresholding algorithm was developed by U. Schiefer and J. Paetzold, university eye clinic Tuebingen, and called “German Adaptive Thresholding Estimation (GATE)”. When compared to the full threshold strategy (FT) and SITA Standard in a multicentre study it achieved comparable results. Accuracy and test-retest reliability have shown to be similar to both other strategies, but GATE showed a much shorter test duration than FT. [13] GATE is able to determine accurate thresholds over the entire sensitivity range and is applicable to all kinds of ophthalmologic pathologies and any test point arrangement.

The algorithm consists of “GATE-i” which is used for the initial examination of a patient and “GATE” which is used for all subsequent examinations. GATE-i begins with the testing of 5 predefined seed points. The achieved DLS values are then compared to the age-corrected normal hill of vision. If necessary, deviations of the seed locations from the normative values are taken to adapt the other starting stimulus intensities.
A modified 4-2-dB staircase strategy follows the testing of the seed points. Two reversals are needed in order to terminate the examination of a test point. A

local threshold is defined as the value between the brightest stimulus not seen and the dimmest stimulus seen. Precocious termination is possible in areas of deep or absolute defect: If the initial stimulus, that is slightly higher than the suspected DLS, is not answered within a given time window, a stimulus of maximum brightness is presented. If this stimulus is also not answered, the testing at that location is terminated.
The GATE algorithm differs only by the fact that the starting values are not based on adjusted age-related normative values, but on previously accessed local thresholds of precedent examinations. By that, GATE needs even shorter examination times than GATE-i.
Thus, GATE is an algorithm that examines visual fields in an adaptive method (adaption of initial stimulus intensities to the age-corrected normative values for GATE-i and adaption of initial stimulus intensities to precedent examinations for GATE) in order to perform threshold estimating perimetry.

1.3 Purpose of this study

Until now GATE has only been available in a prototype version for laboratory use (GATEp). In order to be able to introduce this fast-thresholding algorithm into clinical practice, the license rights have been sold to Haag-Streit AG, Köniz, Switzerland, that integrated the algorithm into their commercially available EyeSuite Perimetry software (GATEe).

The *primary objective* of this study was to assess the *agreement* between this incorporated (GATEe) and the original version (GATEp) of the algorithm regarding local differential luminance sensitivity (DLS). The results were to be related to and evaluated by the also assessed repeatability of GATEp. Furthermore, the examinations were also performed by a perimeter of the newest generation (serial device 1075) with *LED background illumination* and a *higher maximum stimulus luminance* using the GATEe algorithm. Possible effects of the different illumination on the measurement results were assessed. GATEe was therefore performed on two different perimeters: serial device 104 with a maximum stimulus luminance of 1273 cd/m² (4000 asb), i.e. GATEe104,

and serial device 1075 with a maximum stimulus luminance of 3183 cd/m² (10,000 asb), i.e. GATEe1075.
Secondary objectives were the comparison of *examination durations* and *mean sensitivity* (MS).

2. Subjects and methods

2.1 Subjects

Since examinations of the visual field are notably relevant for particular eye diseases concerning prognosis, therapy and monitoring of the disease, four groups of patients representing four different diseases affecting the visual field have been examined. These groups were as follows:

1) 15 patients suffering from manifest glaucoma
2) 3 patients with NAION (Non-Arteritic Anterior Ischemic Optic Neuropathy)
3) 6 patients with chiasmal or postchiasmal lesions of the visual pathway resulting in a homonymous hemianopia or quadrantanopia
4) 6 patients suffering from tapeto-retinal degeneration (Retinitis pigmentosa (RP))

In total, 30 patients were examined, 17 men and 13 women between 22 and 78 years of age (mean: 58.7 years). All patients underwent eight perimetric examinations. Therefore, 240 examinations were completed altogether.
13 right eyes and 17 left eyes were chosen as study eyes according to the inclusion criteria of the different groups of diseases that will be explained in the following.
See Table 1 for an overview of all recruited test subjects.

Table 1: Recruited test subjects

((post-) chiasmal = (post-) chiasmal pathway lesions)

RP = Retinitis pigmentosa, OS = left eye, OD = right eye)

Patient ID	Gender	Age [years]	Disease	Eye
2201	Male	73	NAION	OD
2202	Female	53	Glaucoma	OD
2203	Female	76	Glaucoma	OS
2204	Female	61	Glaucoma	OS
2205	Female	66	Glaucoma	OS
2206	Female	71	Glaucoma	OS
2207	Male	62	Glaucoma	OS
2208	Male	65	Glaucoma	OS
2209	Male	65	Glaucoma	OD
2210	Female	78	NAION	OS
2211	Male	53	Glaucoma	OS
2212	Female	75	Glaucoma	OS
2213	Male	41	(post-) chiasmal	OD
2214	Male	76	NAION	OS
2215	Male	66	Glaucoma	OD
2216	Male	72	(post-) chiasmal	OD
2217	Male	56	Glaucoma	OD
2218	Female	64	(post-) chiasmal	OS
2219	Male	22	RP	OS
2220	Male	55	(post-) chiasmal	OD
2221	Female	38	(post-) chiasmal	OS
2222	Male	65	RP	OD
2223	Male	59	Glaucoma	OD
2224	Female	62	Glaucoma	OS
2225	Male	61	RP	OS
2226	Male	44	RP	OS
2227	Female	67	RP	OD
2228	Female	29	RP	OD
2229	Male	60	Glaucoma	OD
2230	Female	27	(post-) chiasmal	OS

2.1.1 Groups of subjects

2.1.1.1 Glaucoma

Glaucoma is the umbrella term for a number of different ophthalmologic diseases of different aetiology resulting in an optic neuropathy with characteristic morphological changes of the optic disc and typical visual field defect patterns, often accompanied by a high intraocular pressure (IOP). There is no healing therapy for glaucoma, but if left untreated, it may result in blindness of the eye. [31] Diagnostic criteria include: structural damage of the optic nerve head resulting in increasing excavation (which may be detected by measuring the vertical disc-cup-ratio) and/or focal notching with accompanying local atrophy and functional damage, i.e. visual field defects resulting in characteristic nerve fibre bundle defects. [32]
Most of the perimetric examinations in outpatient care are applied to glaucoma patients. Besides other examinations like ophthalmoscopy, morphometry (e.g. optical coherence tomography = OCT) of the retina and tonometry, perimetry is important for glaucoma patients in order to diagnose the disease and for follow-up purposes. Perimetry can help to differentiate, if the patient suffers from a slowly progressive functional loss or a more aggressive form of the disease, which would have an impact on the aggressiveness of treatment. [33] In glaucoma patients threshold estimating static perimetry of the 30° central visual field is recommended. [34] The most important treatments of glaucoma are medical and surgical methods to reduce the IOP depending on the aetiology in order to prevent progression of visual field defects.
Patients were included into the study, if the optic nerve head and/or retinal nerve fibre layer (RNFL) and visual field were abnormal, according to the classification stages I-III (AULHORN classification [35]). The affected eye was chosen as study eye. If both eyes were affected, the worse eye was chosen.

2.1.1.2 AION

The anterior ischemic optic neuropathy (AION) is an acute ischemia of the papilla because of vessel transformation of either inflammatory (arteritic AION)

or arteriosclerotic reasons or because of a hypotonic situation (non-arteritic AION). Patients complain of sudden vision impairment or even blindness in one eye. Altitudinal visual field defects, especially of the lower half, are typical. The extent of the defects depends on the degree of destruction of the papilla and the optic nerve. Typical signs in ophthalmoscopy are optic disc oedema and segmentally blurred disc margin, sometimes combined with hyperaemia and disc haemorrhages followed by segmental atrophy of the optic nerve head. [31] Improvement or further deterioration of visual field and visual acuity may mostly be observed up to 6 months after the ischemic event. Thereafter further significant change is very rare. [36] If the reason for the AION is inflammatory (giant cell arteritis), it is of utmost importance to treat with high-dose steroids as soon as possible in order to protect the fellow eye and prevent occlusion of brain vessels. Important diagnostic signs for giant-cell arteritis are a high erythrocyte sedimentation rate (ESR) and elevated C-reactive protein (CRP). Also, the patients often suffer from pathognomonic jaw claudication. [31]
Follow-up of the visual field defects (under therapy) is the main purpose of perimetric examinations of AION patients. When performing static perimetry threshold-estimating strategies are useful, because of their ability to detect and quantify both local defects and diffuse reduction of sensitivity. [34]
For this group the affected eye was chosen as the study eye.

2.1.1.3 Chiasmal or postchiasmal lesions of the visual pathway

Chiasmal lesions of the visual pathway normally result in heteronymous bitemporal visual field defects with a great variability depending on the aetiology and location of the lesion. In many cases the underlying pathology is a pituitary tumour or craniopharyngioma. Further reasons are other tumours, aneurysms or inflammatory processes. Neurosurgical and medical therapies are applied. Postchiasmal lesions are due to numerous neurologic diseases like tumours, vascular insults, basal meningitis, trauma, abscesses or aneurysms. They all result in contralesional homonymous visual field defects. The most important diagnostic tool in this case regarding topographic information is perimetry, since typical visual field defects occur depending on the location of damage along the

visual pathway (optic tract, lateral geniculate nucleus, optic radiation, visual cortex). Therapy and prognosis depend on the location and aetiology of the damage. Neurosurgery or neurologic treatments are possible, but regression of postchiasmal damage in visual field defects is rare. [31]
Inclusion criteria for this group were homonymous or heteronymous hemianopia or quadrantanopia regardless of origin. Stroke was no exclusion criteria in this group. The study eye was chosen by randomization.

2.1.1.4 Tapeto-retinal degeneration: Retinitis pigmentosa

A heterogeneous group of retinal pathologies leading to nyctalopia and progressive loss of visual acuity and constriction of the visual field is called retinitis pigmentosa. In its classical form a concretion of retinal pigments is a typical symptom. This concretion proceeds from the mid-periphery towards the fovea. So, in the course of the disease the visual field narrows step by step. It usually starts with the destruction of the rods and later also the cones, which first leads to a disorder of colour and contrast vision and later even to optic atrophy. Since there is no curative or prophylactic therapy the disease is of a chronic progressive character and may lead to blindness.
For retinal diseases perimetry is a tool which helps to discern differential diagnoses and helps in trend surveillance. There are often small remaining islands of vision in the periphery apart from the typical small concentric visual field, even though the disease is already quite advanced. This is why grid 84NO has been applied for retinitis pigmentosa patients in this study. Patients were included, if suffering from retinitis pigmentosa of the classical form in different stages of the disease. The study eye was chosen by randomization.

2.1.2 Inclusion criteria

- ✓ Physical, intellectual and linguistic abilities in order to understand the test requirements
- ✓ Willingness to comply with the protocol of the 2 visits
- ✓ 18 years old, informed consent

For the study eye:

- ✓ Spherical ametropia max. ± 8 dpt, cylindrical ametropia max. ± 3 dpt
- ✓ Distant visual acuity better than 10/20
- ✓ Isocoria
- ✓ Pupil diameter > 3mm

2.1.3 Exclusion criteria

- ✓ Pregnancy, nursing
- ✓ Diabetic retinopathy
- ✓ Asthma
- ✓ HIV positive or AIDS
- ✓ History of epilepsy or significant psychiatric disease (e.g. dementia)
- ✓ History of stroke (except for patients in the group of visual pathway lesions)
- ✓ Medications known to affect the visual field sensitivity
- ✓ Acute ocular infections (e.g. keratitis, conjunctivitis, uveitis)
- ✓ Severely dry eyes
- ✓ Miotic drugs
- ✓ Amblyopia
- ✓ Squint
- ✓ Nystagmus
- ✓ Albinism
- ✓ Any ocular pathology in either eye that may interfere with the ability to obtain visual fields, disc imaging or accurate IOP readings
- ✓ Keratoconus
- ✓ Intraocular surgery (except for uncomplicated cataract or glaucoma surgery performed >3 months prior to screening)
- ✓ History or presence of macular disease and/or macular oedema
- ✓ Relevant opacities of central refractive media (cornea, lens, vitreous body)
- ✓ Ocular trauma
- ✓ Suspected lack of compliance

In this study, patients with false-negative or false-positive rates exceeding 30% were not excluded from the analysis, because of the assumption, that false-negative answers are often increased in patients with severe visual field loss and in order to draw conclusions from a very realistic sample of visual field examinations as would be normal for everyday ophthalmologic practice.
These patients all suffered from severe visual field loss, either because of advanced glaucomatous loss (patient IDs 2205, 2211, 2212, 2217, 2223) or from homonymous hemianopia (patient ID 2216).
See Table 16: Elevated catch trial rates (Appendix)

2.1.4 Recruiting of subjects

Recruiting of participants for the study was accomplished via two different ways: Firstly, patients from the outpatient clinics of the university eye hospital in Tuebingen, e.g. the outpatient glaucoma service, were screened individually by the investigator. Secondly, patients meeting all inclusion criteria were contacted via telephone call and asked to participate in the study. Inclusion and exclusion criteria were checked by the information given in the AIS (information system for doctors of the university eye hospital Tuebingen) and by the patients' history.

2.2 Study design

Two visits within 14 days have been scheduled for each of the chosen 30 test subjects. At their first visit the patients received all relevant information about the purpose and the exact procedure of the tests. They all signed an informed consent. Afterwards, four examinations using the GATE-i strategy were performed. In glaucoma patients the IOP was measured with a non-contact tonometer. The first visit took approximately 1.5 hours. At the second visit four examinations were performed using the GATE-strategy, which took approximately one hour.
The four static visual field examinations at each visit were performed with three different Octopus 900 devices:

GATEp1	prototype GATE software, serial device 104, first examination
GATEp2	prototype GATE software, serial device 104, second examination

GATEe894 EyeSuite GATE software, serial device 894

GATEe1075 EyeSuite GATE software, serial device 1075, a perimeter of the newest generation with LED background illumination and higher maximum stimulus luminance (3183 cd/m² = 10,000 asb)

The order of the four examinations was randomized in advance for each test subject and each visit. The patients were free to choose their breaks individually between the examinations, but at least one break had to be taken. All examinations were monitored continuously by the investigator.

Prior to the recruitment of participants and the first examination, the study had been reviewed by the independent Ethics Committee of the faculty of medicine, Tuebingen University, and was approved to comply with the tenets of the Declaration of Helsinki. It was also registered at ClinicalTrials.gov. The clinical trial number was NCT01265628.
This was a non-invasive study, no medication was tested and there was no known additional risk due to the diagnostic equipment. The devices were commonly used in diagnostic procedures. Regular safety tests were done by "Medizintechnisches Servicezentrum des Universitätsklinikums Tuebingen".
The participants received 32 € expense allowance per visit and were granted an accident en route insurance. They were free to stop participation in the tests at any time without any consequences. This validation study with 30 participants was commissioned by the Haag-Streit AG, Köniz, Switzerland.

2.2.1 Perimetric examination procedure

After the calibration of the perimeter the adequate software settings were chosen. A correction of spherical and cylindrical refraction was obtained, if necessary, by thin-rimmed glasses put into a movable holder before the start of the examination for examinations inside the central 30°. Adequate near correction was achieved by an age-dependent near addition in accordance with the following table (table 2) and subsequent fine tuning, so that the patient could see the fixation target in focus.

Table 2: Age-dependent spherical near addition for Octopus 900

Age [years]	Near correction [dpt]
40-44	+ 1.0
45-59	+ 1.5
50-54	+ 2.0
55-60	+ 2.5
>60	+ 3.0

Cylindrical correction was applied if cylindrical ametropia was greater or equal to 1 dpt. For examinations beyond the central 30° of the visual field (as it is for the peripheral part of the grid 84-NO) no correction lenses were applied. A white half-transparent eye patch was used to cover the not-examined eye. The patients' position was maintained with the face resting on the chin and forehead rest.

