STRUCTURAL CHANGES IN THE ANTERIOR VISUAL PATHWAY AND ITS CORRELATION WITH FUNCTIONAL VISUAL OUTCOME

A Case-Control study of children with Optic Pathway Gliomas

Bethany E. Higgins BSc
RADIOLOGICAL SCIENCES, UNIVERSITY OF NOTTINGHAM
ID: 4276267

A Dissertation presented in consideration for the degree of Masters in Research in Medicine

Submitted: 17th November 2017

1.0 Abstract

1.1 Background

Optic Pathway Gliomas (OPGs) are a type of benign tumour found in paediatric cohorts that are typically located in the anterior visual pathway and can result in irreversible visual impairment. Suffering from low vision has a significant negative impact on physical and emotional health, leaving many patients vulnerable and dependent on others (1). Research has been conducted that links structural changes in optic nerves to decreased levels of visual function in diseases such as multiple sclerosis (2,3), but to the author's knowledge, no investigation has been completed into the correlation of optic nerve size and visual function in children with OPGs.

1.2 Methods

A systematic literature review was conducted to investigate what method of assessing optic nerve size was the most reliable. This highlighted the need for three distinct, observational studies. The first retrospective, observational experiment investigated if cross-sectional area assessment of optic nerve size was reliable in measuring the optic nerves of 22 children with OPGs (aged 2-15 years). Optic nerve size was assessed from retrospective, structural MRI scans. The second observational experiment analysed the variation of optic nerve sizes in 30 healthy children (aged 5-16 years) in relation to their age, visual acuity and height. Optic nerve size was assessed from structural MRI scans and visual acuity scores were collected using the Sheridan Gardiner test. The third case-control, observational experiment investigated the correlation between optic nerve size, average retinal nerve fibre layer (RNFL) thickness and visual acuity in 27 children with OPGs (aged 2-15 years). Moreover, the variation of optic nerve size and visual acuity according to age was compared between children with OPGs and 30 healthy controls (aged 5-16 years). Data from the previous studies were utilised, coupled with retrospective visual function assessments of visual acuity and OCT results.

1.3 Results

A measure of optic nerve size using a cross-sectional area technique was established to be reliable using a coefficient of variation and an intraclass correlation coefficient measure. Mean optic nerve size in the healthy paediatric cohort was 8.798mm² and it was found to negatively correlate with both age (R^2=0.396) and height (R^2=0.393). This is the first recording of maturation in a paediatric cohort demonstrating axonal decline, to the author's knowledge. In children with OPGs, optic nerve

size was found to not correlate with RNFL thickness, age or visual acuity across all data in this dataset. However, the data for the left eyes did exhibit a significant negative correlation between these variables, and both optic nerve size and RNFL thickness were found to significantly predict visual acuity (R^2 = 0.402). RNFL thickness was also confirmed to correlate with visual acuity in both the left (R^2=0.380) and the right eyes (R^2=0.394). When optic nerve size and visual acuity of children with OPGs and the healthy cohort were compared, the OPG group was found to have significantly smaller nerves and poorer visual function. Lastly, despite issues with small sample sizes throughout the studies due to the difficulties in collecting usable MR images, information regarding faster or more motion resistant sequences have been accrued which will guide future paediatric imaging of awake cohorts.

1.4 Conclusion

A measurement method of optic nerve size has been validated as reproducible and has aided the researcher in demonstrating a previously unrecorded, negative correlation between optic nerve size with height and age in a healthy cohort. In the OPG cohort, this measurement method was used to find that optic nerve size was significantly smaller than in the healthy cohort, and that RNFL thickness correlated with visual acuity which was supported by the literature, while optic nerve size did not correlate with visual acuity. Moreover, RNFL thickness and optic nerve size were found to predict visual acuity, with RNFL thickness being the more significant predictor. Although, as this study was limited by a small sample sizes, these provide the preliminary findings on which to base a larger study.

2.0 Acknowledgements

I would like to extend thanks to my supervisors Dr Rob Dineen, Professor David Walker and Dr Caroline Blanchard of the Faculty of Medicine & Health Sciences for their invaluable help with my research. I would also like to thank Mr Shery Thomas of the Ophthalmology department and Dr Andrew Cooper from the Sir Peter Mansfield Imaging Centre for their help and guidance with different factors of my study.

Contents

3.0 Abbreviations

CI-Confidence Intervals

CNS-Central Nervous System

CSA-Cross-sectional Area

CV-Coefficient of Variation

CVFQ-Children's Visual Function Questionnaire

DTI-Diffusion Tensor Imaging

DWI-Diffusion Weighted Imaging

EEG-Electroencephalogram

FLAIR-Fluid Attenuation Inversion Recovery

FSE-Fast Spin Echo

GCL-IPL-Ganglion Cell Layer of the Inner Plexiform Layer

GE-General Anaesthetic

GRE-Gradient echo

ICC-Intraclass Correlation Coefficient

LoA-Limits of Agreement

MDC-Modified Dodge Classification

MF-Magnetic Field

MRI-Magnetic Resonance Imaging

NF1- Neurofibromatosis type-1

OCT-Optical Coherence Tomography

ON-Optic Nerve

ONS-Optic Nerve Size

OPG-Optic Pathway Glioma

PD-Proton Density

ROI-Regions of Interest

RNFL-Retinal Nerve Fibre Layer

RNFLt-Retinal Nerve Fibre Layer thickness

SE-Spin echo

SG-Sheridan Gardiner test

SNR-Signal-to-Noise-Ratio

SI-Signal Intensity

T-Tesla

TI-Inversion Time

TR-Repetition time

VA-Visual Acuity

VEP-Visual Evoked Potentials

VF-Visual Field

VP-Visual Pathway

4.0 List of Tables

5.0 List of Figures

6.0 Background

6.1 Optic Pathway Gliomas

6.1.1 Epidemiology

Gliomas roughly constitute half of all central nervous system (CNS) tumours in paediatric patient cohorts (13). Optic Pathway Gliomas (OPGs) are a subtype of glioma, typically classed as low-grade astrocytic neoplasms with an incidence rate of 4 per 100,000 (14). OPGs form 2-4% of CNS gliomas and their peak prevalence is between 2-8 years (15). OPGs infiltrate the visual pathway (VP) of the brain, which constitutes the optic nerves (ON), the optic chiasm, the optic tracts and the optic radiations [see Figure 6.1].

To classify tumours, the histological characteristics are examined and compared to the World Health Organisation grading criteria. OPGs are categorised as low grade pilocytic astrocytomas, classed as Grade 1 indicating their indolence (16). The natural history of OPGs is not straightforward, varying from progression to long-term stability (17). Additionally, occurrences of complete regression have been documented (18). As OPGs are rare and have a characteristically ambiguous clinical course, defining a standardised assessment protocol is difficult.

OPGs can be split into two categories: Neurofibromatosis type-1 (NF1) associated tumours known as **syndromic OPGs** and non-syndromic **sporadic OPGs**. Syndromic tumours tend to be indolent, while sporadic tumours are more likely to be aggressive, despite both being histologically similar (19). This factor can be attributed to their differing, typical locations in the VP, discussed further in section 6.2.

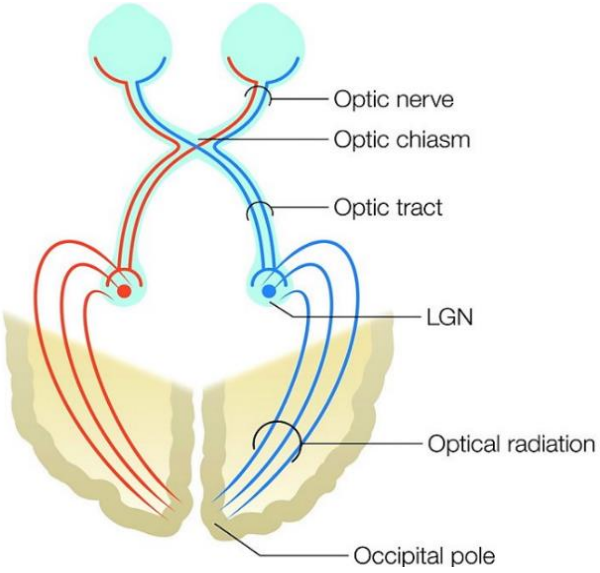

Figure 6.1 *Schematic drawing of the human visual pathway, sourced from Mandelstam* (7)

6.1.2 Histology

Microscopic examination techniques can elucidate the histological characteristics of OPGs, such as bipolar cells with a fibrillary background. The cells demonstrate thin, GFAP-positive processes, Rosenthal fibres, eosinophilic granular bodies and low cellularity (16). In addition, hyalinisation of blood vessels will be frequently evident and the proliferation index will be elevated, despite OPGs being of a low-grade class and increased proliferative activity typically being associated with more aggressive tumour types (20).

The literature on the histological characteristics of OPGs has widened due to advancement in molecular profiling techniques. Investigations utilising whole genome profiling indicate sporadic OPGs appear to exhibit a non-random duplication in the 7q34 region (21). This implicates the BRAF oncogene, typically causing it to fuse with KIAA1549, which subsequently results in the upregulation of the RAS/RAF/MEK/MAPK pathways (22), ultimately increasing risk of tumorigenesis.

Conversely, mutations of the BRAF oncogene are not evident in NF1 associated syndromic OPGs (23). This links to evidence that tumour location impacts the differing profiles of pilocytic astrocytoma gene expression (24). Syndromic tumours are linked to decreased neurofibromin expression, a gene found on chromosome 7 linked to the RAS/RAF/MEK/MAPK pathway, which subsequently links to pilocytic astrocytoma formation (25).

The mTOR pathway also plays a key role in tumorigenesis. Syndromic OPGs are believed to induce hyperactivation of the mTOR pathway through resulting high levels of phosphorylated S6 (19). This postulation has been supported by mouse models of NF1, which have demonstrated an increased percentage of S6 in OPG tissue (26). Sporadic OPGs resulting from BRAF fusion with KIAA1549 have been shown to hyperactivate the mTOR pathway through TSC2 inactivation (19).

Bajenaru *et al.* (27) highlights that the complete inactivation of the neurofibromin 1 gene appears necessary for OPG formation to occur, supported by investigations utilising mouse models of NF1, whereby an NF1 knockout specific to astrocytes was used. However, despite that the mechanisms behind pilocytic astrocytoma formation for both tumour variations being different, both still result in an increase in RAF/RAS signalling which impacts the RAS/RAF/MEK/MAPK pathway, explaining the histological similarities seen between the two tumour types (19).

6.1.3 Anatomic locations of Optic Pathway Gliomas

Weiss *et al.* (28) investigated 16 patients with OPGs, finding 25% occurred within the optic disk or ON, 20-40% occurred within the chiasm and 33-60% extended into the hypothalamus. Twenty years on, Binning *et al*. (29) agrees with Weiss *et al.* (21) to an extent, but cites chiasm involvement is more frequent (40-75%). In addition, there are more recorded occurrences of extension into both the hypothalamic region and the chiasm, opposed to individual involvement (30).

The location of the OPG is critical, influencing the resulting symptoms and ultimately the likelihood of survival, as the location dictates tumour resectability. El Beltagy *et al.* (31) identified OPGs located in the anterior VP tend to cause visual deficits as their main symptom, while if they are located in the hypothalamic region, endocrine function is negatively affected.

OPGs occurring within the ON have limited symptoms compared to other locations, e.g. local visual dysfunction. This is due to compression of the ON, leading to a break down in myelination and nerve atrophy (15). This atrophy is speculated to give rise to visual dysfunction. In addition, Walrath *et al.* (32) report that it is unlikely for an OPG occurring in the ON to extend into the chiasm, further highlighting the indolence of the tumour in this region.

6.1.4 Optic Pathway Gliomas Demography

Demographic factors of OPGs include a higher incidence rate in young females and an earlier symptom presentation age of syndromic tumours in NF1-positive patients (33). This is supported by Helfferich *et al.* (19) who identified syndromic tumours that had a median age of presentation of 4.6 years while sporadic tumours had a median age of 4.8 years. However, Avery *et al.* (34) argue the variation in the literature regarding mean age of symptom onset in syndromic tumours can be attributed to some clinicians performing neuroimaging screening for all children suffering from NF1, regardless of symptom presentation. Therefore, a reported older age of symptom onset can be expected from establishments choosing to not perform neuroimaging screening.