Adequate instructions to the test subject are mandatory in order to achieve reliable and reproducible results [8,10]. The examiner explained the course and purpose of the examination and how to use the patient-response button. A rather conservative policy of only pushing the button, if sure to have seen a point, has been chosen for this study. The same investigator explained and conducted all examinations of all patients with standardized instructions in order to maintain identical conditions for each examination [10,11]. Furthermore, fixation control and supervision of vigilance and reliability during the examination were crucial tasks of the examiner. If, for example, spontaneous waves of contraction and dilation of pupil size were observed which indicate increasing sleepiness of the patient [37], the investigator was responsible to alert the patient. Throughout the examination the position and the vigilance of the patient was monitored via infrared camera. Adjustment of the chin rest and glass holder was done throughout the whole examination in order to ensure an optimal positioning of the patient at any time.

Since poor fixation may lead to underestimation of depth and extent of visual field defects, fixation was controlled by the investigator via an infrared camera.

Poor fixation was addressed by the investigator's verbal feedback to the patient and by correction of the chin rest. Stable fixation was also supported by prior instructions: Patients were told that they will not see half of the stimuli due to the examination strategy, so that they should not give up on fixation if there was no perception of stimulus for several seconds. Stimulus presentation was repeated automatically, if the eye was found to be closed during initial presentation.

After the examination all results were saved as *.txt-files (GATEp) or *.pid-files (EyeSuite Perimetry) on the computers at the laboratory of the eye clinic Tuebingen and were also printed.

2.3 Technical data

2.3.1 Hardware (examination devices)

All examinations were executed using perimeters of the Octopus 900 series. The three perimeters differed mainly with respect to the maximum stimulus luminance, which was 318 cd/m² (1000 asb) for serial device 104 (GATEp), 1273 cd/m² (4000 asb) for serial device 894 (GATEe894) and 3183 cd/m² (10,000 asb) for serial device 1075 (GATEe1075), respectively.

The perimeters were connected via an Ethernet link to a computer or laptop that controlled the perimeters via the perimetry software and stored the obtained data.

The Octopus 900 perimeter is an automatic projection perimeter with a spherical, Goldmann type cupola design which allows testing of the entire visual field including the periphery (temporally). Kinetic as well as static or flicker perimetry may be accomplished both in the 30° and 90° range. The cupola has a radius of 300 mm (by this being in accordance with the Goldmann bowl). Test zones can be measured up to the following levels of eccentricity: nasal 60°, temporal 89°, superior 60° and inferior 70°. The outer dimensions of the

perimeter are: 648 mm (width) x 519 mm (length) x 796 mm (height). It weighs 25 kg.

Background luminance is based on two light sources of several LEDs and controlled by a separate light sensor. DLS up to 47 dB may be measured with a measurement accuracy of 0.5 dB. Various stimulus sizes (Goldmann I-V) can be presented within pre-specified stimulus intervals (adaptive – 4 sec) and for various stimulus durations (100ms, 200ms, 500ms) at different background luminance levels (1.27 cd/m² or 10 cd/m²). In this study, the background luminance was 10 cd/m^2 for all instruments. Green-lighted markers (diamond) were chosen as fixation targets, because they allow testing of the foveal differential luminance threshold.

The most recent generation of Octopus perimeters (serial device 1075) uses LEDs which slightly change the spectrum of background illumination and allow for higher maximum stimulus luminance levels, while the older Octopus perimeters (serial devices 104 and 894) use bulbs.

A permanent infrared videopupillography based fixation control is possible, because the examined eye is illuminated with infrared LEDs throughout the examination and recorded by a CMOS ("Complementary Metal Oxide Silicon") camera. The image of the eye is shown on the LCD (liquid crystal display) display. The investigator is able to monitor the vigilance of the patient and assure a precise positioning of the eye by a motorized fine-adjustment of the chin rest at any time throughout the examination.

Stimuli are projected onto the inner cupola surface via a mirror projection system. The stimulus intensity is controlled by a light sensor that is also a reference point for the system of coordinates for test locations. [38]

2.3.2 Software

Two different software versions of the GATE algorithm were applied:

1) The original prototype GATE software (GATEp) developed by U. Schiefer and J. Paetzold, university eye clinic Tuebingen, a fast-thresholding estimation software based on a 4-2-dB staircase principle.
2) The commercially available EyeSuite© Perimetry software with the incorporated GATE software (GATEe), Haag-Streit AG, Köniz, Switzerland.

The basic settings for the examinations were set as identical as possible for both software versions.

2.3.2.1 Settings that were the same for GATEp and GATEe

Kind of examination:	Static
Stimulus method:	Standard
Stimulus/background:	W/W (white on white)
Stimulus size:	Goldmann III
Presentation duration:	200 ms
Background luminance:	10cd/m² (31.4 asb)
Examination program:	30A, 84NO
Fixation control:	Off
Fixation target:	Cross markers (diamond)
Eye:	OD (right eye) – OS (left eye)

2.3.2.2 Settings that were different for GATEp and GATEe

Due to methodological reasons some settings were not identical for the two software versions (see table 3).
For GATEe the first examination with the chosen strategy “GATE” automatically applies the GATE-i algorithm and the subsequent examinations automatically use the GATE algorithm when selecting the button “same examination as last time”. For the first examination with GATEp, GATE-i and five anchor points

have to be chosen manually. For the second examination GATE is the chosen examination strategy automatically.
Due to a misleadingly labelled scale of the EyeSuite software two different interstimulus intervals were chosen inadvertently: GATEp: 1200ms, GATEe: 1500ms.
Also, the percentage of catch trials was different GATEp: 2%, GATEe: 10%.
Furthermore, only with the GATEp version five test locations were tested twice for evaluating the short-term-fluctuation.

Table 3: **Differing settings for GATEp and GATEe**

	GATEp	**GATEe**
GATE-i initial points	5	4 (automatic)
Response interval	1200 ms	1500 ms
Catch trials	2%	10%

2.3.3 Test grids

Two different concentric test point arrangements were used for the examinations. The stimuli were arranged according to a polar coordinate system straddling the vertical and horizontal median, in order to facilitate detection of visual field defects respecting the nasal step or the vertical midline. [5,39]

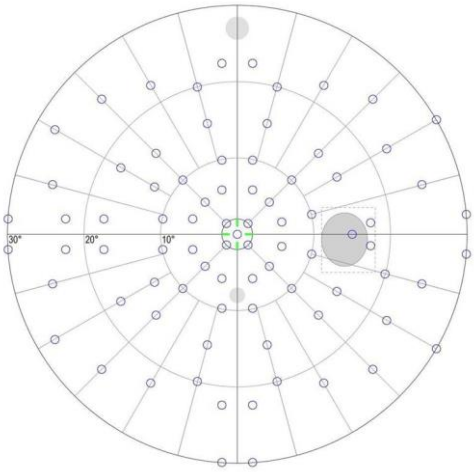

Figure 1: grid 30A

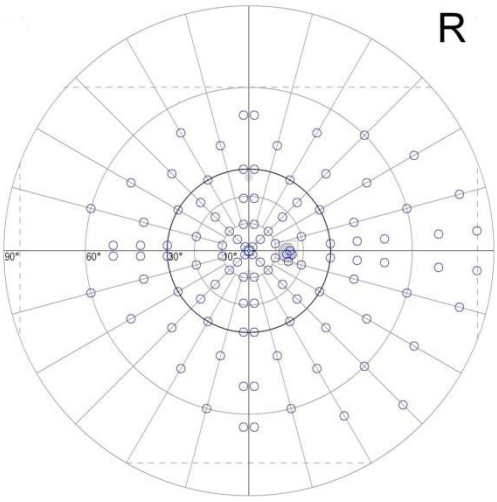

Figure 2: grid 84NO

Test grid 30A (see figure 2) covering the central 30° of the visual field by examining 83 test points was applied for glaucoma patients, AION patients and patients with chiasmal or postchiasmal visual pathway lesions, altogether 24 subjects. Test grid 84NO (see figure 3) examines 109 test points covering almost 90° of the visual field. When applying this grid, the central 30° of the visual field were tested first. Refraction correction glasses had to be removed before continuing with the examination of the peripheral parts of the visual field beyond 30° eccentricity. Grid 84NO was used for patients with retinitis pigmentosa, altogether 6 patients.

When comparing the test point arrangements of GATEp and GATEe several test points were identified that had to be excluded from the analysis for different reasons:

For grid **30A** eleven test points were excluded (see figure 3): The DLS results of five test locations (location IDs: 1, 26, 28, 30, 32) had to be excluded because they were tested twice by GATEp for the assessment of short-term-fluctuation. Their “twin” location IDs (0, 25, 27, 29, 31) were included and analysed. Two points (location IDs: 36 and 52, examining the blind spot) had to be excluded because they were not tested by GATEe at all and four points (location IDs: 80, 81, 86, 87, temporal rim points) had to be excluded because the coordinates of

these location IDs were not identical for the respective test grid versions of the two software versions. The remaining 77 points were analysed.

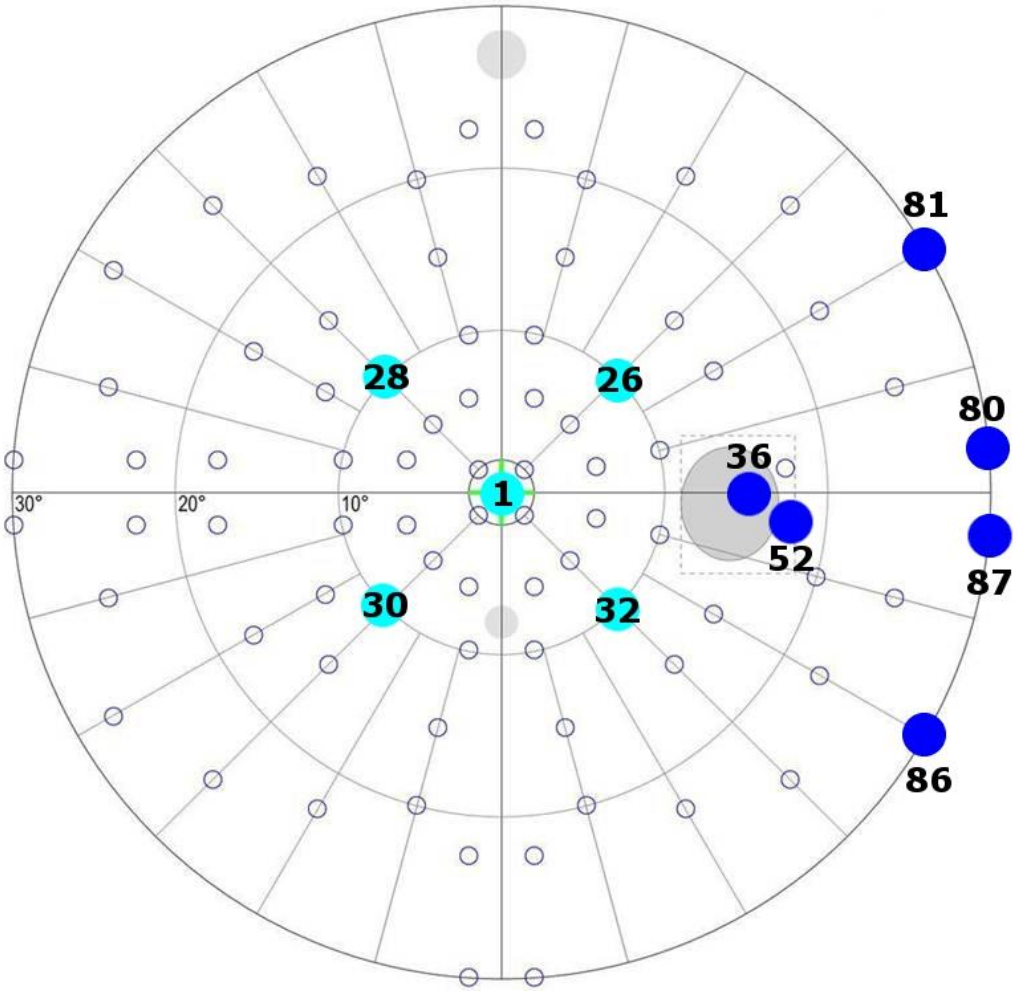

Figure 3: excluded test locations of grid 30A
IDs 1, 26, 28, 30, 32 apply to the second testing of these locations. The results of the first testing have been included into the analysis (ID 0, 25, 27, 19, 31). IDs 36, 52 were not tested by GATEe, IDs 80, 81, 86, 87 had different coordinates for GATEe

For grid **84NO** fourteen test points were excluded from the analysis (see figure 4): Three test points (location IDs: 21, 23, 24, examining the blind spot) were not tested by GATEe and eleven test points (location IDs: 41, 42, 43, 46, 47, 50, 51, 54, 55, located on the 30° rim, and location IDs 107, 108, examining the most peripheral temporal locations) had differing coordinates for the location IDs of GATEp and GATEe. Altogether, 95 points could be included.
See Appendix for tables of location IDs.

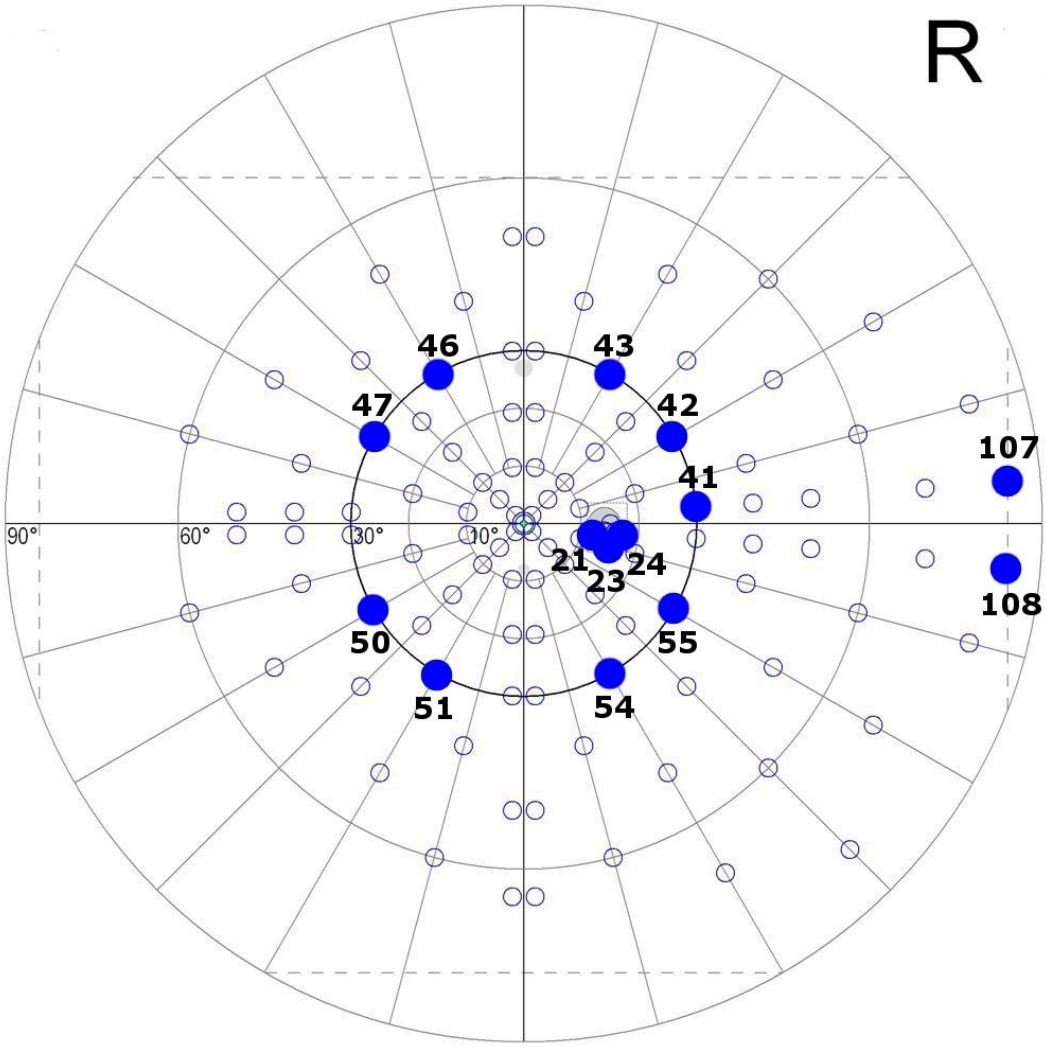

Figure 4: excluded test locations of grid 84NO
IDs 21, 23, 24 were not tested by GATEe, IDs 41, 42, 43, 46, 47, 50, 51, 54, 55, 107, 108 had different coordinates for GATEe

After the examinations and before analysis, all results gained for left eyes were mirrored with respect to the vertical meridian in order to conform to the location IDs of the right eye grids.

2.4 (Statistical) Analysis

The data entry, statistical calculations and design of tables and figures were performed using the statistical software JMP 9.0.0.
All 30 participants concluded both visits. Unfortunately, the EyeSuite Perimetry data could not be saved for two patients (patient IDs 2203 and 2204, both

glaucoma patients) because of problems with the driving laptop. The GATEe results could therefore not be included into the analysis for these two patients.

The statistical analysis was primarily heading for the following variables: local DLS values, test duration, MS values.

2.4.1 Analysis of DLS differences

DLS values were exported from GATEp and GATEe. When the location IDs were checked for conformance, eleven location IDs had to be excluded from analysis for grid 30A and fourteen for grid 84NO (see 2.3.3 Test grids).
The exported DLS values could not be compared immediately for two reasons: First, they had different units of measurement. DLS values were measured in [dBs] by GATEp and in [dB*10] by GATEe. This is why all GATEe results had to be divided by 10 first.
Second, their reference scales were different. For GATEp the reference value was the background luminance of 10 cd/m² (31.4 asb) with the measurement unit dBs (standardized dB scale). For GATEe, however, the scales were referenced to the maximum stimulus luminance, which was 1273 cd/m² (4000 asb) for serial device 894 and 3183 cd/m² (10,000 asb) for serial device 1075, respectively. This correlates with an offset of 4 dB between the two perimeters. Because of the logarithmic scale, a translation from one scale to the other is performed by subtraction or addition of a given value. [1]
Therefore, the DLS values measured by serial device 1075 were subtracted by 4 and the DLS values measured by GATEp were added to 22, in order to achieve comparable dB-scales.
For serial device 1075 all negative results were set to 0 dB, in order to factor out small differences in places of almost absolute defect.

When deciding whether a new method may be used instead of an already established method, there are two essential aspects: First, the amount of agreement between the two methods and second, the clinical evaluation of the differences.

2.4.1.1 Statistical agreement

For the assessment of the statistical agreement of the different examinations the approach by Bland and Altman [40] was applied.

In order to estimate the repeatability of static visual field examinations with the GATE software, replicate examinations were performed with GATEp. [41] The statistical agreement between GATEp1 and GATEp2 represents the retest-reliability of GATE. Secondly we assessed the agreement between GATEp1 and GATEe894 and between GATEp1and GATEe1075.

Bland-Altman plots were modified and drawn for the different comparisons as follows (separately for the different grids): Average DLS values of each test location of the examinations of all included patients were plotted against their differences. The bias was defined as the median difference (instead of the mean) of the DLS values of these examinations and was depicted by a horizontal line. Furthermore, the 2.5 and the 97.5 percentile of the DLS differences (instead of ± 1.96 standard deviation) were established in order to specify the so-called limits of agreement (LOA) and also depicted by horizontal lines. 95% of the differences between the measurements were therefore assumed to lie within these limits.
Like this, the statistical agreement between the methods was specified by the bias that represents a possible systematic error and the limits of agreement that represent the spread of differences between the measurements. [42]

2.4.1.2 Clinical evaluation criteria and literature criteria

This statistical agreement was categorised by clinical evaluation criteria (see table 4). Taking into account the final step size of 2 dB for the GATE algorithm [13,43], the measurement accuracy of 0.5 dB of the Octopus 900 perimeters and an assumed short term fluctuation of 1.5 dB (normative value for Octopus 101) [44], the following criteria have been defined in advance:
Very good agreement was stated for LOA ≤ 3 dB and a bias ≤ 0.5 dB, *good agreement* for LOA ≤ 4 dB and a bias ≤ 1 dB, *acceptable agreement* for LOA

≤ 5 dB and a bias ≤ 2 dB, respectively. LOA > 5 dB and a bias > 3 dB would indicate *insufficient agreement* and were rated as *not acceptable*.

Table 4: clinical evaluation criteria

Clinical evaluation	LOA	Bias
Very good agreement	≤ 3 dB	≤ 0.5 dB
Good agreement	≤ 4 dB	≤ 1 dB
Acceptable agreement	≤ 5 dB	≤ 2 dB
Not acceptable	> 5 dB	> 3 dB

Before the evaluation of the data the LOA were corrected by subtraction of the bias (cLOA, corrected limits of agreement), because the criteria mentioned above are based on the assumption of a bias of 0 dB.

Furthermore, since variance of measurements should not imitate or conceal a real progression of visual field defect, criteria were retrieved from recent literature that would indicate worsening or new detection of visual field defects. Differences exceeding these limits would be rated as not acceptable. According to the recent literature [33,45–48] the aberration between GATEp and GATEe should not exceed ± 5 dB in more than two test locations of the examined part of the visual field. There should not be more than two adjoining test points with an aberration greater than 5 dB. Edge points may be ignored.

2.4.2 Analysis of examination duration

Examination durations of GATE-i and GATE were assessed and compared for GATEp and GATEe. Since the median is more robust concerning outlier values than the mean, we took the median to describe the duration of examination. The statistical spread was specified by the 2.5 and the 97.5 percentile.

Unfortunately, some settings were different for the two versions of the software (see also 2.3.2.2): The stimulus interval for GATEe was 1500ms, whereas it was only 1200ms for GATEp. In GATEp approximately 4% of all questions were catch trials plus fixation controls, whereas GATEe catch trials made up 10% of

all presentations. Furthermore, GATEp had a higher number of questions due to a double testing of 5 test points in order to calculate short-term fluctuation for grid 30A. This is why an additional re-analysis of test times was done assuming the same stimulus interval setting for both software versions and excluding catch trials and stimuli that were presented in order to estimate short-term-fluctuation.
First, catch trial (positive, negative) and fixation control questions were eliminated from the number of questions asked (NQA). Second, for grid 30A and GATEp the mean number of questions needed to re-estimate the 5 points for the calculation of short-term-fluctuation were subtracted. Third, the difference between the stimulus interval of GATEp (1200ms) and GATEe (1500ms) of 300ms was subtracted of the test times per question. Afterwards the adjusted test times per question were multiplied with the adjusted numbers of questions asked and the resulting test times were compared. All this was done in order to assess hypothetical test times for both software versions with as consistent preconditions as possible.

2.4.3 Analysis of MS values

An important index of visual field evaluation is the mean defect (MD), which indicates the mean deviation of the individual hill of vision from the age-adjusted normative hill of vision. The MD is the mean of all defect values of a visual field. Defect values are calculated by subtracting the actually estimated DLS from the normative age-correlated sensitivity values. Since different normative values were underlain for the calculation of defect values in the different software versions (GATEp: normative values for presentation duration of 200ms, GATEe: normative values for presentation duration of 100ms), the MD values were not comparable. This is why, even though MD is the more relevant value to be analysed in perimetry studies, mean sensitivity (MS) values have been compared instead. MS is defined as the mean of all sensitivity values (DLS) of a visual field. Like MD it is a sensitive index of diffuse visual field loss, but can also be influenced by focal defects of sufficient depth or extent. [44]

Median MS values were assessed for the different serial devices and their range was specified by the 2.5 and the 97.5 percentile.

3. Results

3.1 Differential Luminance Sensitivity

3.1.1 GATEp1 vs. GATEp2 (Test-Retest-Reliability)

The comparison of GATEp1 vs. GATEp2 was done in order to specify the statistical agreement of GATEp. See figure 5 for results of grid 30A and figure 6 for results of grid 84NO. The modified Bland-Altman plots do not show trends, the differences do not tend to get larger with increasing average. Also, the variability of differences stays consistent across the graph. The results show very good agreement according to the clinical evaluation criteria.

For a summary of the maximum, minimum and median differences and the LOA of the examinations performed with GATEp as taken from the Bland-Altman plots see table 5. Table 6 shows the results separately for GATE-i and GATE.

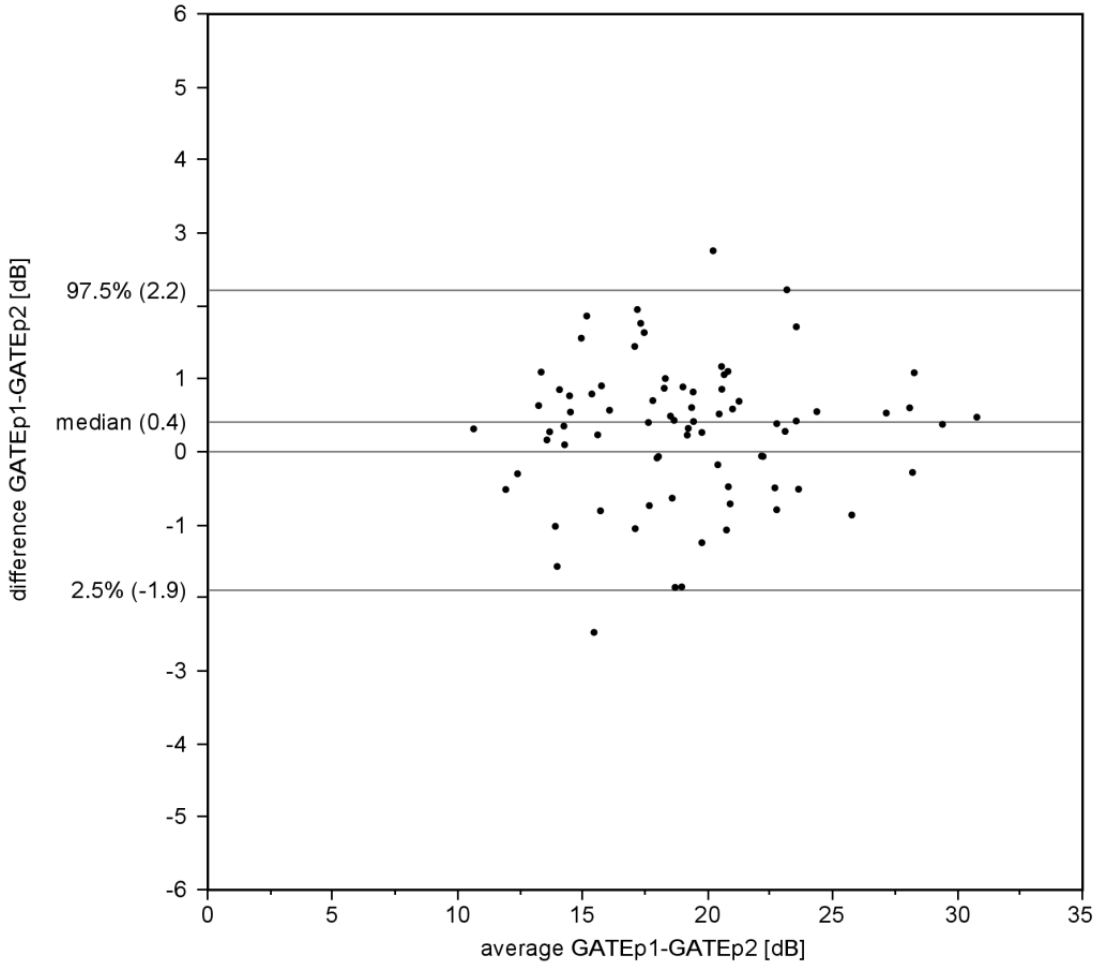

Figure 5: GATEp1 vs.GATEp2, grid 30A
Modified Bland Altman-plot of the examinations performed with GATEp1 and GATEp2 applying grid 30A. The Limits of agreement are depicted by the 97.5- and the 2.5-percentile. The bias is depicted by the median of differences between GATEp1 and GATEp2. For the sake of a faster valuation of the graph, a bias of 0 dB has also been depicted by a narrow line.
The x-axis shows the average DLS values per test location for the two examinations. The y-axis shows the difference of the achieved DLS values per test location of the examination with GATEp1 subtracted by GATEp2. All values are in dB

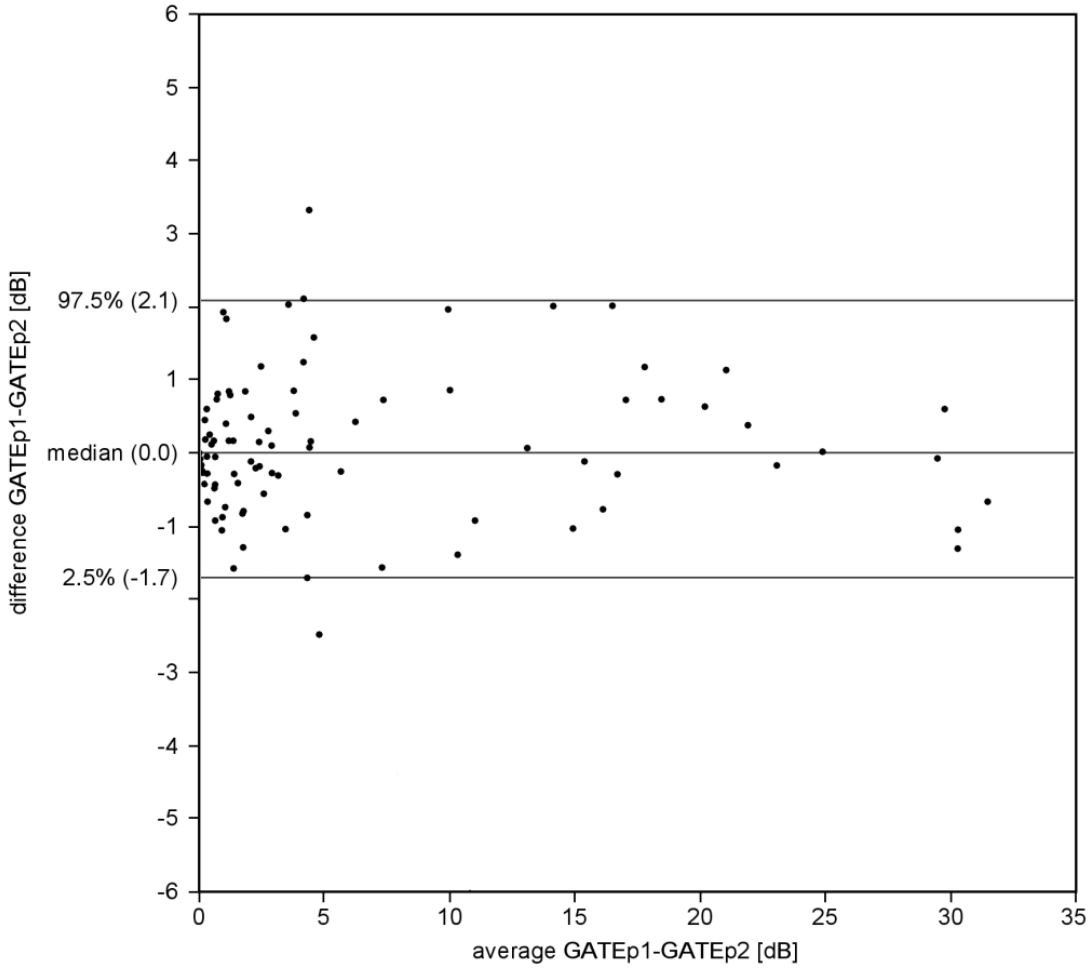

Figure 6: GATEp1 vs. GATEp2, grid 84NO
Modified Bland Altman-plot of the examinations performed with GATEp1 and GATEp2 applying grid 84NO. For explanation of abbreviations see figure 5