6.1.4.1 Syndromic Optic Pathway Gliomas

NF1 is an autosomal, dominant, neurocutaneous genetic disorder impacting multiple systems in the body and induces the growth of tumours as well as causing other symptoms such as learning disabilities, Lisch nodules, café-au-lait spots on the skin, bony lesions and neurofibromas (29). Approximately 50-60% of patients with OPGs have an underlying diagnosis of NF1 (35). 15-20% of NF1-

positive patients will go onto develop an OPG, indicating a relatively high incidence rate for this population cohort (36). However, it is difficult to identify the incidence rate of OPGs as many seen in NF1 patients are asymptomatic. Although, Fisher *et al.* (36) highlights that NF1-positive females are more likely to become symptomatic and three times more likely to need treatment.

Presence of NF1 in a patient with a diagnosed OPG is critical information to clinicians as it infers the resulting OPG may not be typically severe. This is supported by the fact that syndromic tumours in NF1-positive patients have a predilection to involve the ON rather than the chiasm or hypothalamus (33), and this indicates a clinical outcome of minimal severity. Moreover, syndromic tumours have been recorded in some cases to demonstrate spontaneous regression (18).

6.1.4.2 Sporadic Optic Pathway Gliomas

As the name suggests, sporadic OPGs can appear randomly in previously healthy paediatric participant cohorts. These tumours are usually diagnosed following patients seeking advice over visual dysfunction. Sporadic tumours tend to have a more aggressive nature as they tend to involve the chiasm and/or the hypothalamus (37). Sporadic tumours are associated with a greater chance of tumour progression (38) and also yield higher levels of decreased visual acuity (VA) (33).

6.2 Optic Pathway Gliomas Manifestations

6.2.1 Visual Manifestations

Visual tasks characterise everyday life, so when they can no longer be completed with relative ease, they have an enormous impact on the wellbeing of the impaired individual. Visual dysfunction is the most frequently exhibited symptom of OPGs. Thiagalingam *et al.* (39) conducted a retrospective analysis and found 59% of patients demonstrated visual dysfunction at the time of diagnosis, highlighting the prevalence of the manifestation. Visual dysfunction tends to occur in children between 1-10 years (34). Campagna *et al.* (40) report that despite treatments resulting in the stabilisation of the size of the OPG, vision can continue to degrade throughout childhood. Although, Chateil *et al.* (41) disputes this is not the case for syndromic tumours as the associated impairment is more likely to remain stable.

A common type of visual dysfunction seen with OPGs is VA loss (40). This can range from minor deterioration to sudden and complete blindness (6). VA is a measure of the spatial resolution of the visual system and takes into account both neural and optical factors of vision. Loss of VA impacts everyday visual activities such as reading and writing and has a huge detrimental effect on the sufferer.

Another form of debilitating visual dysfunction is visual field (VF) loss, whereby a portion of the VF usually visible either can no longer be resolved clearly or can no longer be seen at all. [See Figure 6.2 for a diagram of VF impairments]. The resulting VF defect depends on OPG location. If the OPG is before the optic chiasm, the consequential scotoma will be expected to be central, asymmetrical and unilateral. If the glioma is located on the chiasm, the expected defect produced would be a bitemporal hemianopia (42). Lastly, if the glioma is located after the chiasm, it may result in a homonymous hemianopia. OPGs have been recorded to induce significant losses in the VF despite in some cases, VA remaining normal in the unaffected field (6).

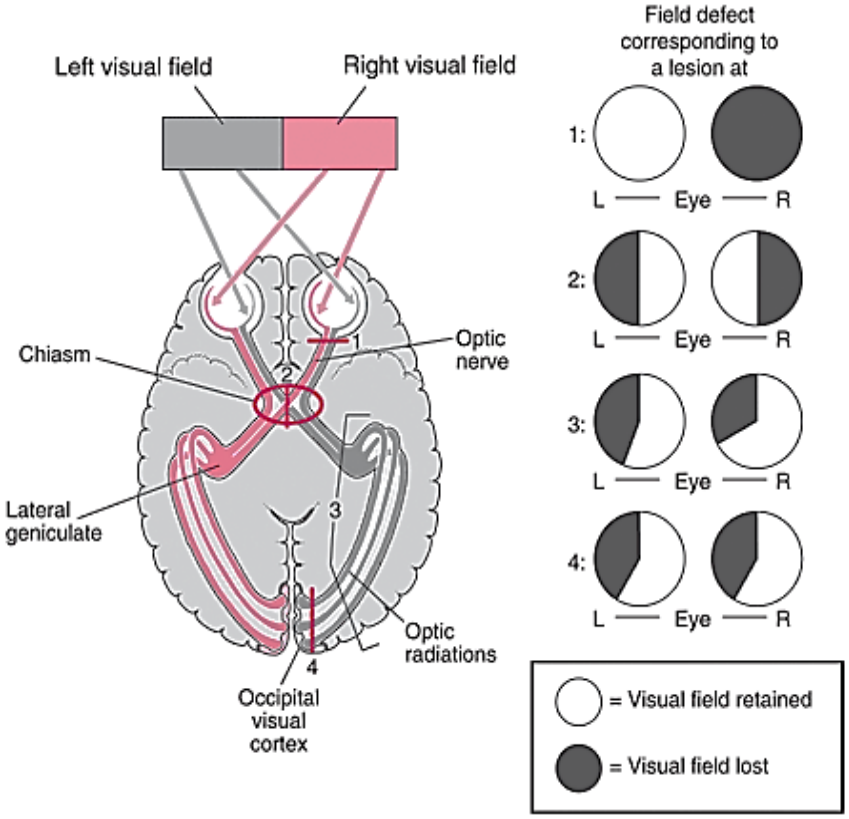

Figure 6.2 *A schematic drawing of visual field deficits, demonstrating the expected defects depending on lesion location sourced from Garrity* (8)

OPGs can also cause nystagmus, whereby the eyes exhibit involuntary movements. The literature describes nystagmus resulting from OPGs can be from side-to-side or rotatory and can be monocular or asymmetrical (43). Other visual manifestations reported include the loss of chromatic vision, relative afferent pupillary defect, proptosis, strabismus, optic disc pallor and in some cases, central retinal vein occlusion (14).

The psychological impact of suffering from low vision has been investigated and the results show the impairment has a significant negative impact on physical and emotional health, leaving many patients vulnerable and dependent on others (1). Moreover, the impact of a child losing functional vision is much more severe, as it can significantly damage their development of social and academic skills (6) and would affect the child's attainment of independence through the development of self-care and mobility skills, increasing dependence on others. Robinson *et al.* (44) identifies the risk of poor vision resulting in poor performance in school, within a cohort already exposed to the possibility of cognitive impairment.

6.2.2 Other Manifestations

OPG symptoms are not exclusively vision related. The gliomas can cause increased intracranial pressure leading to headaches, nausea and papilledema, hydrocephalus (whereby disturbed cerebrospinal fluid flow leads to dilatation of the ventricular system), precocious puberty, diabetes insipidus, diencephalic syndrome causing children to fail to thrive and emaciation, ON tortuosity, endocrine disturbances and increasing head circumference (31,45).

Like the visual features of OPGs, the extent of non-visual features depends upon the tumour type and location. Pre-chiasmatic OPGs on the ON tend to cause visual dysfunction and ON tortuosity, while chiasmatic and post-chiasmatic OPGs can cause hydrocephalus-based symptoms, diencephalic syndrome and endocrine irregularities which can lead to precocious puberty (15). Furthermore, syndromic tumours induce an increased risk of precocious puberty, while sporadic tumours can result in increased intracranial pressure and hydrocephalus symptoms (45).

6.3 Imaging Optic Pathway Gliomas in the Anterior Visual Pathway

6.3.1 MRI: The Basics

This section begins with an overview of MRI image construction and basic physics. MRI is a powerful, non-invasive and non-ionising cross-sectional imaging tool utilising the magnetic properties of tissue within the body. Many types of MRI techniques have been developed that provide different structural and functional information about the tissue being imaged. The MRI output is based upon the behaviour of hydrogen atoms within tissues when placed in the strong magnetic field (MF) and exposed to pulses of radio waves of a specific frequency. The resulting image can reveal integral clinical information regarding tumours and their underlying pathophysiologic mechanisms.

An MRI image is generated through moving polarised atoms acting as tiny magnets, producing an electrical signal in a neighbouring wire coil. MRI systems consist of: a) a strong magnet to produce the main MF, b) shim coils to produce homogeneity in the main MF, c) a radiofrequency coil to emit a measured radio signal into the desired tissue, d) a receiving coil to detect the output signal, e) gradient coils to instigate the spatial localisation of the radiofrequency signals and f) a processing computer to convert the output signals into an image. [See Figure 6.3].

The hydrogen atoms found within the body consist of a charged proton nucleus. The main MF is measured in Tesla (T) and it generates a constant magnetic environment for the protons to exist in equilibrium. Within this constant environment, the protons are pushed into phase together and their

spins align. The general rule is the greater the MF strength, the better the MRI output contrast and signal.

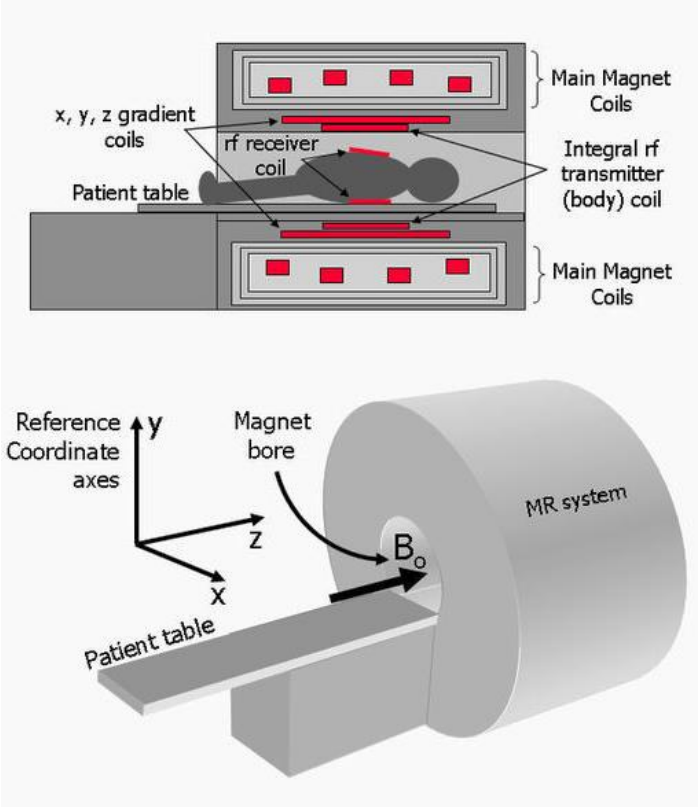

Figure 6.3 *Schematic drawing of an MRI machine, labelled to show the magnet coils in relation to the patient table, sourced from Ridgway* (9)

6.3.1.1 Magnetic Field Strengths

Low field strength magnets vary from 0.15-0.5T. They cost less but have a poorer resolution and limited sequence capability with respect to high field magnets (46). High field strength magnets vary from 1.0-3.0T and are used both clinically and in research. Ultra-high field strengths such as 7.0T upwards are limited to use for research purposes. The advantages of using a magnet with a higher field strength include the faster speed of image acquisition, the increased resolution and the increased signal to noise ratio (SNR).

6.3.1.2 Field of View

The Field of View (FOV) describes the extent of the area the data is collected from (46). The FOV can be limited by factors such as the size of the MRI scanner unit, the study time and the extent of spatial resolution wanted. As a rule, increasing the size of the FOV while everything else remains constant results in an elevated tissue signal but with a reduced spatial resolution.

6.3.1.3 Shim Coils

The main MF must remain constant to retain the equilibrium of the protons. However, sometimes unwanted harmonics can occur. Thus, to ensure field constancy shim coils are used. Shimming acts by generating a mirrored, supplemental MF, produced by passing a current through an active shim gradient. This supplemental field neutralises and therefore cancels out the MF error.

6.3.1.4 Protons and Precession

To produce the signal which initiates the formation of the image, the protons which exist in equilibria must become excited, by introducing a radiofrequency pulse. The protons absorb the energy and their spins shift due to the higher energy change. Once the pulse stops, the spins regulate during a process called precession. The precessional frequency is the rate of precession around the external MF. Protons within different types of body tissues will return to their regular spin at a differing rate. This variation is used to distinguish between varying types of tissue in resulting images. Precession produces an output radiofrequency pulse picked up by receiver coils within the scanner and goes onto formulate the image.

6.3.1.5 Gradient Coils

When forming the resulting MRI image, it is important to produce minor amounts of known variation within the constant, main MF. This gradient is produced by gradient coils. It varies the MF in a linear fashion, so the precessional frequency will subsequently illustrate variation as a function of position of the proton in the *x*, *y* and *z*-axes. This means the signals originating from tissue protons can be spatially localised. Stronger gradients yield rapid imaging times and a higher resolution. Though, there are safety concerns regarding strong gradient coils and excessive nerve stimulation.

6.3.1.6 T1, T2 and Proton Density

The goal of MRI for diagnostic purposes is to demonstrate anatomy and to be able to easily see the differences between normal and abnormal tissues. To achieve this, the image contrast must be altered in the most appropriate way for the tissue type being examined. This reflects the tissues' differing magnetic and physical properties. Signal intensity (SI) refers to the extent of brightness of a certain tissue seen in an MRI image. It is determined by: proton density (PD), T1 relaxation time, T2 relaxation time, and flow. Each individual tissue has a specific PD, T1 and T2 time. When tissues in an MRI image appear bright they are referred to as being hyperintense and have a high SI. On the other hand, when a tissue appears dark on an MRI image, the tissue is described as being hypointense, or has low SI.

T1 and T2 relaxation times refer to the way the protons revert back to their resting states after the radiofrequency pulse. T1 is a measure of the rate of protons returning to equilibrium or b0. T2 is a measure of the rate of protons to losing phase coherence. PD-weighting reflects the number of protons in the tissue. In this case, tissues with the highest concentration of protons will produce to strongest SI. Parameters that influence the relative T1, T2 and PD-weightings of an image acquisition are the repetition time (TR) and echo time (TE).

T1-weighting produces an image in which fat is hyperintense, fluid is hypointense and muscle has an intermediate SI. T2-weighting produces an image in which fat is hyperintense, fluid is also hyperintense and muscle has an intermediate SI. Fast Spin Echo (FSE) is a specific spin echo (SE) technique that allow images to be acquired quickly while retaining image quality.

6.3.1.7 Fat Suppression

Hyperintense fat tissues demonstrate a high SI, with each type of contrast used. This can be counterproductive as lesions tend to also have a raised SI on post-contrast images. It is necessary to be able to differentiate between the two thus, fat suppression is utilised. There are two methods to complete fat suppression: inversion recovery and fat saturation.

Inversion recovery is a commonly used SE sequence. It works by introducing additional radiofrequency pulses to the sequence, flipping the longitudinal magnetisation. Hence, when the SE signal begins, the tissues can be differentiated with regards to their alternate T1 relaxation times, caused by their varying longitudinal magnetisations (47).

Fat saturation is the more widely used method of fat suppression. The fat saturation technique functions by emitting short radiofrequency signals specifically tuned to the resonance frequency of

fat. This occurs immediately before the MRI sequence, meaning the signals from the fat are saturated, thus supressed from the output image.

6.3.1.8 FLAIR

The FLAIR sequence is similar to a T2-weighted image, excluding the fact the TR and TE times are longer. Fluid-attenuated inversion recovery (FLAIR) is utilised to reduce the high intensity signal produced by cerebrospinal fluid (CSF). This is important to use when lesions are covered by a bright CSF signal. The use of FLAIR nullifies the signal, so the CSF subsequently appears dark, meaning the lesion can be differentiated. For fluid suppression, a long inversion time (TI) is usually administered, typically 70% of the T1 value.

6.3.1.9 Enhancement and Contrast Media

Gadolinium-chelate contrast medium is the most commonly used media in MRI. Once injected into the body it enhances lesions, so they can be clearly visualised by clinicians. Gadolinium-chelate contrast agent use is generally considered safe for individuals with normal renal function as the kidneys filter the contrast agent out after 24 hours. When gadolinium is used for MRI of OPGs, the tumour be enhanced, but this can be highly variable. (48). However, it is worth noting both fat and enhancing lesions can produce the same level of brightness on T1 weighted images. Therefore, careful inspection of pre-contrast T1-weighted images or use of fat-saturation sequences is advised (48).

6.3.1.10 Signal to Noise Ratio

SNR describes the magnitude of the signal produced by the desired tissue in comparison to the magnitude of the signal being generated by external causes. The two main sources of noise in MRI are: molecular movement within the body and electrical interreference from the receiver coils, data cables or the other electronic based elements of the MRI measurement system.

The noise itself depends on the type and size of the coils used and the bandwidth of each of the pulse sequences. To reduce the noise and have the highest SNR as possible, the FOV and the slice thickness can be increased meaning a larger volume is being sampled, although there will be a loss of spatial resolution. Another method is to increase the strength of the MF used (within reason). Furthermore, surface coils can be introduced to increase local SI.

6.3.1.11 Spatial Resolution

Spatial resolution defines how sharp an MR image appears, meaning an MR image with low spatial resolution will appear pixelated. Spatial resolution is determined by the size of the imaging voxels defined in the units describing 3D space. Therefore, spatial resolution can be different in three different directions. The size of the voxel, thus the size of the resolution depends upon the size of the MRI matrix, FOV and slice thickness.

6.3.1.12 Imaging Planes

There are three principal imaging planes used in neuroimaging studies: the axial plane, the sagittal plane and the coronal plane. The axial plane will produce transverse images appear like slices of the brain. The sagittal plane will produce images taken perpendicular to the axial plane, separating the right and left hemispheres of the brain giving a lateral view. The coronal plane produces images taken perpendicular to the sagittal plane thus separating the front of the brain from the back, giving a frontal view.

6.3.2 Methods Adapted from MRI

Since the creation of MRI more recently developed methods such as diffusion weighted imaging (DWI), are used to shed light onto the OPG's impact on nerve fibres (49). DWI relies upon the pedesis of water molecules within brain tissues, using their movement in different directions to create contrast on MR images. Diffusion properties of water molecules in the brain can be impacted by temperature, chemical interactions with macromolecules and the held kinetic energy of the water molecules themselves. Hence, as diffusion within the brain is not truly random, it is defined as 'apparent diffusion'.

Diffusion tensor imaging (DTI) evaluates the structural integrity of myelinated nerves in the brain. It is a specialised type of DWI characterising the diffusion of water within the nerves and extracellular spaces. DTI of the anterior VP is demanding as the ON are small and eye movements during the scanning procedure that can result in unusable imaging outputs (50). However, DTI has been successful utilised to identify abnormalities in the anterior VP when OPGs are present (51).

Diffusion tensor tractography (DTT) is based upon DTI technology. DTT allows reconstruction of the white matter pathways and can be used to visualise pathologies that alter white matter tract structure. DTT has the potential to demonstrate the impact of OPG on the white matter tract fibers traversing them. This information could guide clinical decisions to resect lesions or to implement other

treatments (52). Lober *et al.* (53) used DTT to analyse the relationship between OPGs and anterior VP fibers using retrospective data, finding optic fibers either deviate around the lesion or traverse through it. This is important information for clinicians considering resection and can only be demonstrated via the use of DTT. However, DTT is limited by the subjective decisions made by radiologists utilising the technology, as the placement of the ROIs (regions of interest) depends upon the radiologist's experience and understanding of OPGs (51).

6.3.3 Analysing Optic Pathway Gliomas from MRI

To determine the best treatment option for an OPG, factors such as tumour location and size are analysed from MRI images. Ordinarily, these measurements are based upon manual measurements of the tumour. The accuracy of this has been questioned as manual measurements are user-dependent (54). Moreover, manual measurement has been described as being time-consuming and inaccurate, disadvantaged by the inhomogeneous nature of OPGs. This can have detrimental effects on determining disease progression. Recently, automatic delineation of tumour boundaries has been introduced to automise tumour analysis; in order to increase speed of analysis and reduce user-dependent findings. Automatic delineation has been attempted with astrocytomas (55), low-grade gliomas (56) and most recently, OPGs (54).

6.3.4 Optic Pathway Gliomas Classifications

To effectively describe the OPGs for individual cases, the Dodge system was introduced (57). This system identified the location of the OPG with regards to neighbouring structures, indicating it in different stages. The classification was useful as it highlighted patients suitable for resection procedures and those who had a poorer prognosis.

After the introduction of MRI, a modified version of the Dodge classification (MDC) was created (58). MDC allows a detailed description of the OPG with regards to its involvement in locations throughout the anterior VP. Moreover, it adapted the Dodge classification to feature the prospective visual, genetic and surgical risk of the lesion which would aid multidisciplinary paediatric oncology teams in their treatment decisions (59).

The MDC classification has subsequently revealed syndromic tumours are more likely to induce asymmetrical involvement around the chiasm (58). Differentiating between the central or asymmetrical location of an OPG relative to the chiasm is an important factor for clinicians, as it may

indicate the severity of visual dysfunction within the child. Thus, MDC has improved clinician's ability to characterise inhomogeneous OPGs.

However, it has been suggested that the MDC be reviewed to include further clinical factors. Walker *et al*. (59) advises this would aid the standardisation of the surgical assessment of OPGs across the globe and give multidisciplinary paediatric oncology teams a broad overview of the factors of an individual OPG and its prospective impact on the child's quality of life.

6.3.5 MRI in Paediatric Cohorts

Using this technology with a paediatric cohort can be difficult, due to the anatomy of children being smaller than in normative adult populations and head movements during the MRI sequence within the scanner can cause motion artefacts. This can result in the subsequent scan being unusable. However, there are a number of methods utilised to encourage cooperation during MRI.

6.3.5.1 General Anaesthesia

General Anaesthesia (GA) is utilised for all age groups for a variety of procedures, from surgery to dental processes (60). It is particularly useful in MRI as it ensures the child will be still for a long period of time in a loud, confined environment. This ensures cooperation of the child and no motion artefacts, but is associated with potential risks such as anaphylaxis, aspiration and post-procedure nausea. Further disadvantages include the cost associated with GA procedures, including MRI-compatible monitoring and ventilation systems, trained staff, inpatient services and extra time allocated to GA induction and recovery (60). Thus, minimising the need for GA in MRI in a paediatric cohort would be safer and more cost-efficient.

6.3.5.2 Alternative methods

The most apparent alternative would be to administer MRI without sedation, although a method must be found to ensure the diagnostic quality of the resulting scans. For neonates, a dummy and gentle swaddling can also be utilised to reduce movement within a scanner in an effort to encourage sleep. Swaddling can be used in conjunction with feeding to increase sleepiness in the child (60). For older infants and children, sleep manipulation can be used, whereby the children are either deprived of sleep prior to the MRI or encouraged to naturally fall asleep during the scan (61).

The use of mock scanners has increased in paediatric neuroimaging research in an effort to replace the need for GA. Young participants are given the opportunity to 'practice' within a pretend scanner, complete with authentic MRI noises to familiarise the child with the environment within the scanner, in the hope to reduce unsuccessful scanning sessions. De Bie *et al.* (62) confirm the use of mock scanner training prior to the scanning session resulted in 81 out of 90 children yielding good quality structural scans. Moreover, out of 60 children under the age of 7 years, 53 produced diagnostic quality scans. This indicates the use of mock scanners to familiarise children with the MRI scanner environment can benefit clinicians greatly by increasing children's cooperation and understanding of the scanning procedure.

6.4 Clinical Approach to Optic Pathway Glioma Treatment

As OPGs are typically benign tumours, treatment is ordinarily offered to patients who exhibit disease progression. However, it remains debatable what constitutes enough disease progression to warrant treatment. Fisher *et al*. (63) suggests an obvious decline of visual function such as decreasing VA or the swelling or pallor of the optic disc, coupled with visible progression of the OPG evident by the increase in size, enhancement or extension of the tumour, are common factors used to initiate treatments (42). Listernick *et al.* (10) define visual deficit progression as either: a difference of two-lines of a Lea, Snellen or HOTV VA test when compared to the previous evaluation, or a decrease of two octaves in the Teller acuity test.

Guidelines have been established to ensure children receive routine neuro-ophthalmic examinations to track disease progression. The NF1 OPG Task Force recommends recently diagnosed children under 6 years of age receive an ophthalmological evaluation once a year (64). However, there is no consensus over the appropriate duration of ophthalmic screening (65).

6.4.1 Treatment Types

The three major treatment types of OPGs are radiotherapy, chemotherapy and surgery. The routine first treatment response for symptomatic OPGs is chemotherapy, with a mixture of carboplatin and vincristine administered (66). Intravenous chemotherapy is a systemic therapy that involves the intravenous infusion of anticancer drugs which spread throughout the body. The nature of this treatment means it is not localised, and can therefore have effects on other parts of the body. No study has established one chemotherapeutic medium over another for treatment of OPGs. However, some clinicians prefer to administer temozolomide as its side effects are less severe.

Chemotherapy has been established as an effective therapy in a paediatric cohort suffering from progressive OPGs. Mahoney *et al*. (67) identifies that chemotherapy can suspend the need for radiotherapy. However, despite chemotherapy stabilising and reducing the OPG, Janss *et al.* (68) report 60% of children exhibited signs of tumour progression after 5 years post treatment. Hence, children suffering from progressive OPGs may require more chemotherapy, radiotherapy, or both (69). Fisher *et al.* (63) conducted a large multi-centre, retrospective study to investigate the associated visual outcomes with chemotherapy, and found VA improved in 33% and remained stable in 40% of patients when chemotherapy was given. This is supported by Moreno *et al.* (70) who conducted a review and found when chemotherapy was offered, 50% of the subject's retained their VA level, while 14% improved.