Table 5: Overview of bias and LOA and clinical evaluation criteria for GATEp1-GATEp2

Overview of bias and LOA for GATEp1 vs. GATEp2						Clinical evaluation criteria	
Grid	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
30A	-2.5	2.8	0.4	-1.9 to 2.2	-2.3 to 1.8	Very good	Very good
84NO	-2.5	3.3	0.0	-1.7 to 2.1	-1.7 to 2.1	Very good	Very good

Min. = minimum difference of DLS values between the two examinations
Max. = maximum difference of DLS values between the two examinations
Bias = median difference of DLS values between the two examinations
LOA = limits of agreement, i.e. 2.5- and 97.5-percentile
cLOA = corrected limits of agreement: LOA subtracted by the bias

Table 6: Overview of bias and LOA and clinical evaluation criteria for GATEp1-GATEp2 regarding GATE-i and GATE separately

Overview of bias and LOA for GATEp1 vs. GATEp2						Clinical evaluation criteria	
Grid 30A							
Strategy	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
GATE-i	-3.5	3.1	0.4	-2.5 to 2.6	-2.9 to 2.2	Very good	Very good
GATE	-1.8	4.2	0.4	-1.6 to 3.1	-2.0 to 2.7	Very good	Very good
Grid 84NO							
Strategy	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
GATE-i	-3.5	5.2	0.0	-2.9 to 3.9	-2.9 to 3.9	Very good	Very good
GATE	-2.5	3.0	0.1	-1.9 to 2.5	-2.0 to 2.4	Very good	Very good

For explanation of abbreviations see table 5

When evaluating the DLS results that were taken from the modified Bland-Altman plots comparing GATEp1 versus GATEp2 the cLOA values and biases show very good agreement for both grids. For grid 30A the median values of the differences indicate a small bias (0.4 dB), whereas no bias is found for grid 84NO.

The comparison of the LOA between GATE-i and GATE (as shown in table 6) shows larger LOA for GATE-i. This represents a greater variability for the initial examinations.

3.1.2 GATEp1 vs. GATEe894 (Comparison with Eyesuite, serial device 894)

For the assessment of the statistical agreement between GATEp and GATEe the results of the examination with GATEp1 (first examination with GATEp) and GATEe894 (examination with GATEe on serial device 894, maximum stimulus luminance of 1273 cd/m² [4000 asb]) were compared. The Bland-Altman plots for the comparison of GATEp1vs. GATEe894 are shown in figures 7 and 8 for the two different grids. The information that can be drawn from these plots is summarised in table 7.

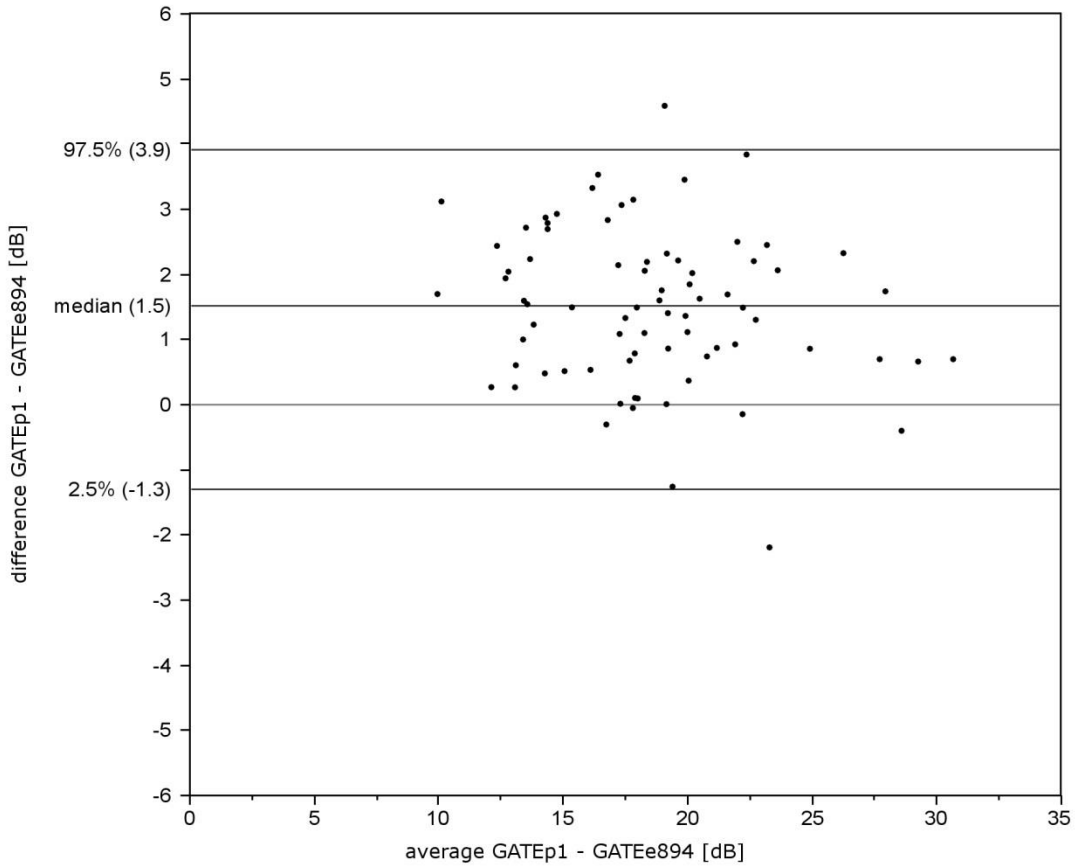

Figure 7: GATEp1 vs. GATEe894, grid 30A
Modified Bland Altman-plot of the examinations performed with GATEp1 and GATEe894 applying grid 30A. For explanation of abbreviations see figure 5

For grid 30A a bias of 1.5 dB is observed. This indicates systematically higher threshold values for GATEp1 compared to GATEe894. The same is observed – but only to a small extent (0.5 dB) - for grid 84NO. The differences between GATEp and GATEe, i.e. the range of the LOA, are greater than those between GATEp1 and GATEp2.

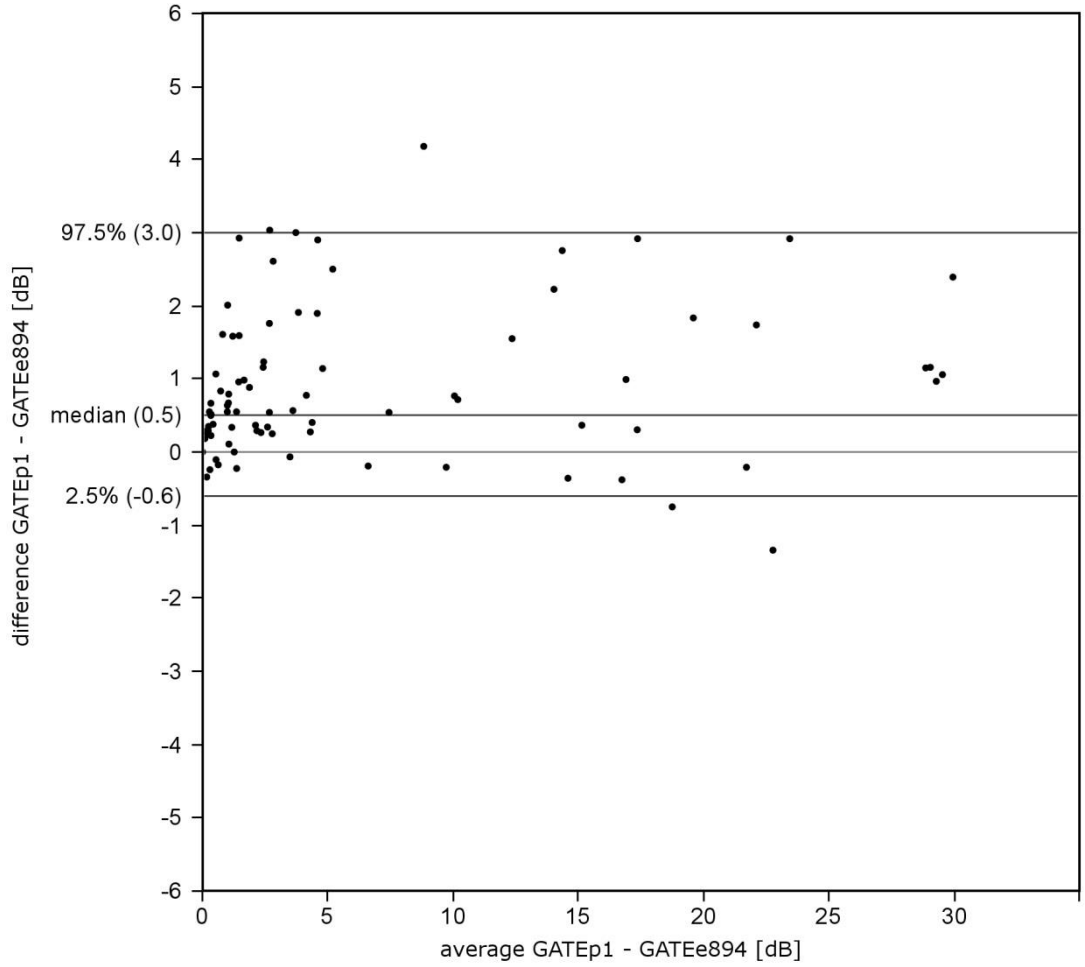

Figure 8: GATEp1 vs. GATEe894, grid 84NO
Modified Bland Altman-plot of the examinations performed with GATEp1 and GATE894 applying grid 84NO. For explanation of abbreviations see figure 5

Table 7: Overview of bias and LOA and clinical evaluation criteria for GATEp1 vs. GATEe894

Overview of bias and LOA for GATEp1 vs. GATEe894						Clinical evaluation criteria	
Grid	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
30A	-2.2	4.6	1.5	-1.3 to 3.9	-2.8 to 2.4	Acceptable	Very good
84NO	-1.3	4.2	0.5	-0.6 to 3.0	-1.1 to 2.5	Very good	Very good

For explanation of abbreviations see table 5

When applying the clinical evaluation criteria, GATEp1 and GATEe894 show acceptable to very good agreement for grid 30A and very good agreement for grid 84NO.

Table 8: Overview of bias and LOA and clinical evaluation criteria for GATEp1 vs. GATEe894 regarding GATE-i and GATE separately

Overview of bias and LOA for GATEp1 vs. GATEe894						Clinical evaluation criteria	
Grid 30A							
Strategy	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
GATE-i	-3.1	4.6	1.5	-2.2 to 4.5	-3.7 to 3.0	Acceptable	Good
GATE	-1.8	5.1	1.6	-1.4 to 4.7	-3.0 to 3.1	Acceptable	Good
Grid 84NO							
Strategy	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
GATE-i	-2.4	6.7	0.4	-2.1 to 5.3	-2.5 to 4.9	Very good	Acceptable
GATE	-1	4.3	0.7	-0.8 to 3.8	-1.5 to 3.1	Very good	Good

For explanation of abbreviations see table 5

The range of LOA is again greater for GATE-i than for GATE (see table 8). The clinical evaluation criteria also show acceptable to very good agreement between GATEp and GATEe.

3.1.3 GATEp1 vs. GATEe1075 (Comparison with Eyesuite, serial device 1075)

The results of the examinations with GATEp1 (first examination with GATEp) and GATEe1075 (examination with GATEe on serial device 1075, maximum stimulus intensity of 3183 cd/m² (10,000asb)) were also compared.

Table 9: Overview of bias and LOA and clinical evaluation criteria for GATEp1 vs. GATEe1075 regarding GATE-i and GATE separately

Overview of bias and LOA for GATEp1 vs. GATEe1075						Clinical evaluation criteria	
Grid 30A							
Strategy	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
GATE-i	-2.4	4.6	1.0	-1.5 to 4.4	-2.5 to 3.4	Good	Good
GATE	-2.9	4.1	1.2	-2.5 to 3.7	-3.7 to 2.5	Acceptable	Good
Grid 84NO							
Strategy	**Min. [dB]**	**Max. [dB]**	**Bias [dB]**	**LOA (2.5%-97.5%) [dB]**	**cLOA [dB]**	**Bias**	**cLOA**
GATE-i	-2.2	6.4	0.5	-1.8 to 5.1	-2.3 to 4.6	Very good	Acceptable
GATE	-0.9	3.9	0.8	-0.4 to 3.3	-1.2 to 2.5	Good	Very good

For explanation of abbreviations see table 5

Table 9 shows the biases and LOA of the comparison of GATEp1 and GATEe1075. Again, the clinical evaluation criteria state acceptable to very good agreement of the examinations. For grid 84NO the biases are comparable to those of the comparison with GATEe894. For grid 30A they are a little smaller.

3.1.5 Results regarding the literature criteria

Five test locations showed deviations between the examinations exceeding 5 dB. See table 10 for an overview of these locations.
They all show greater sensitivities for GATEp1 than for GATEe. There were no adjoining or paracentral test locations exceeding the 5 dB limits.

The grid 30A exception only exceeded the 5 dB by 0.1 dB. Its location ID 39 belongs to a midperipheral point in the superior temporal quadrant (see figure 9). Except for this one, all other test points exceeding the 5 dB deviation limits occurred with RP patients (grid 84NO) performing GATE-i. See figure 10.
Location ID 44 and location ID 53 are locations at the border of the central 30° visual field potentially interfering with the rim of the near correction glasses. As stated above, border points may therefore be ignored.
Location ID 27 is a midperipheral point in the lower nasal quadrant, not adjoining the other location IDs. This test location is either part of or directly adjacent to a scotoma in the visual fields of all patients examined with this grid. It is the only point, where the 5 dB limit is exceeded twice (for the comparison of GATEp1 with both GATEe894 and GATEe1075).
Location ID 10 is a rather central, but not paracentral point of the superior nasal quadrant. It was perceived by only half of the patients.

The following figures show the location of these test points within the grids.