Radiotherapy utilises high-energy rays to destroy cancerous cells. It can damage healthy cells but ordinarily these can regenerate. Radiotherapy is an efficient treatment of OPGs. Horwich and Bloom (71) report radiotherapy yields a progression free survival rate of 66-90%, while Combs *et al*. (72) report a progression free survival rate of 71-90%, with visual deficits stabilising at a rate of 69-81%. Moreover, several studies have suggested an early introduction of radiotherapy before evident visual decline is associated with better visual outcomes (72). However, this is contradicted by Wong *et al.* (73) who reports while radiotherapy did reduce failure rate, there was no evidence of survival advantage. In addition, Kelly *et al.* (69) found there was only an improvement of 15%, while there was no change in 70% of patients undergoing radiotherapy. Implying there is no relationship between VA and response to radiological treatment.

Despite it being efficacious, most clinicians avoid radiotherapy due to its harmful side-effects on a young population. It can risk further visual decline, endocrine impacts (74), cerebrovascular disease (75) and neurocognitive deficits (76). Hence, radiotherapy is not regularly given to children under 5 years and is only considered for patients who experience progression following surgery. Kelly *et al.* (69) summarises after 10 years, patients with progressing OPGs (with or in absence of NF1) exhibited more disease-control if they received radiotherapy. However, after 20 years the rates are equivalent for those treated with radiotherapy and those not. Therefore, other treatment options are often utilised.

Surgery would only be appropriate if the OPG had extrinsic components (34). Other reasons necessitating resection include painful proptosis and the exposure of the cornea. Occasionally, debulking surgery is utilised for large chiasmatic and hypothalamic tumours impacting surrounding structures (34). Complete resection is ordinarily only considered for OPGs limited to the ON, as the procedure can be associated with risking blindness in the affected eye (15).

6.5 Measuring Visual Deficits

It is important for visual deficits to be assessed and monitored accurately, particularly in a cohort as susceptible to visual impairments as children with OPGs. Blanchard *et al.* (77) reports at the time of tumour diagnosis, children with OPGs were found to have three major ophthalmological signs: reduced VA (37), optic atrophy or swelling and proptosis. Other signs include VF deficits (78) such as, visual-evoked potential (VEP) abnormalities (10) and irregularities of the retinal fibre layer (14). Finding the most accurate way of assessing these abnormalities across the OPG clinical population would encourage standardisation across clinics and ultimately, the available literature.

6.5.1 Measuring Visual Acuity

Avery *et al.* (34) identify the quantitative assessment of VA as the most important part of an ophthalmological assessment for a child with an OPG. VA has been found to be reproducible in many studies measuring OPG progression (79) and due to its quantitative nature, it allows the measure to clearly illustrate changes in visual function. VA indicates a person's clarity of vision and ability to resolve details in a visual scene. This is dependent on both optical and neural factors. When a child is born, their VA is low, but it improves as they grow. The maturation of a child's visual system between birth and 8 years of age can make longitudinal analysis of VA difficult, but there are general guidelines for what the normal VA range should be. Normal VA in children aged 3 upwards are illustrated in Figure 6.4.

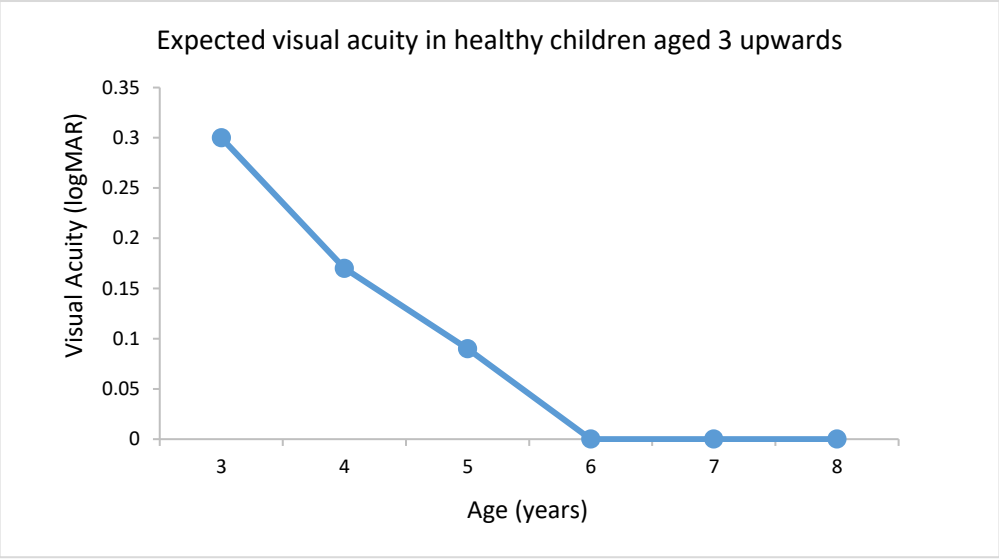

Figure 6.4 *The expected visual acuity values in healthy children, aged ≥3years. Adapted from Listernick et al.* (10)

When assessing VA, patients are ordinarily shown letters or images to identify, known as optotypes (80). Their VA is analysed from the smallest optotype the subject can correctly identify. If a child is unable to report the largest letter of the test, the examiner uses the 'Counting Fingers' test. Routinely, VA testing is completed using high contrast conditions, with black optotypes on a white screen in a low-lit room. VA is assessed using the subject's best corrected vision, using optical correction.

The distance at which the subject performs the test from depends on the type of test chosen and whether the test is measuring near or far vision. This distance is set for each measurement method and must be adhered to, as the size of the optotypes corresponds to the chosen distance. Furthermore, the eyes should be assessed individually (monocular VA opposed to binocular VA), through the occlusion of one of the eyes (80). Light levels should remain constant throughout testing. Clinicians must standardise their chosen to avoid yielding invalid results.

6.5.1.1 Snellen vs LogMAR

VA can be reported in different ways. The two major styles are Snellen scoring and LogMAR scoring. Snellen scoring is reported as a fraction, whereby the numerator indicates the viewing distance in feet from the optotypes being read, while the denominator indicates the distance in feet the average person can see the same optotype. This can be subsequently converted into decimals (10).

LogMAR scoring was introduced to standardise VA testing. It is a linear scale, measured by taking the base 10 logarithm of 1/VA notation, in the logarithm of the minimal angle of resolution (81). Ordinarily, each line on a VA chart is divided into 0.1 LogMAR units, increasing the ease of calculating LogMAR VA scores. Moreover, using the LogMAR notation for VA aids longitudinal data analysis by making statistical analysis easier and avoids the issue of relying on measurement choice between meters and feet. Thus, the LogMAR unit increases the cohesion across low vision clinics and research (82). Table 6.1 indicates the conversion of the two metrics. Low vision is defined by <0.5 logMAR, while blindness is defined by >1.3 logMAR (83).

Table 6.1 *The conversion from Snellen fractions to logMAR units*. *Values above the red line indicate blindness. Values above the orange line indicate low vision. *value for LP and NLP derived from Grover et al.* (156) ***value for HM and CF derived from Schulze-Bonsel et al.* (157)

Snellen VA	LogMAR VA	Low Vision notations
20/6060	2.48	No light perception (NLP)*
20/4000	2.3	Hand Motion (HM)**
20/2500	2.1	
20/2000	2.0	Light perception (LP)*
20/1600	1.9	Counting Fingers (CF) **
20/1300	1.82	
20/1250	1.8	
20/1000	1.7	
20/800	1.6	
20/640	1.52	
20/500	1.4	
20/400	1.3	
20/300	1.2	
20/240	1.08	
20/200	1.0	
20/160	0.9	
20/120	0.78	
20/100	0.7	
20/80	0.6	
20/63	0.5	
20/60	0.48	
20/50	0.4	
20/40	0.3	
20/32	0.2	
20/30	0.18	
20/25	0.1	
20/22	0.04	
20/20	0.0	
20/16	-0.1	
20/15	-0.12	
20/13	-0.18	
20/12.5	-0.2	
20/10	-0.3	

6.5.1.2 Test Types

There are many different types of VA tests, differing in their optotypes, distance and nature of delivery. The type of VA test chosen depends on the age and ability of the participant. For very young children, a preferential looking test such as Teller Acuity test can be used. For pre-literate children, the Lea figure, Kay picture or HOTV matching can be used, which utilise pictures. While for literate children, a Snellen or LogMAR chart such as a Keeler chart, Bailey-Lovie chart or an EDTRS chart can be used, which utilise letters. Some tests such as the Sheridan Gardiner test can be adapted to be used in literate and illiterate children. Each VA test has been shown to demonstrate a high test-retest reproducibility in young children (36,79).

The Response Evaluation in Neurofibromatosis and Schwannomatosis Visual Outcomes Committee have recently confirmed the need for consistent quantitative testing methods to produce the most accurate vision function outcome measures (77). The committee recommended the use of Teller Acuity cards (also known as grating acuity cards) in children with NF1, with HOTV matching used when the child is proficient enough (77).

Teller Acuity Cards have been cited as a reliable measure of VA and are significantly correlated with recognition acuity VA tests (84). However, they have also been reported to overestimate VA. This is because it is a grating acuity test rather than a recognition acuity test like other VA testing methods. Thus, when VA is not in the normal range, it may overestimate the VA level (85), while recognition tests may underestimate VA when vision is considered within the range of normal (86). Therefore, despite the sufficient agreement between the tests, they are not absolutely equivalent (87). This highlights the necessity of a standardised testing between clinics worldwide.

The Sheridan Gardiner test is an example of VA tested for reliability, adapted to suit preliterate and literate children. The child is shown a flipbook of different sized letter optotypes, beginning with the far-vision booklet, and progressing onto the next size appropriate booklet to accurately determine their VA. Like the HOTV and Lea Figure test, the child has a lap card which allows them to point to the corresponding optotype they can see, if they are unable to verbalise it. This makes the Sheridan Gardiner test particularly useful in testing a large cohort of children as it reduces the need to swap between tests for different ages. Ultimately, this is very important as VA tests are not interchangeable and their results are not always comparable, due to differences in test complexity.

Table 6.2 *VA norms for paediatric populations by test type, adapted from Anstice and Thompson (4)*

VA Test Type	Authors	Age-group	LogMAR equivalent score	Snellen equivalent score
Teller Acuity Cards	Mayer *et al.* (88)	1 month	1.5	6/180
		6 months	0.7	6/30
		4 years	0.07	6/7
	Hargadon *et al.* (89)	5-6 years	0.1	6/7.5
Single Lea Symbol	Becker *et al.* (90)	2-6 years	0.07	6/7
Crowded Lea Symbols		2-6 years	0.12	6/8
	Chen *et al.* (79)	4.5-8.5 years	0.08±0.09	6/7.5
Single Kay Pictures	Norgett and Siderov, (11)	4-6 years	-0.15±0.11	6/4
		7-9 years	-0.18±0.13	6/4
Crowded Kay Pictures		4-6 years	-0.1±0.09	6/5
		7-9 years	-0.17±0.11	6/4
	Jones *et al.* (12)	2.5-16 years	-0.04	6/6
Sheridan Gardiner uncrowded	Norgett and Siderov, (11)	4-6 years	-0.18±0.08	6/4
		7-9 years	-0.17±0.14	6/4
Sheridan Gardiner crowded	Simmers, Gray and Spowart, (91)	5-6 years	-0.08±0.09	6/5
HOTV crowded	Drover *et al.* (92)	3-4 years	0.08	6/7.5
		5 years	0.03	6/6
		6 years	-0.03	6/6
	Pan *et al.* (93)	3-4 years	0.17±0.13	6/9
		4-6 years	0.08±0.11	6/7.5
	Birch *et al.* (94)	5.5-12 years	-0.1 to 0.3	6/9
Crowded Keeler logMAR	Jones *et al.* (12)	2.5-16 years	0.04	6/6
	Simmers, Gray and Spowart, (94)	5-6 years	0.1±0.08	6/7.5
	Norgett and Siderov, (11)	4-6 years	0.0±0.08	6/6
		7-9 years	-0.04±0.11	6/6
ETDRS	Birch *et al.* (90)	5.5-12 years	-0.10 to 0.40	6/5

6.5.1.3 Studies Measuring the Variation of Visual Acuity in Healthy Children

It is generally agreed that VA in healthy children increases with age until 6 years and then plateaus (Figure 6.4). Multiple studies have been completed using different VA methods to confirm this and a general consensus has been reached regarding expected levels of VA. Table 6.2 adapted from Anstice and Thompson (4), identifies the VA norms for different tests and their appropriate age-group. The findings agree with Listernick *et al.* (10), despite evidence of slight variation in scores between different visual tests.

Norgett and Siderov, (11) aimed to identify the crowding effect in the following VA tests: Crowded Keeler, Single and Crowded Kay, Sheridan and Gardiner and Sonksen (See Figure 6.5). 103 children were tested and divided into two age groups: 4-6 and 7-9years. A paradigm of testing the left eye was adopted. The older age group performed marginally better than the younger age group, as expected. This supports the literature regarding the maturation of the VP results in better VA (10).

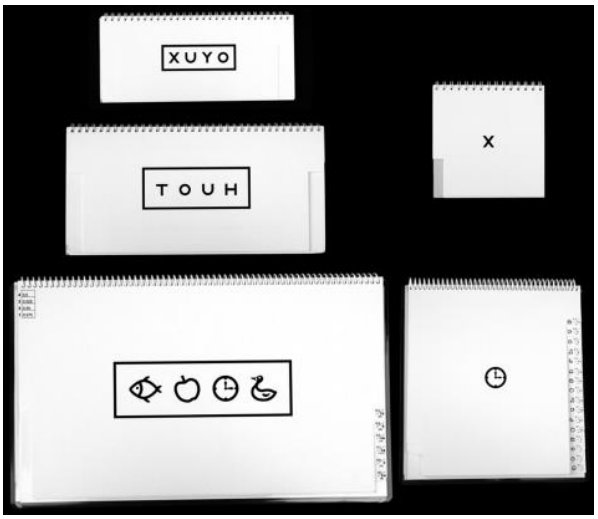

Figure 6.5 *sourced from Norgett and Siderov, (11), clockwise from top: Crowded Keeler, Sheridan Gardiner, Single Kay, Crowded Kay and Sonksen*

Jones *et al.* (12), aimed to compare Kay Picture test and Crowded Keeler logMAR. 103 children (range 2.5-16 years) were tested binocularly. Good agreement was found between the tests using an Intraclass Correlation Coefficient and Bland and Altman Plot [see Figure 6.6]. There was a mean difference of 0.08logMAR, but this is less than one-line difference thus insignificant. The data was divided into affected eye (poorer VA) and unaffected eye (contralateral). The mean VA for the letter

test was 0.12logMAR for the affected eye and 0.04logMAR for the unaffected eye for the Letter Test and 0.039logMAR and -0.04logMAR (respectively) for the Kay Picture Test.

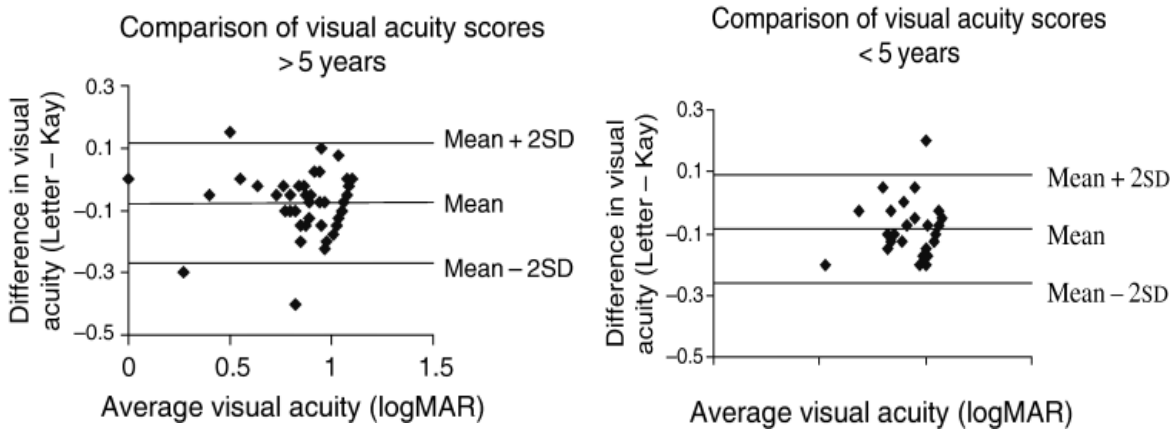

Figure 6.6 *sourced from Jones et al. (12), Bland and Altman plots comparing the Kay Picture Test and the Letter Acuity Test*

Simmers, Gray and Spowart, (94) aimed to compare the Sheridan Gardiner test and the Glasgow acuity cards (now Keeler Crowded). The binocular testing of 702 primary school children with a mean age of 5.4. was conducted, and they found 633 (90.2%) children had a VA better than 0.0 logMAR. The Keeler crowded test was found to be more sensitive than the Sheridan Gardiner test, identifying 100% of amblyopia cases, compared to 55% for the Sheridan Gardner test.

6.5.1.4 Studies Measuring the Variation of Visual Acuity in Children with Optic Pathway Gliomas

There are limited studies looking at the variation of VA in children with OPGs. This is due to VA developing as the child matures and as the tumour changes. Fisher *et al.* (36) conducted a multi-centre, retrospective study to assess the changes in visual function following chemotherapy treatment. 88 patients were evaluated, with a median age at OPG diagnosis of 2.66 years. It was found VA had improved in 32% of patients, remained stable in 40% of patients and declined in 28% of subjects. The methods used to assess VA were Teller, Lea, HOTV and Snellen.

Chang *et al.* (95) conducted a study using sweep Visual Evoked Potentials (VEP) to assess visual function in children with OPGs. Out of 16 children with OPGs, 10 had a VA of 0.0 logMAR recorded from both eyes, 1 had a VA of 0.18 logMAR recorded from both eyes, while the remaining participants had one eye with a LogMAR score of 0.0 and the contralateral eye had a score of 0.1-0.18 logMAR

(95). Although this appears to imply the VA of the children is relatively normal, consider that corrected vision was utilised and when both eyes are measured together rather than individually, it is not apparent if there is a discrepancy between the two which could indicate low visual function.

Parrozzani *et al.* (5) measured VA when assessing the feasibility of using optical coherence tomography to diagnose OPGs. As seen in Table 6.3, reproduced from Parrozzani *et al.* (5), it is clear low levels of VA are evident when it has been possible to assess it. Lea symbols and Snellen charts were used to assess VA. However, the data is limited by the reported limited cooperation of some participants, identified as uninformative when assessment was not possible. Due to the small sample size, it is unclear if this is representative of the OPG population.

Table 6.3 *VA outcomes from Parrozzani et al. (5)*

Patient	Age	VFA Threshold: Lowest Fifth Percentile by Age, logMAR	VFA Result in the Worst Eye	Classification by VFA
3	4	0.26	Uninformative	Uninformative
6	7	0.17	0.1	Normal
11	4	0.26	Uninformative	Uninformative
12	3	0.46	Lack of fixation in RE	Suspected of OPG
14	7	0.17	Uninformative	Uninformative
19	4	0.26	Lack of fixation in LE	Suspected of OPG
26	13	0.17	Hand motion	Suspected of OPG
30	3	0.46	0.2	Normal
31	15	0.17	0.1	Normal
32	10	0.17	0.4	Suspected of OPG
36	10	0.17	Light perception	Suspected of OPG
40	5	0.17	0.1	Normal
46	6	0.17	Hand motion	Suspected of OPG
51	11	0.17	0.1	Normal
56	14	0.17	0.8	Suspected of OPG

During a study looking into vision specific Quality of Life, conducted by Avery and Hardy (6), VA was assessed. Although values were not individually provided, the extent of impairment in 35 participants with OPGs was reported. It is clear 34-37% of assessed individuals exhibited impaired VA levels. This is illustrated by Figure 6.4, sourced from Avery and Hardy (6).

VA scores were also compared the Children's Visual Function Questionnaire (CVFQ), used to measure the children's visual Quality of Life in four main areas: Family Impact, Personality, Competence, and Treatment Difficulty. Children who were found to have mild to severe monocular VA loss had significantly lower ratings on the Competence and Family Impact subscales compared to children with normal VA levels. This finding contradicts Angeles-Han *et al.* (96) study which speculates monocular vision loss does not impact visual quality of life as the fellow seeing-eye counteracts the impairment.

Avery and Hardy (6) findings are supported by numerous studies in the literature also finding monocular vision loss leads to lower quality of life scores (97).

Table 6.4 *VA outcomes, from Avery and Hardy, (6) and key adapted from the report*

	Study subjects
OD category of VA—no. (%)	
Normal	23 (66)
Borderline	4 (11)
Mild impairment	3 (9)
Moderate impairment	1 (3)
Severe impairment	4 (11)
OS category of VA—no. (%)	
Normal	22 (63)
Borderline	4 (11)
Mild impairment	3 (9)
Moderate impairment	3 (9)
Severe impairment	3 (9)

Key Term	logMAR equivalent
Borderline	within 0.1 logMAR of age-based norms
Mildly Impaired	within 0.2–0.3 logMAR of age-based norms
Moderately Impaired	0.4–0.9 logMAR of age-based norms
Severely Impaired	C1.0 logMAR of age-based norms

6.5.2 Measuring the Visual Field

The VF is the area in the visual periphery seen when focussing on a central point. A defect in the VF is defined by Cubbidge (98) as '*a departure from the normal topography of the hill of vision'*. Different types of VF losses can occur, depending on the area of the anterior VP damaged. Figure 6.2 identifies possible VF defects that can occur. The routine VF test used with children is a Kinetic Manual Perimetry test, whereby a clinician sits facing the patient using their eyes as the focus point and uses either an object or their fingers to assess which areas of the VF the patient cannot see (99). Listernick *et al.* (10) suggests VF testing using a computerised method can be used reliably in a paediatric cohort, although notes many tested children had yielded false results during the assessment, implying the test may be unsuitable (56).

Some children with OPGs demonstrate VF deficits. However, this VF loss does not appear to occur in children without the simultaneous experience of clinically significant VA loss (56). Moreover, it is difficult to test for VF deficits in a young cohort and longitudinal, significant changes are hard to quantify due to test variability and lack of the child's cooperation (10). Despite the appeared low confidence in VF testing in children with OPGs, Avery *et al.* (14) recommends VF should be tested during routine ophthalmological assessments. It can be postulated this is due to VA only reflecting

damage in the central two degrees of the VF, while VF measures can indicate damage across the entire FOV.

6.5.3 Measuring Visual Evoked Potentials

Despite VA being a reliable measure of the visual function, Avery *et al*. (34) points out the measure can only detect damage in the central two degrees of the VF, while a more sophisticated measure such as VEP can reveal further damage across the VF. A VEP is an induced coritcal potential caused by a visual stimulus, measured via an electroencephalogram (EEG) recorded at the scalp over the occipital cortex. The EEG is sensitive to electrical signals are recorded at the scalp over the occipital cortex. VEP are regarded as an objective measure of visual function, and have been found to be more sensitive VA in detecting ON impairment, by assessing the reduced amplitudes in children with exisiting VP damage. Moreover, VEP have been found to be more attainable from a paediatric cohort whom are unsuitable for VF testing (69).

However, Kelly *et al.* (69) found there is limited ability to measure progressive visual impairment in a patient with reduced EP, making the measurement unsuitable. This is supported by Listernick *et al.* (10) who reports it is unclear the extent of a change and in what VEP parameters constitutes as a negative decline. Moreover, in an effort to correlate VEP and vision loss, it has been reported VEP deficits in OPG participants overlap with findings in healthy participants (95). Hence, experts in functional outcomes in OPG participants diagree with the technique's use in this cohort.

6.5.4 Measuring Other Visual Deficits

Refractive error can impact a child's ability to perform well in visual tests. Thus, refractive error should always be assessed, and if necessary their vision should be corrected before performing in visual assessments. Types of error include hyperopia, myopia, astigmatism, and anisometropia. Correction is important in children with OPGs as it should be ensured any reports of visual decline should be due to loss of VA attributed to the OPG, and not because the child requires glasses.

Testing colour vision should be considered when a child is old enough to comply, as it aids clinicians by identifying if visual deficits stem from amblyopic vision loss due to optic neuropathy, or due to OPG progression (34). For example, disruption of the functionality of the ON may be characterised by average VA but deficient colour vision, while amblyopia may be characterised by the reverse. In addition, Avery *et al.* (34) adds pupillary reflex, ocular alignment, optic disc pallor and an examination of the fundi should be carried out by practitioners when assessing OPG patients.

6.5.5 Issues with Assessing Visual Deficits

Despite VA testing being reliable in children (100), the success of each method depends on the child's age. This is true of normative groups of children and children with OPGs. For instance, in young children it is not unusual for a repeated test using the same method to yield different results. Thus, the testing relies on the child's willing cooperation, making it a psychophysical measure (101).