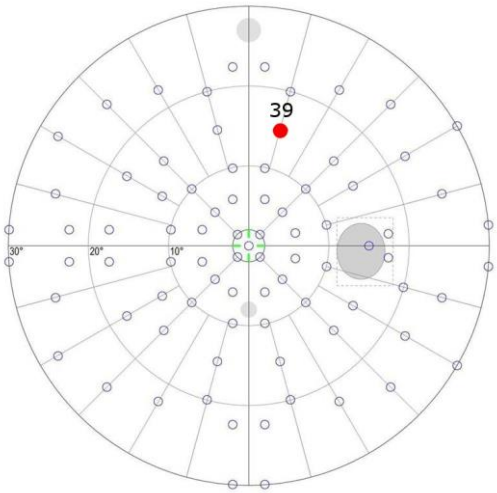

Figure 9: Test location exceeding the 5 dB deviation limit, grid 30A

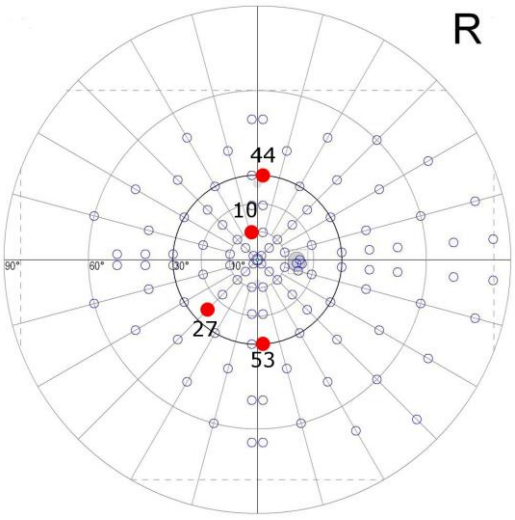

Figure 10: Test locations exceeding the 5 dB deviation limit, grid 84 NO

Table 10: Overview of the test points exceeding the 5 dB deviation limit

Grid	Examination	Location ID	Eccentricity/angle [°/°]	Difference [dB]
84NO	GATEp1 vs. GATEp2 GATE-i	44	(30.0/86.2)	5.2
	GATEp1 vs. GATEe894 GATE-i	53	(30.0/-86.2)	5.6
		27	(17.5/-135.0)	6.7
	GATEp1 vs. GATEe1075 GATE-i	27	(17.5/-135.0)	6.4
		10	(9.9/78.3)	5.5
30A	GATEp1 vs. GATEe894 GATE	39	(15.0/74.9)	5.1

3.2 Examination duration

Table 11 shows all test durations. GATE-i (i.e. the *initial* session) took 1.5 min (median value) longer than GATE (i.e. the *subsequent* session) for all perimeters as had been expected.

Overall median test duration (including both grids and examinations) of GATEp was 8.6 min (2.5, 97.5 interval: 5.5 min, 11.6 min), for GATEe 9.3 min (2.5, 97.5 interval: 6.3 min, 12.4 min).

Table 11: Overview of examination durations [min]

	GATE-i			GATE		
	Median	2.5-percentile	97.5-percentile	Median	2.5-percentile	97.5-percentile
30A						
GATEp1	9.2	7.8	11.6	7.7	6.0	9.6
GATEp2	9.3	7.4	11.1	7.9	6.0	8.7
GATEe894	9.9	7.8	11.2	8.8	7.1	10.7
GATEe1075	9.4	8.1	10.9	8.7	6.3	12.2
84NO						
GATEp1	9.9	9.2	12.5	6.4	5.3	8.4
GATEp2	9.6	9.0	12.9	7.0	5.2	8.3
GATEe894	10.8	7.2	12.4	7.6	6.3	9.1
GATEe1075	10.5	8.6	13.3	7.1	6.2	9.1

3.2.1 Simulation of test times assuming identical settings.

Table 12: Examination duration per question for the four perimeters

Serial device	GATE-i			GATE		
	Test time [min]	NQA	Time / question [s]	Test time [min]	NQA	Time / question [s]
GATEp1	9.4	326	1.74	7.5	283	1.60
GATEp2	9.5	338.5	1.69	7.7	280	1.65
GATEe894	10.1	290	2.09	8.6	250.5	2.06
GATEe1075	9.7	298.5	1.94	8.6	266	1.94

NQA = number of questions asked

When dividing the test time by the number of questions asked (NQA) - see table 12, it becomes clear that the test time per question was shorter for GATEp than for GATEe by approximately 300 ms for GATE-i and 375 ms for GATE. This was probably due to the different settings for the interstimulus interval. However, the test time per question was 120 ms - 150 ms longer for GATEe894 compared to GATEe1075, even though the settings for those examinations were exactly the same. Reasons for that should probably be sought in special characteristics of serial device 894 that used to pause shortly during some examinations without apparent reason, maybe because of problems with the driving laptop.

By subtracting the number of catch trials and the number of questions asked for the estimation of short-term-fluctuation from the overall number of questions

asked, we assessed the number of questions that were just and only asked in order to determine the DLS. Short term fluctuation was only assessed for grid 30A on serial device 104. Table 13 shows the results. All values in this table are median values.

Table 13: Calculation of adjusted number of questions

Serial device	GATE-i			GATE		
	grid	- catch trials	-SF	grid	- catch trials	-SF
1104	30A	310	292.4	30A	275	259.4
	84NO	292.5		84NO	200.5	
2104	30A	322	303.7	30A	285.5	269.6
	84NO	285		84NO	214	
894	30A	260.5		30A	228.5	
	84NO	273.5		84NO	193.5	
1075	30A	259		30A	240.5	
	84NO	288.5		84NO	194	

SF = number of questions asked for the assessment of short term fluctuation

After that the test times per questions of GATEe were adjusted by subtracting 300 ms, which is the difference between the stimulus intervals of GATEp and GATEe. Afterwards the adjusted test times per question were multiplied with the adjusted numbers of questions. See table 14 for the resulting adjusted test times. All values are median values.

Table 14: Adjusted test times

Serial device	GATE-i			GATE		
	Adjusted test time [min]	Adjusted NQA	Adjusted time / question [s]	Adjusted test time [min]	Adjusted NQA	Adjusted time / question [s]
30A						
1104	8.5	292.4	1.74	6.9	259.4	1.60
2104	8.6	303.7	1.69	7.4	269.6	1.65
894	7.8	260.5	1.79	6.7	228.5	1.76
1075	7.1	259	1.64	6.6	240.5	1.64
84NO						
1104	8.5	292.5	1.74	5.3	200.5	1.60
2104	8.0	285	1.69	5.8	214	1.65
894	8.2	273.5	1.79	5.7	193.5	1.76
1075	7.9	288.5	1.64	5.3	194	1.64

NQA = number of questions asked

After these transformations for enhancing comparability table 14 now shows similar test times for GATEp and GATEe.

3.3 Mean sensitivity (MS)

When looking at the MS values of the different groups of patients, AION and glaucoma patients, who suffer from similar visual field defects, show similar MS values (about 17.7dB). The MS of patients with (post-) chiasmal lesions is higher than for the other patients (about 20.0 dB). RP patients who suffer from an essential constriction of the visual field achieve the lowest MS values of approximately 5.2 dB in median.

When analysing the median MS values for the different perimeters, serial devices 104 and 894 show comparable median MS values with only slightly higher results for 894. The median MS value of serial device 1075 (GATEe1075, higher maximum stimulus intensity) is about 5 dB higher (median value) than that of the perimeters with lower maximum stimulus intensity. This effect can still be observed, even if median MS values are calculated for the different visits (GATE and GATE-i) and for the different diseases (see table 15). The greatest differences between serial devices 1075 and 104 (GATEp) were found for AION and glaucoma patients (approximately 6 dB), who are mainly suffering from relative scotomas. For hemianopia patients the difference is approximately 4.5 dB and for RP patients only 2.5 dB (see also Table 15). Median MS values are smaller for GATE compared to GATE-i by approximately 1 dB, except for serial device 1075.

In summary, MS values of 104 and 894 are comparable, but examinations done with serial device1075 achieve higher MS values, especially for patients with relative scotomas.

Table 15: Median MS values overview

	1104 (GATEp1)	**2104 (GATEp2)**	**894 (GATEe894)**	**1075 (GATEe1075)**
Median MS (2,5 percentile, 97.5 percentile) [dB]				
Overall results for the different perimeters				
	15.6 (0.8, 23.7)	14.7 (0.6, 23.4)	15.9 (2.0, 23.8)	20.2 (2.9, 28.2)
Groups of diseases				
AION	16.1 (14.1, 20.0)	15.3 (13.6, 20.8)	15.8 (13.7, 22.8)	22.0 (19.5, 28.2)
Glaucoma	16.9 (5.7, 23.7)	16.4 (5.2, 23.7)	16.8 (7.9, 23.8)	21.9 (9.3, 28.2)
HH	19.5 (13.7, 22.6)	19.4 (14.0, 23.3)	19.9 (15.5, 23.6)	24.0 (17.2, 27.2)
RP	4.1 (0.8, 5.9)	4.3 (0.6, 6.6)	5.2 (1.8, 7.0)	6.7 (2.7, 9.3)
Visit strategy				
GATE-i	15.9 (0.8, 23.6)	15.0 (0.7, 23.5)	16.2 (1.8, 23.8)	20.5 (2.7, 28.2)
GATE	14.5 (0.9, 23.7)	14.6 (0.6, 23.3)	15.9 (2.3, 23.8)	19.8 (3.1, 28.2)

4. Discussion

4.1 DLS differences

4.1.1 Clinical evaluation criteria

At first glance, an aberration of 3 dB seems a rather big difference to be rated as very good agreement. However, a short-term fluctuation of 3 dB is supposed to be normal or at least inside the 95% reference interval [44]. For glaucoma patients (i.e. half of the patients examined in this study) short-term fluctuation is usually even greater than for ophthalmologically normal subjects. [49]
Furthermore, the LOA represent a 95% reference range, which means that most differences were smaller than these values. But in order to not obscure potential progression of the disease, LOA greater than 5 dB were rated as not acceptable.
Regarding the bias, differences < 0.5 dB are not relevant in clinical practice, since the perimeters only assess DLS with a measurement accuracy of 0.5 dB [38]. A bias exceeding the supposed normal short-term fluctuation (i.e. > 3 dB) was rated as not acceptable.

4.1.2 Literature criteria

There are manifold definitions regarding progression of glaucoma in various studies. Since in this study only local DLS values were assessed, the search was restricted to clinical criteria that referred to DLS.
Several different glaucoma studies defined a worsening by at least 5-10 dB at 2-3 adjacent test points outside the central visual field as a progression or a new manifestation. [45,46] For example, Anderson and colleagues defined a progression of glaucoma, if there was a worsening of ≥ 3 points by ≥ 10 dB in an existing defect or if a worsening of ≥ 2 new adjacent points by at least 10 dB had taken place. [47] A minimum depression of 9 dB in peripheral test locations and a depression of 5 dB in paracentral points are needed to elevate the score

indicating a progression of visual field defects in the AGIS score (Advanced Glaucoma Intervention Study). [48]
This is why all test points exceeding the 5 dB deviation limit were checked for adjacency and their location in the visual field. Two of them were negligible border points. Two other test points (location ID 27 and10) were either part of or adjacent to scotomas for all examined patients (RP). So, due to small fluctuations in fixation these test points could be “swallowed up” by the adjacent scotomas, which could lead to great fluctuation in the perceived DLS level. Maximum differences between the procedures were far below 10 dB. For all test conditions, the upper (i.e. 97.5%) LOA values were below 5.3 dB.

4.1.3 DLS results

Visual field results represent a large and complex physiologic variability [50–52]. This threshold variability has been shown to increase with progressive eccentricity [53]. For glaucoma patients inter- and intra-subject variability is even larger [49,54] and local increase of variability within and between tests may even be the first visual field disturbance detectable [55,56]. Various studies showed that test-retest variability of threshold perimetry increases with decreasing sensitivity until it declines again near 0 dB (floor effect) [22,57–59]. All patients included in this study suffered from moderate to severe visual field loss. An acceptable to very good agreement could be stated for all comparisons between the examinations, i.e. regarding intra- *and* inter-subject agreement. So, from a clinical point of view, the agreement between GATEp and GATEe is sufficient.

The agreement of two methods is limited by the repeatability of these methods. Furthermore, since variability increases with decreasing sensitivity of the visual field of the patients [59] and the agreement of two different methods is limited to their repeatability [40,41] it is not surprising that the repeatability of GATEp itself is better than the agreement between GATEp and GATEe.

A bias of approximately 1.5 dB was observed for grid 30A indicating a systematic tendency of GATEp to assess higher DLS values than GATEe. This may be due to the methodological differences between the two procedures [42]. For grid 84NO this bias is only approximately 0.5 dB. The comparatively small bias is probably the result of a ceiling effect due to the extended visual field losses resulting from the advanced disease of the RP patients. When comparing SITA algorithms with the FT strategy, biases of approximately 1 dB have been found [57,60]. Since perimetric examinations should usually be evaluated by follow-up examinations and trend analysis, they should be performed with the same software and strategy for each examination. Therefore, such systematic errors should not relevantly deform visual field results or impact visual field evaluation.

GATE utilizes local thresholds from previous examinations instead of testing starting points (like GATE-i). This saving of time is very valuable in a clinical setting. Furthermore, the results may be more accurate and repeatable [30]. However, the risk to bias the results toward previous findings could increase in cases of immediate and pronounced change, like for example inflammation, trauma or infarction. Such a case would result in a prolonged threshold approach due to the assumption of (in the meantime) invalid previous local threshold values. However, the majority of ophthalmological diseases that need follow-up shows chronic progression like for example glaucoma, compressive or hereditary optic neuropathy or degenerative retinal diseases like age-related macular degeneration or tapeto-retinal degeneration. For these patients immediate changes are rarely the case. To assume normal conditions in these patients and neglect previous findings could result in a considerable prolongation of test duration and could therefore provoke fatigue.

4.2 Examination duration

4.2.1 Comparison of GATE-i and GATE

Longer test durations for GATE-i than for GATE were to be expected, since GATE-i examines more test points and refers to age-related standard values instead of taking the patients' last examination as the basis for the starting luminance of the test points.

4.2.2 Comparison of GATEp and GATEe

The unexpectedly longer examination durations of GATEe of partly up to one minute (i.e. approximately 8%) are probably explained by the different settings, especially due to the fact that the presentation interval for GATEp was only 1200 ms compared to 1500 ms for GATEe (300 ms difference, i.e. 25% longer intervals). These settings were different, because the stimulus interval could be typed in for GATEp, but had to be chosen from a regulator with predetermined values for GATEe. The labeling of the regulator led to believe that the same stimulus interval had been chosen. The real stimulus interval was not accessible before the end of the examinations. Hence, a simulation of test times assuming identical settings that indicated similar test times, was performed.
It has to be mentioned, however, that the calculations only offer an approximation of possible examination durations, since a change in stimulus interval may possibly have an influence on response behavior. Assuming an adequate approximation, the incorporated GATEe strategy offers the same advantages as have been shown for GATEp in a prior study regarding test time [13]. GATEe is therefore a good alternative to any other fast-thresholding estimation strategy. Consistent settings should be realized in future studies in order to confirm the simulated results.

4.2.3 Comparison with SITA Standard

In several studies with patients suffering from visual field loss average or median test times of about 6-8 min were achieved for SITA Standard with a 24-2 pattern (52 test locations (TL), 0.12-0.15 min/TL, i.e. 7.2-9.0 s/TL). Shorter

test times of about 5 min could be achieved for healthy test subjects [13,22,58]. The median test times for GATEe in this study with visually impaired test subjects were 8.8 min for grid 30A (83 TL, 0.11 min/TL, i.e. 6.6 s/TL) and 7.6 min for grid 84NO (109 TL, 0.07 min/TL, i.e. 4.2 s/TL) [42]. If comparing the needed time per test location, GATEe might even be faster than SITA standard if the same patients were tested with the same grid. When compared to the full threshold (FT) strategy, GATE showed considerably shorter test times. [13]

For GATE no increase of test time is observed for increasing visual field loss. RP patients with extended visual field loss show the shortest examination durations of the four groups. This is probably due to the comparatively high proportion of test points with absolute scotomas which do not need a time-consuming thresholding strategy. In contrast, increasing test times for increasing visual loss have been shown for SITA [13,22].

4.2.4 Test times compared to a prior study

In a prior study of Schiefer et. al. examination durations for a group of 40 patients with manifest glaucoma, 10 patients with suspected glaucoma and 10 patients with ocular hypertension were 5.7 min for GATE-i and 4.7 min for GATE, while SITA Standard took 5.6 min and FT needed 9.0 min in the first study that involved the GATE algorithm [13]. In this study, however, GATE-i needed 9.3 min and GATE needed 7.6 min. These longer test times were probably due to the use of different test point arrangements. 83 and 109 test locations have been tested in this study, whereas pattern 24-2 grid that was used in the study of Schiefer et.al. only tested 53 test locations covering the central 24° visual field. The longer test times could also be due to the different sample of subjects.