This makes testing children with OPGs harder, as the tumour may coincide with cognitive dysfunction, learning difficulties or delays in development (102). Avery *et al.* (80) highlights some children with NF1 may only be able to complete a VA test designed for a younger aged cohort. If the child is receiving chemotherapy, this lack of cooperation may be intensified due to fatigue. However, this is not excluded to children with OPGs and/or NF1, as healthy children of the same age can differ in their emotional control and thus may require different VA testing methods (80).

Despite this, skilled paediatric ophthalmologists can repeatedly obtain reliable visual function measures in children with OPGs. Methods such as incorporating parent's help during the visual test such as aiding occlusion and encouraging parents to mimic visual testing at home can help reduce stress and encourage cooperation. This is in addition to issuing constant reassurance and praise to the child throughout the examination (34).

Another issue with testing VA is the inability to interchange between tests, despite the outcome measure being the same. This stems from the differing complexity of different tests and the sometimes overestimation of VA by some e.g. Teller Acuity Cards (10). For example, tests incorporating a lap card have an increased chance of identifying the correct optotype via guesswork, while in chart-based tests, some optotypes look similar e.g. E and F. Lastly, the use of different equipment can cause VA measures to vary within tests, such as the use of a chart versus a lit monitor screen (80).

Furthermore, practitioners are also faced with the issue of defining what a significant change in VA is, as it is dependent on the participant. Generally, if a child is no longer able to read two lines above what is their norm, this ordinarily is decided as VA decline (10). However, unless more research is conducted on VA outcomes in children with OPG, an absolute definition will not be found.

6.6 Link Between Anterior Visual Pathway Structural Changes and Visual Function

6.6.1 Retinal Nerve Fibre Layer

There is speculation as to whether VA and VF loss in OPG patients are attributed to damage to the axons in the VP. Therefore, structural changes are expected to be correlated to alterations in visual function. This has been demonstrated by using optical coherence tomography (OCT) of the retinal nerve fibre layer (RNFL) (5). The RNFL is an extension of the ON fibers into the globe, while OCT is an imaging facility using near-infrared light to take micrometre resolution images of the biological tissues within the eye. OCT uses concepts of interferometry to enable clinicians to identify the structures within the eye and thus, identify any changes.

OCT technology has improved as the use of differing wavelengths of light was incorporated opposed to using temporal delay to ascertain the spatial location of reflected light. SD-OCT has been demonstrated itself to be a reliable, non-invasive technique, capable of measuring anterior VP integrity. This is true for a variety of diseases including glaucoma, multiple sclerosis, and optic neuropathy (103–106).

Utilising OCT in the OPG cohort may be promising as it could give a reliable structural biomarker of longitudinal visual function (107). OCT is now routinely used in the OPG Clinic run in Queens Medical Centre, Nottingham, to measure longitudinal changes in RNFL thickness (RNFLt) and ON head volume. However, until more research into the longitudinal changes in the OPG population indicated by OCT is investigated, OCT findings will not be relied upon for clinical decisions as of yet.

Avery *et al.* (108) highlight the necessity to evaluate and understand the correlation between visual functional decline and structural malformations in the OPG population, before the modality can be used in clinical decision-making. Research into glaucoma has provided the current information available on the relationship between anterior VP structure and visual function, thus more work is needed for the disease specific relationships such as the impact of benign tumours (109).

Research teams headed by Avery from the Children's National Medical Centre, Washington, have investigated the use of OCT in the OPG cohort, reporting on peripapillary RNFL atrophy. The peripapillary RNFL is the immediate area around the ON head. Using cross-sectional studies, they have found that not only is peripapillary RNFL atrophy linked to the magnitude of vision loss, but it can distinguish between healthy controls and the case group (14,108). For example, Avery *et al.* (108) indicates children experiencing loss in visual function ordinarily illustrate a ≥10% decline of RNFLt. Moreover, the team reported that 56% of their participants with a VA of 0.3 LogMAR showed a deficit

in their RNFL (34). Hence, this measure appears to be applicable for predicting visual dysfunction.

Although, accurately measuring the peripapillary RNFL can be impeded by swelling of the axons. Thus, while the past literature appears to focus on the peripapillary RNFL, more current research is investigating whether the ganglion cell layer (GCL) and inner plexiform layer (IPL) may be used as a visual biomarker of visual function (103,108). The GCL-IPL is found within a packed area of fibrils, which consists of ganglion cells and displaced amacrine cells. It is not impacted by the disadvantages faced by measuring the peripapillary RNFL. Moreover, it is unaffected by refractive error and reports less between-individual variance (103,108).

Gu *et al.* (103) conducted an investigation to see if they could become a feasible biomarker for visual dysfunction in an OPG population. It was demonstrated that GCL-IPL can be used to differentiate between patients with visual dysfunction and patients with normal vision. It has been postulated this measure from OCT is better for the OPG population as the results are not marred by changes in optic nerve size (ONS) that would misleadingly increase peripapillary RNFLt (103). It has been postulated a decrease in GCL-IPL thickness in children who report normative visual function may be used as an indication for earlier treatment. However, the temporal characteristic between atrophy and visual dysfunction is unknown and requires longitudinal investigation to quantify (103,108).

6.6.2 Studies Examining the Variation in Anterior Visual Pathway Structure in an Optic Pathway Gliomas cohort

Minimal studies have been conducted to assess structural changes in the anterior VP in the OPG cohort. Levin *et al.* (110) assessed the correlation between ON tortuosity and ON sheath thickening on the likelihood of developing an OPG. 132 children were evaluated and it was found patients exhibiting tortuosity were significantly more at risk to developing an OPG, but does not indicate if this would lead to an aggressive tumour which impacts on visual function. Tortuosity has been investigated in this cohort previously (111), however these structural changes were not correlated with visual function.

Tractography, a subtype of diffusion weighted imaging which models the structure of the neural tracts, has been utilised in cohorts with OPGs. This method can identify structural changes of neuronal networks in the anterior VP. Lober *et al.* (53) utilised this method and found visual fibres either stopped before the glioma, traversed irregularly or diverged around it. However, these findings were not correlated with visual function.

6.6.3 Optic Nerve Structure and Visual Function

The ON is a cranial nerve connecting the globe to the diencephalon and subsequent visual processing areas, and as such damage to the nerve can result in visual dysfunction. The nerve can be damaged in an OPG population via swelling due to OPG location or ON atrophy. Although it has been reported atrophy of the structures within the eye is correlated with visual deficits, it appears from the literature that no research exists investigating the direct correlation between ONS and visual function in children with OPGs. It is expected that ON atrophy reflects deficits in vision, particularly in a cohort suffering from tumours situated within or around the ON, causing disruption to the regular structure.

Avery *et al.* (112) investigated the volume of the anterior VP in its entirety and found a greater volume including OPG itself correlated with atrophy in the RNFL. However, when measuring the structures of the pathway and the tumour, the study reported that a larger-sized OPG is correlated with more severe visual deficits, which has previously been established. This is supported by Kelly *et al.* (69) who reported progressive visual dysfunction was found to be predicted by larger OPG volume at presentation.

ONS has been correlated to visual dysfunction in other diseases such as: optic neuritis (2,3,113–117), ischemic optic neuropathy (106,117) and increased intracranial pressure (43,118). These three symptoms can result due to glaucoma (106,119–121) or multiple sclerosis (2,3,104,122,123). These studies demonstrate the ONS decreases due to atrophy, and infers it results in diminished visual function. However, when looking at OPGs in a clinical population, there are more factors to consider. For example, if the OPG is located on one ON, would the contralateral ONS be regular or atrophic due to the presence of the tumour?

To investigate the relationship between visual function and ONS in patients with OPG, it is first necessary to identify a reliable method of measuring ONS using MRI.

7.0 Systematic Literature Review

7.1 Abstract

This review aimed to analyse the reliability of manual methods of assessing ONS using MRI. Publications featured a measure of repeatability and/or reproducibility, providing an approved reliability index was reported (intraclass correlation coefficient or the coefficient of variation). PubMed and Web of Science research databases were searched in July 2017. Quality of research and risk of bias was evaluated using a self-designed assessment featuring 15 weighted criteria, ensuring incorporated publications had a high level of reporting and methodological quality. 13 papers were assessed in this review (598 participants in total). The publications were assessed particularly on the type of reliability measure used, the re-test interval utilised, the type of ON measure included and the location on the ON where the measure was taken. It was found coefficient of variation index was used most frequently, along with cross-sectional area being the most popular choice for ON measurements. However, major differences in study design did not allow for extensive quantitative analysis. Overall, repeatability was found to be good, but reproducibility was not reported on as frequently as expected.

7.2 Introduction

The ON functions as the starting point of the mammalian VP, connecting the eyes to the visual cortex in the occipital lobe. The ON is anatomically different to the other cranial nerves: it stems out from the diencephalon of the brain opposed to the brain stem, it is coated in three meningeal layers and it does not possess much ability to regenerate. Therefore, malformation of the ON due to disease or congenital disorder is correlated with irreversible visual dysfunction (110). Hence, a reproducible measurement method to determine ONS could aid clinicians in identifying ON disease progression and help determine the cause of visual impairment (122).

ONS can be altered by several congenital abnormalities such as ON hypoplasia (123–125) and diseases such as optic neuritis (2,3,113–117), ischemic optic neuropathy (106,117) increased intracranial pressure (43,118), glaucoma (106,119–121) or multiple sclerosis (2,3,122,123). ONS can also be impacted by tumours present in the VP, causing ONS increase due to the location of a tumour such as an OPG, stretching the nerve or decrease due to resulting white matter atrophy (31,113,116,125). The alteration of the ON structure subsequently impacts the ON's functionality, and as the nerve is the only pathway from the retina to the visual cortex, the damage can have devastating results (117).

ON atrophy from demyelination and axonal loss, resulting in changes in size, can be measured using various methods, including ultrasound and MRI. Using ultrasound, the retrobulbar ON diameter has been ascertained to be within the range of 2.86-3.67mm (124). Despite its convenience, ultrasound has its limitations. For example, the accuracy of the measurements rely upon the angle of incidence of the sound wave being perpendicular to the structure being imaged, a factor that is hard to quantify when assessing deep set, internal structures (124). Moreover, when the structures of interest exhibit tortuosity such as the ON when impacted by an OPG, measurements using ultrasound can be unreliable.

MRI offers an alternative imaging modality that has proved highly useful in imaging the ON. MRI's high spatial resolution is beneficial when attempting to image a small structure. Ordinarily, a sequence incorporating thin slices and a high resolution FSE sequence to yield T2-weighted images is used clinically when imaging the ON and orbits (126). This can be with or without fat suppression. Moreover, using a 3T MRI scanner gives a better SNR over a 1.5T scanner. Using an FSE or a TSE will limit the available time for eye movement to impact the usability of the MRI image (119). The use of MRI for quantifying ONS has resulted in the discovery of a significant decrease in ON diameter from the retrobulbar portion to mid orbit, from 3.5-3.1mm respectively (124).

It is speculated OPGs and their treatments such as radiotherapy, chemotherapy and resection may impact the structure of the anterior VP, thus potentially resulting in adverse impacts of the areas of the pathway that are not directly involved with the tumour e.g. the contralateral ON. This may go on to impact visual function. If evidence is found for this postulation, neuroprotective strategies can be implemented during treatment aiming to preserve vision. Hence, it is invaluable to have a reproducible, gold-standard measurement method of ONS that can be used in clinical trials of tumour treatments and neuroprotective strategies, and potentially in clinical practice. An agreed upon, standardised method would help to encourage cross-centre cohesion, a benefit not limited to OPG research, but to centres studying visual deficits as a result of ON malformation.

As there is no gold-standard measurement method for assessing ONS, studies in the literature have measured different dimensions of the nerve, including the diameter, the radius, the volume and the cross-sectional area (CSA). Moreover, these measurements are taken at different points on the ON, varying the results further. Due to the multiple techniques available to measure ONS, the literature is varied on the average size of the ON in healthy cohorts (118,124,127) and in case studies (113,117,123,125,128). Moreover, minimal literature is available for normative ONS in different age cohorts.

Before a measurement method can be widely adopted, its reliability must be proven. Mathematically speaking, reliability is the true variance, divided by the true variance and the error variance (129). To quantify reliability, the method must demonstrate its repeatability and reproducibility. There are varying definitions of the above terms, resulting in the confusion and sometimes subsequent misuse of them in text. In an effort for transparency, these definitions will be explicitly stated, referring to Bartlett & Frost's (130) classification:

Reliability is the degree at which a method will result in the same result when repeated multiple times. Thus, reliability indexes determine the extent of error within results. Reliability can infer reproducibility or repeatability.

Repeatability is the degree of variation between results collected successively from the same source, using the exact same conditions by the same observer. Hence, the evident variation can be attributed to the method itself. This is also known as test-retest reliability or intra-observer reliability.

Reproducibility is the degree of variation between results collected from the same source using different conditions such as a different observer or within a different setting. This is also known as inter-observer reliability.

Hence, the degree of both repeatability and reproducibility should be stated for a proposed measurement method to assure other research teams can use the method reliably. To statistically test the repeatability and reproducibility of a measurement method, either a reliability index such as an intraclass correlation coefficient (ICC) or a Coefficient of Variation (CV) should be concluded, with an explicitly reported test-retest interval.

In this report I have systematically reviewed the available, published literature on reliability of methods for measuring ONS. From the literature assessed as relevant and reliable, I will ask the following questions: (i) What type of reliability measure was taken? (ii) what re-test interval was used (iii) What type of measurement(s) are utilised when assessing ONS, and (iv) Where on the ON was the measure(s) taken?

7.3 Materials and Methods

7.3.1 Search strategy

A systematic literature search was performed by one researcher (B.E.H.) The databases searched were *MEDLINE/PubMed* (https://www.ncbi.nlm.nih.gov/pubmed/), *Web of Knowledge* (http://wok.mimas.ac.uk/), and CENTRAL (*http://www.cochranelibrary.com/about/central-landing-*

page.html). Utilising keywords stipulated by the previously read literature, I used the following search strings: ("reproducibility" [All Fields] OR "repeatability" [All Fields] OR "reliability" [All Fields] OR "test-retest" [All Fields] OR "TRT"[All Fields]) AND ("measurement" [All Fields] OR "measure" [All Fields]) AND ("radii" [All Fields] OR "radius" [All Fields] OR "diameter" [All Fields] OR "volume" [All Fields] OR "size" [All Fields] OR "area" [All Fields] OR "cross-sectional area" [All Fields]) AND ("optic nerve" [All Fields] OR "optical nerve" [All Fields] OR "ON" [All Fields]) AND ("MRI" [All Fields] OR "MR" [All Fields] OR "Magnetic Resonance Imaging" [All Fields]). I included all languages and no limits on dates within the search and the references of each paper read were combed for further relevant papers. For Web of Science, the articles were ordered by relevance and the top 100 were assessed.

In the first screening phase, papers focussing on a manual method of measuring the human ON were identified using the title and abstract. Here, manual method is defined as incorporating human interaction with the process, hence investigations use semi-automated methods will be included as long as sufficient human interaction is involved.

During the second phase of screening, I used the following exclusion criteria to filter the potential articles:

(a) the study must use a reliability index to measure the reliability of their method in measuring ONS, using either an ICC or a CV as to ensure easy comparison between studies;
(b) the study must utilise MRI sequences yielding structural images, excluding variations such as fMRI, DTI, DWI, perfusion MRI and MR spectroscopy
(c) the article must not be a review or meta-analysis;
(d) and the full article must be available in English.

7.3.2 Qualification of Researchers

The literature search was performed by the author B.E.H., who has a background in psychology, visual sciences and medicine, and has extensive experience in reviewing literature. I was supervised by R.A.D., D.A.W. and C.B. whom have collective experience in radiology, oncology, and physics and have trained and advised B.E.H. on correct systematic literature review techniques.

7.3.3 Data Extraction and Synthesis

The articles have been assessed for their quality of reporting of their observational study design and the quality of the study methodology using a self-developed scaled criteria [see Appendix A] to

effectively assess research quality in the observational studies specifically aiming to assess the reliability of a method of measurement. It was influenced by the Cochrane Handbook for Systematic Reviews of Interventions (131) and QUADAS (132) and QUADAS-2 (133).

Research 'quality' is not easily defined, but refers to the appropriateness of the study design, procedure, analysis and presentation with regards to providing a valid answer to the research question (134). Hence, the designed scaled criteria looks specifically at elements relating to the studies assessing a method of measurement of the ON assures reputable research quality. Each criterion was allocated a weight, from 0 to 2. Depending on the weight assigned, the studies were scored on each criterion. If a criterion was not met, a score of 0 was issued. If it is unclear if a criterion is met, or it partially fits the criterion, a score of 1 was issued. If the criterion was agreeably met, a score of 2 was issued. The highest possible score was 30 and the lowest was 0. If an overall score of <15 was given, the paper was discounted. A scaled assessment of research quality was chosen opposed to a simple, oversimplified checklist as it offers a more descriptive analysis of the research adhering to the criterions, opposed to a simple yes or no (135).

7.3.4 Risk of Bias Assessment in Included Studies

A bias is a systematic error or deviation away from the truth in results. An assessment of research quality is not synonymous with assessing research bias, as a study perceived to be conducted well may still be at risk of exhibiting a significant bias. Thus, the criteria assessing research quality has been adapted to consider all aspects of assessing possible bias within the selected literature, drawing influence from the specified sources of bias from the Cochrane handbook for systematic reviews of interventions. This includes assessing for allocation or indication bias, selection bias, performance bias, detection bias, reporting bias and other bias (131). This has been completed to ensure the selected reports are of high standard.

The data was abstracted by B.E.H. and a qualitative synthesis based upon the selected article results was written for each of the following themes:

(a) the type of reliability measure taken

(b) the re-test interval used

(c) the type of ON measure taken

(d) where on the ON was the measure(s) taken

7.4 Results

7.4.1 Literature search

The database search of PubMed and Web of Science using the stipulated keywords returned 63 papers: 32 from PubMed and 23 from Web of Science. 9 were found to be duplicates and removed, leaving 46. Following this, 30 publications were found from the reference lists. This gives a total of 76 publications. In the first phase of screening, 41 of the 76 papers were removed on the grounds of not meeting the first set of inclusion criteria, leaving 35. During the second phase of screening, 20 of the remaining 35 papers were removed as they did not meet the second set of stipulated criteria for inclusion. This leaves 15 publications for this review to focus upon. Figure 7.1 shows the results of the literature search process at each specified stage.

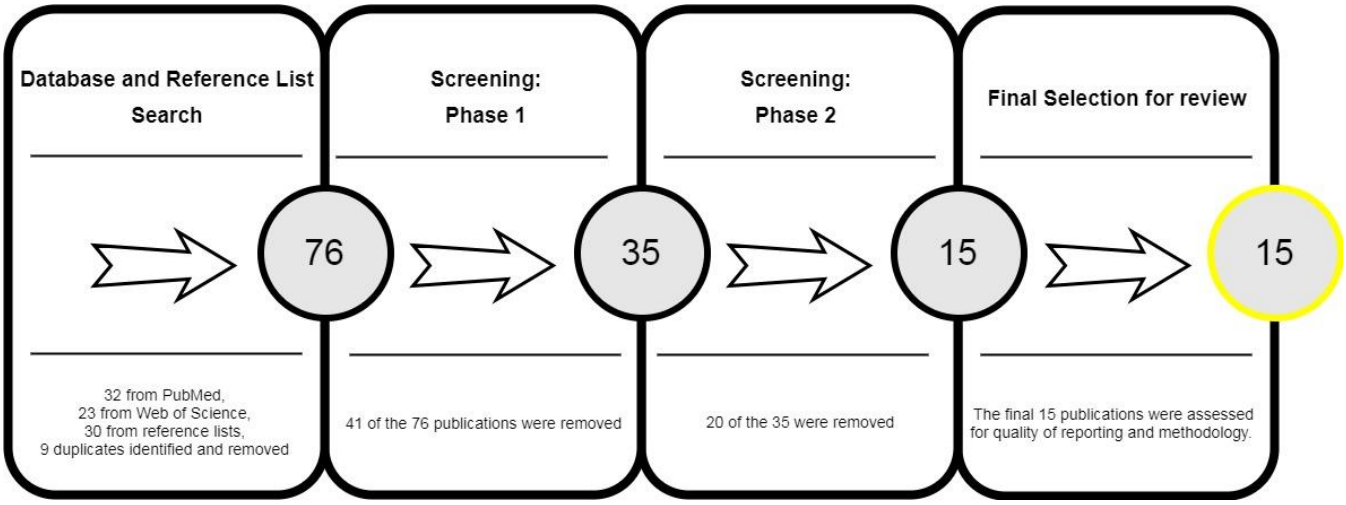

Figure 7.1 *The number of results at each stage of the literature search*.

7.4.2 Risk of Bias

The criteria used to assess the research quality and bias for each of the 15 studies is illustrated in Appendix A. A total of 15 criteria were used to assess the 15 publications: 6 criteria on the quality of reporting and 9 on the quality of the methodology. Meaning, 225 checks were carried out. The highest score available for quality of reporting was 12. The scores ranged from 7 to 12, with a median of 10. The highest score of quality of methodology was 18. Scores ranged from 5 to 18, with a median of 13. Thus, the highest score available for the overall research quality was 30.

The range in research quality was 14 to 26, with a median score of 23. Six of the publications were awarded a score of 80% or higher (113,115,116,118,119,123). The highest score achieved by a publication reviewed overall was shared by Hickman *et al.* (113) and Ramli *et al*. (119), while the lowest was Karim *et al.* (124). Following this assessment, publications that were not awarded a score ≥15, thus achieved a score of 50% or lower were discounted (124,136), leaving 13 publications [see Figure 7.2]. Figure 7.3 to 7.5 feature graphs depicting the results for quality of reporting, quality of methodology and overall research quality.

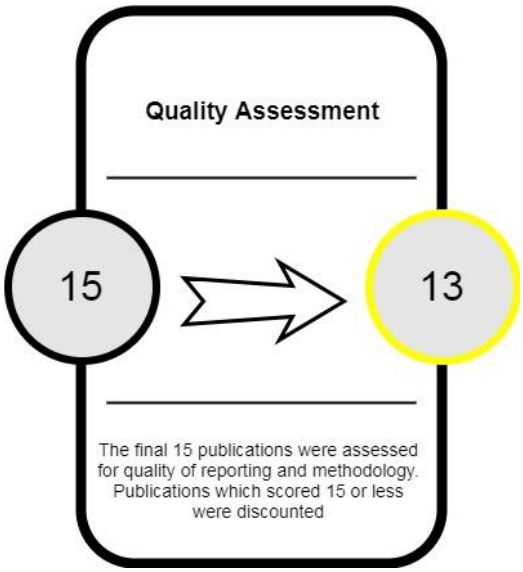

Figure 7.2 *The number of results after the quality of research assessment*

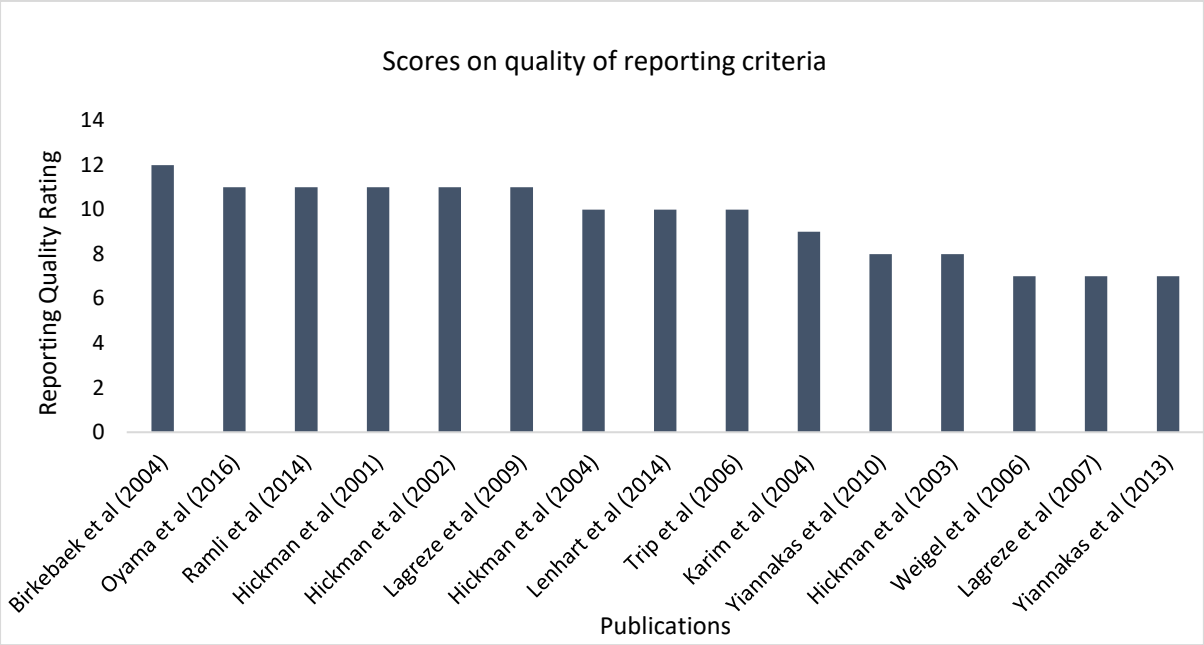

Figure 7.3 *A graph illustrating the individual scores of the reviewed publications for their quality of reporting*