4.2.5 The effect of shorter test times on perimetric performance

Shorter test times are supposed to reduce fatigue and reliability problems. Marra et al. suggested that for test times of 5-8 minutes no major trend of either learning or fatigue effect is observed in a single session [4]. All this should lead

to a higher reliability of test results due to shorter test times. Furthermore, patients are supposed to be more motivated and attentive and less bored by the examination and even patients who can only concentrate for a short period of time may adequately perform threshold estimating visual fields. Another advantage of shorter test times is time saving in clinics and practices, which is practical, but also economically important.

4.3 Mean Sensitivity

We unfortunately cannot provide the Mean Defect (MD) for the GATEe examinations, since the normative data base for MD of EyeSuite is referring to a stimulus duration of 100 ms, whereas the stimulus duration was 200 ms in our study, which was chosen for comparability reasons with the standard setting over decades for all other perimetric examinations in the university eye hospital in Tuebingen. This is why we assessed and compared MS values.

4.3.1 Background luminance

A background intensity of 10 cd/m² was chosen in this study in order to operate under photopic conditions. This is reasonable because for photopic adaption of the eye Weber's law is valid which states that the necessary differential luminance (level) for stimuli rises linearly with the background luminance [34]. Under photopic conditions exclusively the cone photoreceptor system can be tested. It furthermore offers a fast adaptation for patients who have usually been exposed to a bright environment beforehand. Also, examination results are less depending on pupil size than for non-photopic conditions [61].
This is why a background luminance of 10 cd/m² (31.4asb) was recommended by the International Perimetric Society in 1978 to be the standard for perimetric examinations [62].

4.3.2 Maximum stimulus luminance and its influence on MS values

The three different perimeters offered three different maximum luminance levels. This is important, because the maximum luminance is the reference value for the logarithmic relation scale which is the measurement unit for DLS

values (max. stimulus intensity = 0dB) [11] and this is the reason why the DLS results in dB of the different perimeters had to be translated by the above mentioned formulas in order to be comparable.

For a comparatively dim background luminance of 1.27 cd/m² (4 asb) stimulus intensities higher than 318 cd/m² (1000 asb) may produce disturbing stray light. This can lead to an irradiation into scotoma areas and by that falsify sensitivity estimates and lead to an underestimation of defect depth and size [63]. This is why the maximum level of stimulus luminance is limited in order to prevent stray light, which increases with brighter stimuli leading to a replacement of local responses by stray-light responses [64]. In this study a remarkable increase of MS values was found for the perimeter with a maximum stimulus luminance of 3183 cd/m² (10,000 asb, serial device 1075) compared to the other perimeters with maximum stimulus luminances of 318 cd/m² (1000 asb, serial device 104) and 1273 cd/m² (4000 asb, serial device 894), respectively. It is not quite obvious, however, if this effect is due to stray light, because in this study for all perimeters a background luminance of 10 cd/m² (31.4 asb) instead of 1.27 cd/m² (4 asb) was used. If the increased MS values were due to stray light and not only to the reduction of scotomas, there should rather also be a noticeable effect on the size of the blind spot.

The effect of higher MS values for higher maximum stimulus intensities was smallest for extended absolute scotomas (RP patients) and most pronounced for scotomas with higher portions of relative scotomas (glaucoma patients). The greatest differences of MS values between serial devices 1075 and 104 were shown for AION and glaucoma patients (approximately 6 dB), who quite often suffer from relative scotomas. For RP patients, who mostly suffer from extended absolute scotomas the MS difference was only approximately 2.5 dB. The results therefore support the theory that the influence of high maximum stimulus intensities is greater for relative scotomas than for absolute scotomas. That means, even if there seemed to be an effect of software on MS values at first glance, this effect should rather be attributed to the higher maximum

stimulus intensity than to the GATE algorithm itself. This point is supported by the fact that MS values of serial device 894 and serial device 104 were comparable, but MS values estimated by serial device 1075 were relevantly higher.

It seems logical that higher stimulus luminance levels lead to higher global MS in patients with areas of deep visual field loss, because stimuli with higher luminance levels rather provoke a reaction in these areas than stimuli with lower luminance levels. The importance of the comparison of global MS values should therefore be discussed. Other circumstances that could possibly have an effect on the MD of standard automated perimetry (and by that also on MS values) could be time of day, season, experience of the investigator, the rate of false-positive responses [8] and pupil size [65].

4.4 Examination parameters

4.4.1 Stimulus size

Common standard stimulus sizes have been introduced by Goldmann. Five different sizes are defined by Roman numerals I (0.25mm², i.e. 6.5 min of arc) to V (64mm², i.e. 104 min of arc), each covering a 4-fold greater area than the previous stimulus size. The most commonly used stimulus size in standard automated perimetry is the Goldmann size III stimulus (4 mm², i.e. 26 min of arc) [6,11].
This stimulus size was also chosen in this study, because it is big enough to offer a good dynamic range and reduce refraction errors, but at the same time small enough to avoid missing detection of small scotomas. Furthermore, this is the mandatory stimulus size for examinations regarding expert opinion examinations as recommended by the Transport and Traffic Committee of the scientific association of ophthalmology in Germany (Deutsche Opthalmologische Gesellschaft) for standard automated perimetry [66].

The larger the stimulus, the smaller is the influence of refraction errors and the greater is the dynamic range, especially in the peripheral visual field. Taravati et al. observed an exponential rise in variability with decreasing sensitivity for glaucoma patients, whereas variability stayed almost constant, if a very large stimulus was used [67]. Wall et al. also observed a reduction of variability and a slightly better repeatability of MD for the use of size V stimuli when testing glaucoma patients with normal or moderately damaged visual fields [68,69].

The Goldmann III stimulus size was found to be more useful for the detection of field aberrations in RP patients compared to a bigger stimulus size. If progression surveillance in RP patients would have been the aim of the examinations, a size V stimulus would have been more useful, however [70].

4.4.2 Stimulus presentation duration

Stimulus presentation duration should not be shorter than 100 ms in order to prevent temporal summation (Bloch's law) [34], but should not be longer than 200 ms in order to prevent eliciting gaze movements towards the stimulus [1]. Both effects would possibly affect the results of field examination. Other studies say that there seems to be no relevant effect of longer stimulus presentation on fluctuation of perception for stimulus durations between 65 ms and 500 ms [71]. However, for this study the stimulus presentation duration of 200ms has been chosen in order to compromise the above mentioned possible limitations.

4.4.3 Acoustic cueing

The mirror units of perimeters with a mirror-projection system produce ambient noise for mechanical reasons that may indirectly announce a following stimulus presentation. In addition to that, each stimulus (or false-positive catch trial) was presented accompanied by a beep. The projection noise and the beeps are both acoustic cues for the patients and may potentially have either stimulated "trigger happiness", annoyance of the patients and disturbance of the patients' concentration or they could possibly have been a help to concentrate on the

other hand. Different reactions came from the patients concerning acoustic cueing dependent on their subjective sensations.
For serial device 104 the signal sound beeps were presented at the same time as the stimulus presentation, whereas for serial devices 894 and 1075 the signal sounds did not coincide with stimulus presentation.
The effect of the stimulation of trigger happiness by acoustic cueing and by that shorter reaction times and higher rates of false-positive responses could not be confirmed by Rauscher et. al. for kinetic perimetry [72].
Lewald et al. stated that passive auditory stimulation of the blind side may improve vision in hemianopia patients probably due to an activation of residual visual pathways [73]. However, this should have had little effect on the results of this study, because hemianopia patients made up for only 24% of the examined subjects and the acoustic cues were always presented in the same way (not according to the blind side of each patient). Another study showed an improvement of vision in hemianopia patients by improving oculomotor patterns after audio-visual stimulation [74]. This may probably rather affect kinetic perimetry than static perimetry. It has been shown, however, that co-occurring acoustic stimuli may affect visual sensitivity by influencing the perception of visual stimuli [75]. Further studies focussing on the relationship between visual and acoustic perception would be very interesting and important.

4.5 Patient-related parameters

4.5.1 Inclusion and exclusion criteria

The selection of inclusion and exclusion criteria is always a delicate issue and crossroads between the risk of homogeneity and stratification and the wish to ensure representativeness and transferability to real-life conditions. In fact, pregnancy and breastfeeding, may affect vision and visual field [76]. In order to establish stable conditions, especially with regard to test-retest reliability we decided to exclude pregnant and lactating women as well as subjects with asthma.

Diabetic retinopathy is characterized by a variety of potentially confounding impacts on ocular, cerebral and other structures, which may critically interfere with the intended stratification of the sample. Cerebrovascular accident has not been flagged as an exclusion criterion for the fourth group.

4.5.2 Pupil size

The pupil is the natural optical aperture of the eye, controlling the amount of light that is allowed to enter the eye in order to improve imaging on the retina. However, if pupil diameter falls to less than 2.5 mm, the edge of the pupil causes light diffraction resulting in an impaired resolution [77]. Webster et al. have shown that miosis has a considerable effect on Mean Deviation (MD), but does not affect pattern standard deviation, which indicates a uniform reduction of sensitivity caused by small pupil size [65]. Pupil size should therefore exceed 3 mm, in order to prevent influences on the differential luminance sensitivity. Pupil dilatation, however, should be avoided, because the collection of normative data was done with normal, undilatated eyes and furthermore, dilatation results in changes in refraction (e.g. due to spherical aberration), which can also influence differential luminance sensitivity [78].

Another interesting issue about pupil size and perimetry is that fatigue wave amplitude and miosis are indicators of decreasing vigilance of patients and can be documented by infrared pupillography [37,79].

4.5.3 Refraction

Inadequate refractive correction may lead to perimetric outcome errors due to a blurred retinal image formation of the stimulus. This effect is more pronounced for smaller stimuli and may result in refraction scotomas [11].

4.5.4 Patient reliability indices

Patient reliability is an important factor for reproducibility in perimetry, because poor reliability may result in under- or over-estimation of differential luminance sensitivity. A patient’s inattention results in false-negative responses, since the

patient does not respond to a supra-threshold stimulus that was previously perceived. The rate of false negative answers is increased in case of severe visual field loss and leads to an artificial increase in visual field loss [80]. False-positive catch trials, which are characterized by patients' responses without stimulus presentation, are a good predictor for "trigger happiness" or "guessing" of the test subject. Increased false-positive response rates result in an artificially reduced visual field loss. If patient reliability is decreasing during examination, the patient should be reinstructed [77]. Catch trials, however, offer a surprisingly imprecise prediction of real reproducibility of the field status [81].

4.5.5 Learning effect

Learning effects in inexperienced subjects performing automated static perimetry are common. Studies have shown that this effect is rather small and usually limited to the first sessions. Sensitivity may increase with perimetric training. These effects could be shown for normal subjects [82,83] and also for patients with glaucoma [84]. It is therefore important, to perform more than one test in order to create a reliable baseline for perimetric follow-ups.
All patients in this study had undergone perimetric testing before. Learning effects should therefore have been minimized.

4.5.6 Fatigue effect

As perimetry is normally a rather long-lasting examination the effect of fatigue and decreasing vigilance has been tested before and proved to be existent and to influence the results of perimetric examinations. The fatigue effect, i.e. decreasing measured DLS values during examination, has been found in normal subjects [85], but also in glaucoma patients, where it resulted in the increasing of depth and/or size of defects. This effect is usually more pronounced with increasing eccentricity [86], mostly within the midperipheral field and with increasing age of the test subjects [87]. Shortening of test duration may reduce fatigue in perimetry.
Fatigue effects should not have caused systematic errors in this study, because the sequence of the various methods has been randomized.

4.6 Statistical analysis

Bland and Altman state that the correlation coefficient would be no indicator of agreement at all and the test of significance would not be relevant when asking for agreement [88]. Since perimetry is a technique rather biased by coincidental measurement errors (for example because of vigilance problems), analysis of regression is not an adequate approach and is also dependent on the range of measured values, which is also the case for the determination of the intraclass-correlation-coefficient [41]. The plotting of average results against the differences between the results, however, offers an easy-to-interpret graphical method to explore and illustrate statistical relationships, which can also be done in a nonparametric approach [40], for example for small samples of data. This was another reason for choosing the graphical approach of Bland and Altman for the comparatively small sample of 30 test subjects, only allowing descriptive statistics.

4.7 Future trends and perspectives

Since GATE is not restricted to a special type of ophthalmological disease, it can be applied to all kinds of visual pathway lesions – in contrast to the SITA strategy. A great advantage of GATE is the possibility to test any (arbitrary) test location. This allows for individual adding of test points in regions of interest, thereby enhancing spatial resolution in these areas [13].

4.8 Conclusion

This study shows a very good repeatability of the prototype version of the new fast thresholding algorithm GATE (German Adaptive Thresholding Estimation) with regard to differential luminance sensitivity (DLS) values and a good agreement between the prototype version of GATE (GATEp) and the commercially available version of GATE incorporated into the EyeSuite software package (GATEe). The results suggest that they can be used interchangeably. The GATE thresholding algorithm offers short examination durations and is not restricted to glaucomatous field loss.

5. Summary

Purpose

To validate the commercially available “EyeSuite” version of the new fast thresholding algorithm GATE (German Adaptive Thresholding Estimation) for automated static perimetry.

Methods

Thirty patients suffering from visual pathway lesions of various origin (anterior ischemic optic neuropathy [n=3], glaucoma [n=15], (post-) chiasmal visual pathway lesion [n=6], retinitis pigmentosa (RP) [n=6]) were tested on three Octopus 900 perimeters (Haag-Streit AG, Köniz, Switzerland) with various maximum stimulus intensities (serial device 104: 318 cd/m², 894: 1273 cd/m² and 1075: 3183 cd/m²). Grid 84NO (90° eccentricity, 109 test locations) was applied for patients with RP, grid 30A (30° eccentricity, 83 test locations) for all other patients. Repeatability of the prototype version of GATE (GATEp) and agreement between GATEp and the commercially available EyeSuite version (GATEe) were assessed by comparing local differential luminance sensitivities (DLS) and median test durations by means of modified Bland-Altman plots. Mean sensitivities [MS] were compared.

Results

The comparison of DLS values showed very good repeatability for GATEp (bias <0.5 dB, limits of agreement [LOA] <3 dB) and a very good to acceptable agreement between GATEp and GATEe (bias <2 dB, LOA <5 dB). Median examination durations for GATEp and GATEe were 7.8 min and 8.8 min for grid 30A, 6.7 min and 7.8 min for grid 84NO. MS values were comparable for both software versions, but higher values were assessed by the perimeter with the highest maximum stimulus luminance (serial device 1075).

Conclusion

The prototype version of GATE (GATEp) shows a very good repeatability. GATEp and the software version implemented in the EyeSuite software (GATEe) show good agreement regarding local differential luminance sensitivity and examination duration. GATEe can therefore be recommended for clinical practice.

5.1 Zusammenfassung

Ziel

Validierung der kommerziell erhältlichen „EyeSuite“-Version von GATE (German Adaptive Thresholding Estimation), eines neuen Algorithmus für schnelle Schwellenbestimmung in der automatischen statischen Perimetrie.

Methoden

30 Patienten mit Sehbahnläsionen unterschiedlicher Ursachen (anteriore ischämische Optikusneuropathie [n=3], Glaukom [n=15], (post-) chiasmale Sehbahnläsion [n=6], Retinitis pigmentosa (RP) [n=6]) wurden an drei Octopus 900 Perimetern (Haag-Streit AG, Köniz, Schweiz) mit verschiedenen maximalen Stimulusleuchtdichten (Seriengerät 104: 318 cd/m², 894: 1273 cd/m² und 1075: 3183 cd/m²) untersucht. Raster 84NO (90° Exzentrizität, 109 Prüfpunkte) wurde für RP-Patienten angewandt, Raster 30A (30° Exzentrizität, 83 Prüfpunkte) für alle anderen. Die Reproduzierbarkeit der Prototypversion von GATE (GATEp) und die Übereinstimmung zwischen GATEp und der käuflich erwerbbaren EyeSuite-Version (GATEe) wurde bestimmt, indem lokale Lichtunterschiedlichkeitsempfindlichkeiten (LUE) und die mediane Untersuchungsdauer mit Hilfe modifizierter Bland-Altman-Diagramme verglichen wurden. Mean Sensitivity-(MS) Werte wurden verglichen.

Ergebnisse

Der Vergleich der LUE-Werte zeigte eine sehr gute Reproduzierbarkeit der Messungen für GATEp (Bias <0.5 dB, limits of agreement [LOA] <3 dB) und eine sehr gute bis akzeptable Übereinstimmung zwischen GATEp und GATEe (Bias <2 dB, LOA <5 dB). Die medianen Testzeiten betrugen für GATEp und GATEe 7.8 min und 8.8 min (Raster 30A), 6.7 min und 7.8 min (Raster 84NO). Die MS-Werte waren für beide Softwareversionen vergleichbar, aber höher für das Gerät mit der größten maximalen Stimulusleuchtdichte (Seriengerät 1075).

Fazit

Die Prototypversion von GATE (GATEp) zeigt eine sehr gute Wiederholbarkeit. GATEp und die in EyeSuite implementierte Version, GATEe, zeigen in Bezug auf die LUE-Werte und Testzeiten eine gute Übereinstimmung. Daher kann GATEe für die klinische Praxis empfohlen werden.

6. References

1 Schiefer U, Pätzold J, Dannheim F. Konventionelle Perimetrie - Teil 1: Einführung - Grundbegriffe [Conventional techniques of visual field examination. Part I: Introduction - basics]. *Ophthalmologe* 2005;**102**:627–646.

2 Traquair HM. *An introduction to clinical perimetry*. Kimpton 1938.

3 Wall M. What's new in perimetry. *J Neuroophthalmol* 2004;**24**:46–55.

4 Marra G, Flammer J. The learning and fatigue effect in automated perimetry. *Graefes Arch Clin Exp Ophthalmol* 1991;**229**:501–4.

5 Schiefer U, Schiller J, Flad M. Konventionelle Perimetrie - Aktueller Stand und künftiges Entwicklungspotential. In: *Augenärztliche Diagnostik*. Thieme Verlag: Stuttgart 2003. 93–108.

6 Schiefer U, Pätzold J, Dannheim F. Konventionelle Perimetrie - Teil 2: Konfrontationsperimetrie - Kinetische Perimetrie, [Conventional techniques of visual field examination Part 2: confrontation visual field testing - kinetic perimetry]. *Ophthalmologe* 2005;**102**:821–7.

7 Gardiner SK, Demirel S. Assessment of patient opinions of different clinical tests used in the management of glaucoma. *Ophthalmology* 2008;**115**:2127–31.

8 Junoy Montolio FG, Wesselink C, Gordijn M, *et al.* Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day, and season. *Invest Ophthalmol Vis Sci* 2012;**53**:7010–7.

9 Wild J. IPS Standards and Guidelines 2010. Imaging and Perimetry Society (IPS)

10 Kutzko KE, Brito CF, Wall M. Effect of instructions on conventional automated perimetry. *Invest Ophthalmol Vis Sci* 2000;**41**:2006–13.

11 Flammer J. Theoretische Grundlagen der automatischen Perimetrie. In: Gloor B. *Automatische Perimetrie, Bücherei des Augenarztes, Band 110,* Stuttgart: Enke 1987. 1–31.

12 Schiefer U, Pätzold J, Wabbels B, *et al.* Konventionelle Perimetrie - Teil 3: Statische Perimetrie: Raster - Strategien - Befunddarstellung. *Ophthalmologe* 2006;**103**:149–63.

13 Schiefer U, Pascual JP, Edmunds B, *et al.* Comparison of the new perimetric 'German Adaptive Threshold Estimation' (GATE) strategy with

conventional full-threshold and SITA Standard strategies. *Invest Ophthalmol Vis Sci* 2009;**50**:488–94.

14 Barton JJS, Benatar M. *Field of vision: a manual and atlas of perimetry*. Totowa, N.J.: Humana Press 2003.

15 Flanagan JG, Moss ID, Wild JM, *et al.* Evaluation of FASTPAC: a new strategy for threshold estimation with the Humphrey Field Analyser. *Graefes Arch Clin Exp Ophthalmol* 1993;**231**:465–9.

16 Schaumberger M, Schäfer B, Lachenmayr BJ. Glaucomatous visual fields - FASTPAC versus full threshold strategy of the Humphrey Field Analyzer. *Invest Ophthalmol Vis Sci* 1995;**36**:1390–7.

17 Morales J, Weitzman ML, Gonz lez de la Rosa M. Comparison between Tendency-Oriented Perimetry (TOP) and octopus threshold perimetry. *Ophthalmology* 2000;**107**:134–42.

18 Gonzalez-Hernandez M, Morales J, Azuara-Blanco A, *et al.* Comparison of Diagnostic Ability between a Fast Strategy, Tendency-Oriented Perimetry, and the Standard Bracketing Strategy. *Ophthalmologica* 2005;**219**:373–8.

19 King AJ, Taguri A, Wadood AC, *et al.* Comparison of two fast strategies, SITA Fast and TOP, for the assessment of visual fields in glaucoma patients. *Graefes Arch Clin Exp Ophthalmol* 2002;**240**:481–7.

20 Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma 13. *Acta Ophthalmol Scand* 1998;**76**:268–72.

21 Olsson J, Bengtsson B, Heijl A, *et al.* An improved method to estimate frequency of false positive answers in computerized perimetry. *Acta Ophthalmol Scand* 1997;**75**:181–3.

22 Wall M, Punke SG, Stickney TL, *et al.* SITA standard in optic neuropathies and hemianopias: a comparison with full threshold testing. *Invest Ophthalmol Vis Sci* 2001;**42**:528–37.

23 Shirato S, Inoue R, Fukushima K, *et al.* Clinical evaluation of SITA: a new family of perimetric testing strategies. *Graefes Arch Clin Exp Ophthalmol* 1999;**237**:29–34.

24 Nordmann JP, Brion F, Hamard P, *et al.* [Evaluation of the Humphrey perimetry programs SITA Standard and SITA Fast in normal probands and patients with glaucoma]. *J Fr Ophtalmol* 1998;**21**:549–54.

25 Weber J, Klimaschka T. Test time and efficiency of dynamic strategy in glaucoma perimetry. *German J Ophthalmol* 1995;**4**:25–31.

26 Maeda H, Nakaura M, Negi A. New perimetric threshold test algorithm with dynamic strategy and tendency oriented perimetry (TOP) in glaucomatous eyes. *Eye (Lond)* 2000;**14 Pt 5**:747–51.

27 Zulauf M, Fehlmann P, Flammer J. Perimetry with normal Octopus technique and Weber 'dynamic' technique. Initial results with reference to reproducibility of measurements in glaucoma patients. *Ophthalmologe* 1996;**93**:420–7.

28 Wabbels BK, Wilscher S. Feasibility and outcome of automated static perimetry in children using continuous light increment perimetry (CLIP) and fast threshold perimetry. *Acta Ophthalmol Scand* 2005;**83**:664–9.

29 Wabbels BK, Diehm S, Kolling G. Continuous light increment perimetry compared to full threshold strategy in glaucoma. *Eur J Ophthalmol* 2005;**15**:722–9.

30 Turpin A, Jankovic D, McKendrick AM. Retesting visual fields: utilizing prior information to decrease test-retest variability in glaucoma. *Invest Ophthalmol Vis Sci* 2007;**48**:1627–34.

31 Lang GK. *Augenheilkunde : verstehen - lernen - anwenden ; 50 Tabellen*. Stuttgart; New York: Thieme 2004.

32 Foster PJ, Buhrmann R, Quigley HA, *et al.* The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol* 2002;**86**:238 –242.

33 Nouri-Mahdavi K, Nassiri N, Giangiacomo A, *et al.* Detection of visual field progression in glaucoma with standard achromatic perimetry: a review and practical implications. *Graefes Arch Clin Exp Ophthalmol* 2011;**249**:1593–616.

34 Lachenmayr BJ, Vivell PMO. *Perimetrie*. Stuttgart: Thieme 1992.

35 Aulhorn E, Karmeyer H. Frequency distribution in early glaucomatous visual field defects. *Docum Ophthal Proc Series* 1977;**14**:75–83.

36 Hayreh SS, Zimmerman MB. Nonarteritic anterior ischemic optic neuropathy: natural history of visual outcome. *Ophthalmology* 2008;**115**:298–305.e2.

37 Lowenstein O, Feinberg R, Loewenfeld IE. Pupillary movements during acute and chronic fatigue. *Invest Ophthalmol Vis Sci* 1963;**2**:138–57.

38 Haag-Streit AG 2008-04-09. Octopus 900 - User manual.

39 Hermann A, Paetzold J, Vonthein R, *et al.* Age-dependent normative values for differential luminance sensitivity in automated static perimetry using the Octopus 101. *Acta Ophthalmol* 2008;**86**:446–55.

40 Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;**8**:135–60.

41 Grouven U, Bender R, Ziegler A, *et al.* Vergleich von Messmethoden. *DMW - Deutsche Medizinische Wochenschrift* 2007;**132**:e69–e73.

42 Luithardt AF, Meisner C, Monhart M, *et al.* Validation of a new static perimetric thresholding strategy (GATE). *Br J Ophthalmol* 2015;**99**:11–5.

43 Wild JM, Pacey IE, Hancock SA, *et al.* Between-algorithm, between-individual differences in normal perimetric sensitivity: full threshold, FASTPAC, and SITA. Swedish Interactive Threshold algorithm. *Invest Ophthalmol Vis Sci* 1999;**40**:1152–61.

44 Schiefer U, Pätzold J, Wabbels B, *et al.* Konventionelle Perimetrie - Teil 4: Statische Perimetrie: Befundauswertung - Indizes - Verlaufskontrolle - Perimetrie im Kindesalter. *Ophthalmologe* 2006;**103**:235–54.

45 Anderson DR. Collaborative normal tension glaucoma study. *Curr Opin Ophthalmol* 2003;**14**:86–90.

46 Arnalich-Montiel F, Casas-Llera P, Muñoz-Negrete FJ, *et al.* Performance of glaucoma progression analysis software in a glaucoma population. *Graefes Arch Clin Exp Ophthalmol* 2009;**247**:391–7.

47 Anderson DR, Chauhan B, Johnson C, *et al.* Criteria for progression of glaucoma in clinical management and in outcome studies. *Am J Ophthalmol* 2000;**130**:827–9.

48 Douglas E, *et al.* (The Advanced Glaucoma Intervention Study Investigators). Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. *Ophthalmology* 1994;**101**:1445–55.

49 Flammer J, Drance SM, Zulauf M. Differential light threshold. Short- and long-term fluctuation in patients with glaucoma, normal controls, and patients with suspected glaucoma. *Arch Ophthalmol* 1984;**102**:704–6.

50 Katz J, Sommer A. Asymmetry and variation in the normal hill of vision. *Arch Ophthalmol* 1986;**104**:65–8.

51 Heijl A, Lindgren G, Olsson J. Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol* 1987;**105**:1544–9.

52 Lewis RA, Johnson CA, Keltner JL, *et al.* Variability of quantitative automated perimetry in normal observers. *Ophthalmology* 1986;**93**:878–81.

53 Zulauf M, Flammer J, LeBlanc RP. Normal visual fields measured with Octopus programm G1. I. Differential light sensitivity at individual test locations. *Graefes Arch Clin Exp Ophthalmol* 1994;**232**:509–15.

54 Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. *Am J Ophthalmol* 1989;**108**:130–5.

55 Werner EB, Drance SM. Increased scatter of responses as a precursor of visual field changes in glaucoma. *Can J Ophthalmol* 1977;**12**:140–2.

56 Werner EB, Drance SM. Early visual field disturbances in glaucoma. *Arch Ophthalmol* 1977;**95**:1173–5.

57 Artes PH, Iwase A, Ohno Y, *et al.* Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. *Invest Ophthalmol Vis Sci* 2002;**43**:2654–9.

58 Bjerre A, Grigg JR, Parry NR, *et al.* Test-Retest Variability of Multifocal Visual Evoked Potential and SITA Standard Perimetry in Glaucoma. *Invest Ophthalmol Vis Sci* 2004;**45**:4035–40.

59 Henson DB, Chaudry S, Artes PH, *et al.* Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci* 2000;**41**:417–21.

60 Bengtsson B, Heijl A. SITA fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand* 1998;**76**:431–7.

61 Flammer J. Automatische Perimetrie -Theoretische Grundlagen. In: *Perimetrie mit besonderer Berücksichtigung der Automatischen Perimetrie*. Stuttgart: Enke 1993. 34–59.

62 Enoch JM, Aulhorn E, Dubois-Poulsen A, *et al. Perimetric standards and perimetric glossary of the international council of ophthalmology*. The Hague: Dr. W. Junk Publishers 1979.

63 Fankhauser F, Haeberlin H. Dynamic range and stray light. An estimate of the falsifying effects of stray light in perimetry. *Doc Ophthalmol* 1980;**50**:143–67.

64 Wilhelm H, Neitzel J, Wilhelm B, *et al.* Pupil perimetry using M-sequence stimulation technique. *Invest Ophthalmol Vis Sci* 2000;**41**:1229–38.

65 Webster AR, Luff AJ, Canning CR, *et al.* The effect of pilocarpine on the glaucomatous visual field. *Br J Ophthalmol* 1993;**77**:721–5.

66 Lachenmayer B, Wilhelm H *et al. Fahreignungsbegutachtung für den Straßenverkehr 2013 Empfehlung der Deutschen Ophthalmologischen Gesellschaft (DOG) und des Berufsverbandes der Augenärzte Deutschlands (BVA)*. 6. Auflage. BVA (Berufsverband für Augenärzte) 2013.

67 Taravati P, Brito C, Woodward K, *et al.* The Effect of Stimulus Size on Perimetric Variability. *Invest Ophthalmol Vis Sci* 2005;**46**:3718.

68 Wall M, Kutzko KE, Chauhan BC. Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. *Invest Ophthalmol Vis Sci* 1997;**38**:426–35.

69 Wall M, Doyle CK, Zamba KD, *et al.* The Repeatability of Mean Defect with Size III and Size V Standard Automated Perimetry. *Invest Ophthalmol Vis Sci* 2013;**54**:1345–51.

70 Swanson WH, Felius J, Birch DG. Effect of stimulus size on static visual fields in patients with retinitis pigmentosa. *Ophthalmology* 2000;**107**:1950–4.

71 Pennebaker GE, Stewart WC, Stewart JA, *et al.* The effect of stimulus duration upon the components of fluctuation in static automated perimetry. *Eye* 1992;**6(Pt 4)**:353–5.

72 Rauscher S, Sadowski B, Vonthein R, *et al.* Assessment of reaction times in order to enhance quality of semi-automated kinetic perimetry (SKP) - an age-related normative study. In: *Perimetry Update 2002/2003*. The Hague, The Netherlands: Kugler Publications 2003. 353–8.

73 Lewald J. Passive Auditory Stimulation Improves Vision in Hemianopia. *PLoS ONE* 2012;**7**.

74 Passamonti C, Bertini C, Làdavas E. Audio-visual stimulation improves oculomotor patterns in patients with hemianopia. *Neuropsychologia* 2009;**47**:546–55.

75 Romei V, De Haas B, Mok RM, *et al.* Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Front Psychol* 2011;**2**:215.

76 Khawla Abu Samra. The eye and visual system in pregnancy, what to expect? An in-depth review. *Oman J Ophthalmol 2013 May-Aug;6(2):87–91*.

77 Bosworth CF, Sample PA, Johnson CA, *et al.* Current practice with standard automated perimetry. *Semin Ophthalmol* 2000;**15**:172–81.

78 Lindenmuth KA, Skuta GL, Rabbani R, *et al.* Effects of pupillary dilation on automated perimetry in normal patients. *Ophthalmology* 1990;**97**:367–70.

79 Henson DB, Emuh T. Monitoring vigilance during perimetry by using pupillography. *Invest Ophthalmol Vis Sci* 2010;**51**:3540–3.

80 Bengtsson B, Heijl A. False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? *Invest Ophthalmol Vis Sci* 2000;**41**:2201–4.

81 Bengtsson B. Reliability of computerized perimetric threshold tests as assessed by reliability indices and threshold reproducibility in patients with suspect and manifest glaucoma. *Acta Ophthalmologica Scandinavica* 2000;**78**:519–22.

82 Wood JM, Wild JM, Hussey MK, *et al.* Serial examination of the normal visual field using Octopus automated projection perimetry. Evidence for a learning effect. *Acta Ophthalmol Copenh* 1987;**65**:326–33.

83 Heijl A, Lindgren G, Olsson J. The effect of perimetric experience in normal subjects. *Arch Ophthalmol* 1989;**107**:81–6.

84 Heijl A, Bengtsson B. The effect of perimetric experience in patients with glaucoma. *Arch Ophthalmol* 1996;**114**:19–22.

85 Hudson C, Wild JM, O'Neill EC. Fatigue effects during a single session of automated static threshold perimetry. *Invest Ophthalmol Vis Sci* 1994;**35**:268–80.

86 Johnson CA, Adams CW, Lewis RA. Fatigue effects in automated perimetry. *Appl Opt* 1988;**27**:1030–7.

87 Gonzalez de la Rosa M, Pareja A. Influence of the 'fatigue effect' on the mean deviation measurement in perimetry. *Eur J Ophthalmol* 1997;**7**:29–34.

88 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.

7. Erklärung zum Eigenanteil

Die Konzeption der Studie erfolgte durch Prof. Dr. med. Ulrich Schiefer, Leiter der Arbeitsgruppe Sehbahn, Universitätsaugenklinik Tübingen, und mich, Annette Luithardt, in Zusammenarbeit mit Herrn Dr. biol. hum. M.A. Christoph Meisner, Institut für medizinische Biometrie, Tübingen.

Sämtliche Versuche wurden nach Einarbeitung durch Elke Krapp, damalige Mitarbeiterin der Arbeitsgruppe Sehbahn, eigenständig von mir durchgeführt.

Die Octopus 900 Perimeter und die untersuchte EyeSuite-Software wurden von der Firma Haag-Streit AG, Köniz, Schweiz, gestellt. Bei Fragen zu Soft- und Hardware wandte ich mich an die zuständigen Mitarbeiter der Firma, Mathias Monhart und Iwan Eicher.

Die statistische Auswertung erfolgte nach mehrfacher Beratung durch Herrn Dr. biol. hum. M.A. Christoph Meisner vom Institut für Biometrie selbstständig durch mich.

Hilfe bei der Interpretation der Ergebnisse erhielt ich von Prof. Dr. med. Ulrich Schiefer und Dr. biol. hum. M.A. Christoph Meisner.

Ich versichere, das Manuskript selbstständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

8. Publication

Results of this study have been published in part in the British Journal of Ophthalmology in 2015.

Luithardt AF, Meisner C, Monhart M, *et al.* Validation of a new static perimetric thresholding strategy (GATE). *Br J Ophthalmol* 2015;**99**:11–5.[42]

9. Danksagung

Zunächst möchte ich meinen Dank gegenüber meinem Doktorvater und Betreuer, Prof. Ulrich Schiefer, aussprechen für seine freundliche, aber beharrliche Motivation, Begleitung und Korrektur dieser Arbeit. Danke für alle Hilfe und Unterstützung!
Außerdem möchte ich Herrn Dr. biol. hum. M.A. Christoph Meisner, Institut für medizinische Biometrie, Tübingen, für seine sehr hilfreichen und aufschlussreichen Ratschläge bezüglich der statistischen Analyse danken und für seine Freundlichkeit und Geduld, mir statistische Zusammenhänge begreifbar zu machen. Bereits vor der ersten Untersuchung beeinflusste er die Studie wegweisend mit seinen Ratschlägen.
Darüber hinaus gilt mein Dank den früheren und aktuellen Mitgliedern der Arbeitsgruppe Sehbahn (Elke Krapp, Janko Dietzsch, Andrea Mast, Regine Grund) für ihre Hilfe und Unterstützung bei der praktischen Durchführung der Untersuchungen, der Literaturrecherche und für ihre Worte der Motivation und des Rates.
Für die Beantwortung von Fragen bezüglich der Perimeter und der EyeSuite Software danke ich Mathias Monhart und Iwan Eicher.
Ein besonderer Dank geht an die Probanden dieser Studie. Ihre Bereitschaft, die Wissenschaft und mich durch die Teilnahme an dieser Studie zu unterstützen, hat mich sehr begeistert und motiviert.

Dr. med. Sarah Ziefle und Jonathan Sunkersing danke ich für das Durchlesen und sprachliche Korrigieren dieser Arbeit im Vorfeld.
Zum Abschluss möchte ich noch meinem Ehemann, meiner Familie und meinen Freunden danken, die mich während der Arbeit an dieser Dissertation immer wieder ermutigt und ertragen haben.

10. Appendix

Table 16: Elevated catch trial rates

Catch trials	Rate	Patient ID	Group of patients
False-negative	31%	2205	Glaucoma
False-negative	40%	2211	Glaucoma
False-negative	33%	2212	Glaucoma
False-negative	33%	2212	Glaucoma
False-negative	36%	2216	(post-) chiasmal pathway lesions
False-positive	33%	2216	(post-) chiasmal pathway lesions
False-negative	40%	2217	Glaucoma
False-negative	33%	2223	Glaucoma

Table 17: Location IDs grid 30A

Grid 30A				
Location ID	X - Coordinate	Y - Coordinate	Eccentricity [°]	Angle [°]
0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0
2	1.4	1.4	2.0	45.0
3	-1.4	1.4	2.0	135.0
4	-1.4	-1.4	2.0	-135.0
5	1.4	-1.4	2.0	-45.0
6	5.8	1.6	6.0	15.4
7	4.2	4.2	5.9	45.0
8	-4.2	4.2	5.9	135.0
9	-4.2	-4.2	5.9	-135.0
10	4.2	-4.2	5.9	-45.0
11	5.8	-1.6	6.0	-15.4
12	2.0	5.8	6.1	71.0
13	-2.0	5.8	6.1	109.0
14	-5.8	2.0	6.1	161.0
15	-5.8	-2.0	6.1	-161.0
16	-2.0	-5.8	6.1	-109.0
17	2.0	-5.8	6.1	-71.0
18	2.0	9.7	9.9	78.3
19	-2.0	9.7	9.9	101.7
20	-9.7	2.0	9.9	168.3
21	-9.7	-2.0	9.9	-168.3
22	-2.0	-9.7	9.9	-101.7
23	2.0	-9.7	9.9	-78.3
24	9.7	2.6	10.0	15.0
25	7.1	7.1	10.0	45.0
26	7.1	7.1	10.0	45.0
27	-7.1	7.1	10.0	135.0
28	-7.1	7.1	10.0	135.0
29	-7.1	-7.1	10.0	-135.0
30	-7.1	-7.1	10.0	-135.0
31	7.1	-7.1	10.0	-45.0

32	7.1	-7.1	10.0	-45.0
33	9.7	-2.6	10.0	-15.0
34	-10.8	6.2	12.5	150.1
35	-10.8	-6.3	12.5	-149.7
36	1.5	0.0	15.0	0.0
37	13.0	7.5	15.0	30.0
38	10.6	10.6	15.0	45.0
39	3.9	14.5	15.0	74.9
40	-3.9	14.5	15.0	105.1
41	-10.6	10.6	15.0	135.0
42	-10.6	-10.6	15.0	-135.0
43	-3.9	-14.5	15.0	-105.1
44	3.9	-14.5	15.0	-74.9
45	10.6	-10.6	15.0	-45.0
46	13.0	-7.5	15.0	-30.0
47	17.4	1.5	17.5	4.9
48	-15.2	8.7	17.5	150.2
49	-17.4	2.0	17.5	173.4
50	-17.4	-2.0	17.5	-173.4
51	-15.2	-8.8	17.6	-149.9
52	17.4	-1.5	17.5	-4.9
53	5.2	19.3	20.0	74.9
54	-5.2	19.3	20.0	105.1
55	-5.2	-19.3	20.0	-105.1
56	5.2	-19.3	20.0	-74.9
57	19.3	-5.2	20.0	-15.1
58	19.5	11.2	22.5	29.9
59	11.3	19.5	22.5	59.9
60	2.0	22.4	22.5	84.9
61	-2.0	22.4	22.5	95.1
62	-11.3	19.5	22.5	120.1
63	-22.4	2.0	22.5	174.9
64	-22.4	-2.0	22.5	-174.9
65	-11.3	-19.5	22.5	-120.1
66	-2.0	-22.4	22.5	-95.1
67	2.0	-22.4	22.5	-84.9
68	11.3	-19.5	22.5	-59.9
69	19.5	-11.3	22.5	-30.1
70	24.1	6.5	25.0	15.1
71	17.7	17.7	25.0	45.0
72	-17.7	17.7	25.0	135.0
73	-24.1	6.5	25.0	164.9
74	-24.1	-6.5	25.0	-164.9
75	-17.7	-17.7	25.0	-135.0
76	17.7	-17.7	25.0	-45.0
77	24.1	-6.5	25.0	-15.1
78	-23.8	13.7	27.5	150.1
79	-23.8	-13.8	27.5	-149.9
80	29.8	2.6	30.0	5.0
81	25.9	15.0	30.0	30.0
82	-29.9	2.0	30.0	176.2

83	-29.9	-2.0	30.0	-176.2
84	-2.0	-29.9	30.0	-93.8
85	2.0	-29.9	30.0	-86.2
86	25.9	-15.0	30.0	-30.0
87	29.9	-2.6	30.0	-5.0

Table 18: Location IDs grid 84NO

Grid 84NO				
Location ID	X - Coordinate	Y - Coordinate	Eccentricity [°]	Angle [°]
0	0.0	0.0	0.0	0.0
1	1.4	1.4	2.0	45.0
2	-1.4	1.4	2.0	135.0
3	-1.4	-1.4	2.0	-135.0
4	1.4	-1.4	2.0	-45.0
5	4.2	4.2	5.9	45.0
6	-4.2	4.2	5.9	135.0
7	-4.2	-4.2	5.9	-135.0
8	4.2	-4.2	5.9	-45.0
9	2.0	9.7	9.9	78.3
10	-2.0	9.7	9.9	101.7
11	-9.7	2.0	9.9	168.3
12	-9.7	-2.0	9.9	-168.3
13	-2.0	-9.7	9.9	-101.7
14	2.0	-9.7	9.9	-78.3
15	9.7	2.6	10.0	15.0
16	7.1	7.1	10.0	45.0
17	-7.1	7.1	10.0	135.0
18	-7.1	-7.1	10.0	-135.0
19	7.1	-7.1	10.0	-45.0
20	9.7	-2.6	10.0	-15.0
21	13.9	-1.2	14.0	-4.9
22	15.0	0.0	15.0	0.0
23	14.5	-3.9	15.0	-15.1
24	15.9	-1.4	16.0	-5.0
25	12.4	12.4	17.5	45.0
26	-12.4	12.4	17.5	135.0
27	-12.4	-12.4	17.5	-135.0
28	12.4	-12.4	17.5	-45.0
29	2.0	19.3	19.4	84.1
30	-2.0	19.3	19.4	95.9
31	-2.0	-19.3	19.4	-95.9
32	2.0	-19.3	19.4	-84.1
33	19.3	5.2	20.0	15.1
34	-19.3	5.2	20.0	164.9
35	-19.3	-5.2	20.0	-164.9
36	19.3	-5.2	20.0	-15.1
37	17.7	17.7	25.0	45.0
38	-17.7	17.7	25.0	135.0
39	-17.7	-17.7	25.0	-135.0
40	17.7	-17.7	25.0	-45.0

41	29.8	2.6	30.0	5.0
42	25.9	15.0	30.0	30.0
43	15.0	25.9	30.0	60.0
44	2.0	29.9	30.0	86.2
45	-2.0	29.9	30.0	93.8
46	-15.0	25.9	30.0	120.0
47	-25.9	15.0	30.0	150.0
48	-29.9	2.0	30.0	176.2
49	-29.9	-2.0	30.0	-176.2
50	-25.9	-15.0	30.0	-150.0
51	-15.0	-25.9	30.0	-120.0
52	-2.0	-29.9	30.0	-93.8
53	2.0	-29.9	30.0	-86.2
54	15.0	-25.9	30.0	-60.0
55	25.9	-15.0	30.0	-30.0
56	29.8	-2.6	30.0	-5.0
57	-39.8	2.0	39.9	177.1
58	-39.8	-2.0	39.9	-177.1
59	39.8	3.5	40.0	5.0
60	38.6	10.4	40.0	15.1
61	28.3	28.3	40.0	45.0
62	10.4	38.6	40.0	74.9
63	-10.4	38.6	40.0	105.1
64	-28.3	28.3	40.0	135.0
65	-38.6	10.4	40.0	164.9
66	-38.6	-10.4	40.0	-164.9
67	-28.3	-28.3	40.0	-135.0
68	-10.4	-38.6	40.0	-105.1
69	10.4	-38.6	40.0	-74.9
70	28.3	-28.3	40.0	-45.0
71	38.6	-10.4	40.0	-15.1
72	39.8	-3.5	40.0	-5.0
73	2.0	49.8	49.8	87.7
74	-2.0	49.8	49.8	92.3
75	-49.8	2.0	49.8	177.7
76	-49.8	-2.0	49.8	-177.7
77	-2.0	-49.8	49.8	-92.3
78	2.0	-49.8	49.8	-87.7
79	49.8	4.4	50.0	5.0
80	43.3	25.0	50.0	30.0
81	25.0	43.3	50.0	60.0
82	-25.0	43.3	50.0	120.0
83	-43.3	25.0	50.0	150.0
84	-43.3	-25.0	50.0	-150.0
85	-25.0	-43.3	50.0	-120.0
86	25.0	-43.3	50.0	-60.0
87	43.3	-25.0	50.0	-30.0
88	49.8	-4.4	50.0	-5.0
89	58.0	15.5	50.0	15.0
90	42.4	42.4	60.0	45.0
91	-58.0	15.5	60.0	165.0

92	-58.0	-15.5	60,0	-165.0
93	-15.5	-58.0	60,0	-105.0
94	15.5	-58.0	60,0	-75.0
95	42.4	-42.4	60,0	-45.0
96	58.0	-15.5	60,0	-15.0
97	-2.0	-64.8	64,8	-91.8
98	2.0	-64.8	64,8	-88.2
99	69.7	6.1	70,0	5.0
100	60.6	35.0	70,0	30.0
101	35.0	-60.6	70,0	-60.0
102	60.6	-35.0	70,0	-30.0
103	69.7	-6.1	70,0	-5.0
104	77.3	20.7	80,0	15.0
105	56.6	-56.6	80,0	-45.0
106	77.3	-20.7	80,0	-15.0
107	84.0	7.8	84,0	5.0
108	84.0	-7.8	84,0	-5.0