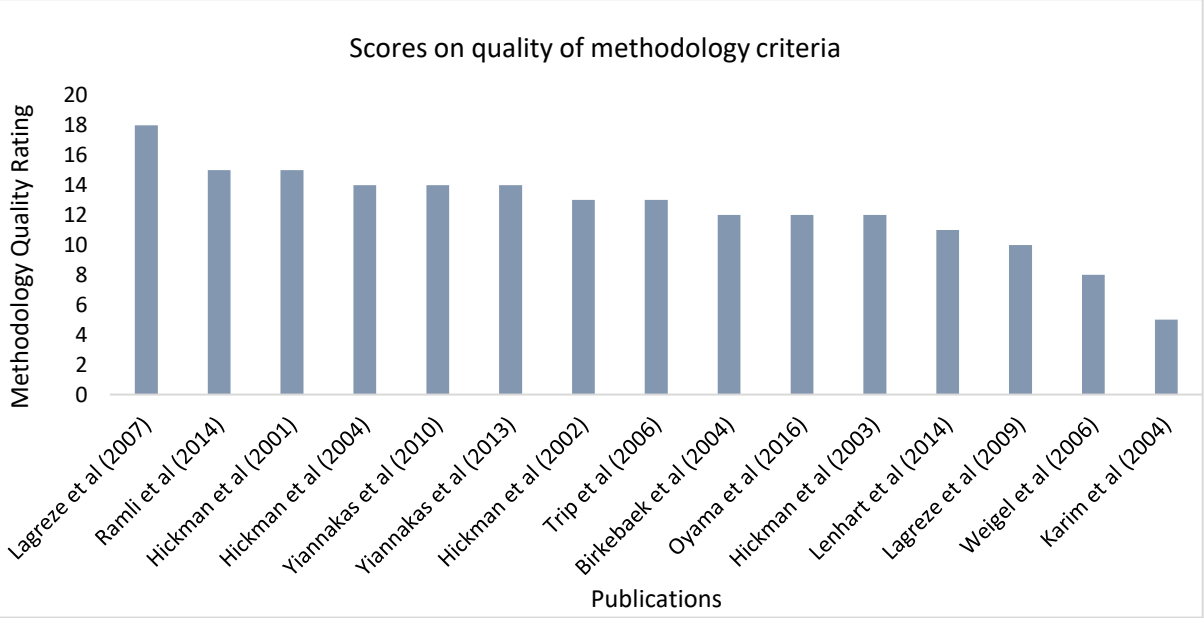

Figure 7.4 *A graph for the individual scores of the reviewed publications for their quality of methodology.*

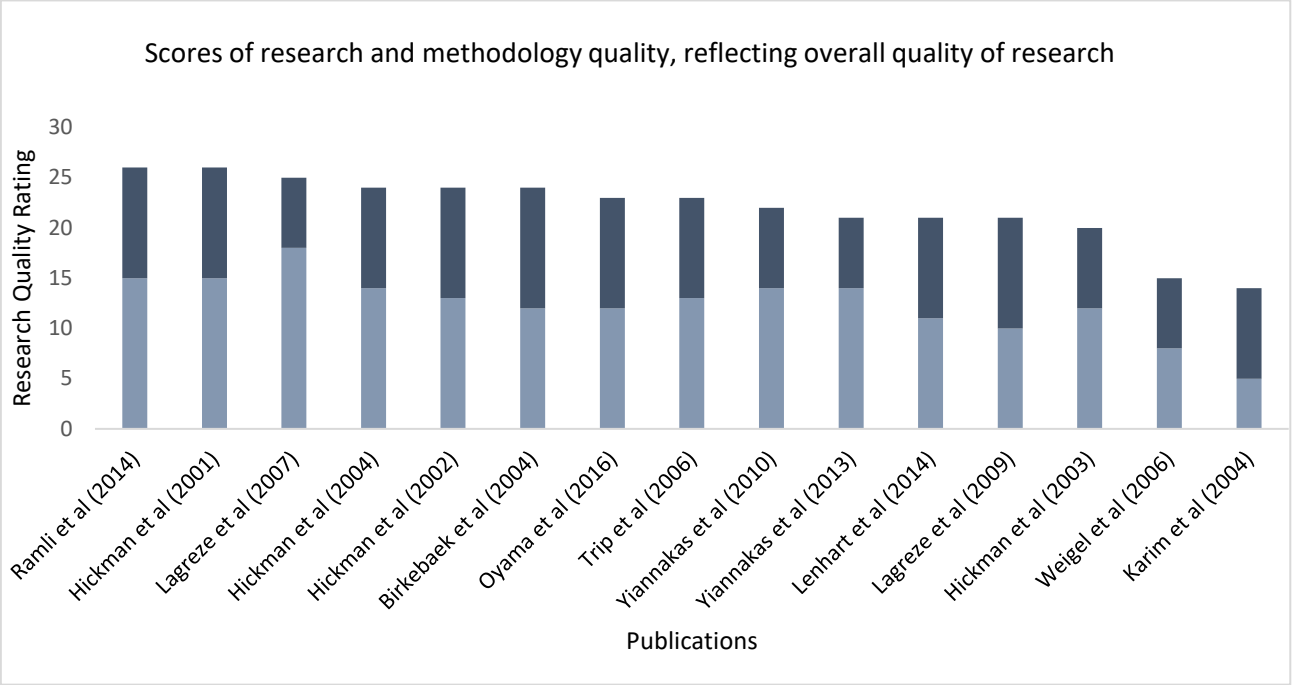

Figure 7.5 *A graph for overall scores of the reviewed publications for their quality of research, derived from their scores in quality of reporting and quality of methodology.*

Frequent limiting factors found when assessing the quality of reporting included publications not explicitly defining the design type of the study and publications not explicitly defining the method of recruitment utilised. Limiting factors found when assessing the quality of methodology included publications not explicitly defining their test-retest interval with regards to repeat measures and not including an inter-observer test to validate the reproducibility of the measurement method.

Furthermore, higher scores could not be awarded to publications citing blinding methods in their studies, as it was not explicitly stated whether the investigators taking repeated measurements were blinded to the previous measurement value taken, as well as the identity of the participant and their health status in case-control studies. Although, several publications illustrated their investigation's strength through reporting intra and inter observer reliability of the utilised measurement methods and further validating them through the correct use of reliability indexes.

7.4.3 Study Characteristics

Table 7.1 to 7.3 summarise the study characteristics. Table 7.1 gives a summary of each study's sample size, age-range study design and conclusions, in addition to the type of measure taken and the location

on the ON it was measured from. The sample size of the included investigations ranged from 8-90 with a mean of 46. The total number of participants equals 598 overall. All of the studies included had an observational study design, with almost half being case-control studies (6 instances, 46%) (2,113,116,119,125,127). The method types can be split into manual and semi-automatic. 46% of the publications used a manual method (118,119,123,125,127,128), while 54% used a semi-automatic method (2,3,113–116,137). The most common disease included was optic neuritis patients (5 instances, 38%) (2,113–116) while the remaining case publications focussed on ON hypoplasia (2 instances, 15%) (123,125), glaucoma (119,127) and other diseases such as multiple pituitary hormone deficiency, isolate growth hormone deficiency and idiopathic short stature (123).

Table 7.2 illustrates the MRI sequences, magnet strength, planes and image resolution used by the investigations. With regards to the MRI sequence used, the most frequent method was a T2-weighted FLAIR, which was used in 5 investigations (2,3,113,115,116). Fat suppression techniques were also frequent, used with T2-weighted FSE sequences, Dual Echo FSE sequences and Short T1 Inversion Recovery sequences (STIR) (113–116,118,137). Furthermore, the majority of the sequences were 2D, with only (119) and (125) opting to acquire 3D images. The majority of the studies utilised a 1.5T magnet strength scanner (9 instances, 69%) (2,3,113–116,125,128,137), while only four studies opted for a 3T magnet strength scanner (31%) (118,119,125,127). Additionally, most of the studies also used an appropriate image resolution of <1mm in the coronal plane 69% (9 instances) (2,3,113,115,116,118,127,128,137) Of these nine publications, four took images in the coronal – oblique plane, thus orthogonal to the ON (31% of the 13 papers included).

Table 7.1 *Summary of included studies in review, *results taken from coronal-oblique images*

	n	*Ages (median and range)*	*Study Design Type*	*Type of measure*	*Location of measure*	*ONS*
Ramli et al (2014)	**90** (30, 30, 30)	68.40 ±8.86, 65.5 ±9.04, 63.47 ±8.38, 55-84years	Case (2x)& Control	volume	Retrobulbar portion to chiasm	Case(1)OS=264.76mm^3±78.88, OD=264.03mm^3±78.53 Case(2)OS=167.40mm^3±45.36, OD=168.7mm^3±46.28 ControlOS=296.56mm^3±8.38, OD=297.8mm^3±71.45,
Hickman et al (2004)	**59** (29, 30)	30, 19-53years; 32 , 21-58years	Case& Control	CSA	Intraorbital portion	Diseased ON=16.1mm^2±3.1, Contralateral ON=13.4mm^2±2.0 Control=13.6mm^2±1.8
Hickman et al (2002)	**10**	44 25-55years	Case	CSA	Intraorbital portion	DiseasedON= 11.1mm^2±2.7; ContralateralON = 12.9mm^2±2.0
Lagreze et al (2007)	**33**	25 22-67years	Healthy	diameter	5, 10, 15mm posterior to globe	5mm=3.23mm (CI 3.14-3.32), 10mm=2.94mm (CI 2.82-3.05), 15mm=2.67mm (2.57-2.77)
Hickman et al (2001)	**33** (17, 16)	40 23-51years, 26 22-39years,	Case& Control	CSA	Intraorbital portion	DiseasedON= 11.2mm^2±2.3, ContralateralON= 12.9mm^2±1.8 Control=12.8mm^2±2.2
Oyama et al (2016)	**62**	Neonates (<1 year)	Healthy	Diameter& CSA	widest section at least 0.5mm from optic chiasm	Width= 2.7mm±0.2; CSA=3.5mm^2±0.5
Birkebaek et al (2004)	**85** (8, 38, 14, 15, 10)	25, 1-204months; 25, 1-221months; 18, 1-210months; 88, 37-192months; 84, 13-204months	Case (5x)	CSA	"not extracranial"	>12months = >2.9mm^2 =groups 3, 4 & 5 and in one or both ONs in groups 1 & 2; ON size lower in <12months
Trip et al (2006)	**40** (25, 15)	40 22-57years 36 30-56years	Case& Control	CSA	Intraorbital portion	DiseasedON=9mm^2±1.5, ContralateralON= 11.8mm^2±1.2 Control= 12.7mm^2±2.4
Yiannakas et al (2010)	**13**	32 ± 4	Healthy	CSA	3mm from globe to orbital apex	*9mm globe to apex=8.4mm^2 ±0.9; apex to 9mm = 11.6mm^2±2.9
Hickman et al (2003)	**66**	18-50years	Case	Mean area	Intraorbital portion	DiseasedON=18.4mm^2±3.8, Contralatera ON=17.8mm^2±3.6
Yiannakas et al (2013)	**8**	31, 29-33years	Healthy	CSA	3mm from globe to orbital apex	0-3mm=7.8mm^2, 3-6mm=6.2mm^2, 6-9mm=5.2mm^2, 9-12mm=4.7mm^2, 12-15mm=4.4mm^2, 15-18mm=4.2mm^2, 18-21mm=4.1mm^2, 21-24mm=3.8mm^2, 24-27mm=4.2mm^2
Lenhart et al (2014)	**52** (26, 31)	1 5.5	Case& Control	diameter	5mm posterior to optic disc and just posterior to optic canal	DiseasedON=1.36mm (CI 1.19-1.54) < Controls; ContralateralON=0.22mm (CI 0.03-0.07) < Controls; increase of 0.05mm per year
Lagreze et al (2009)	**47** (16, 11, 11, 9)	55y ±13, 63y ±7 68 ±8 54y ±16	Case (x3)& Control	diameter	5, 10, 15mm posterior to globe	Case(1)5mm=3.19±0.33, 10mm=2.66±0.38, 15mm=2.48±0.21; Case(2)5mm=2.98±0.26, 10mm=2.43±0.2, 15mm=2.26±0.25; Case(3)5mm=2.82±0.26, 10mm=1.98±0.2, 15mm=1.96±0.26; Control5mm=3.11±0.33, 10mm=2.66±0.3, 15mm=2.64±0.24

Table 7.2 *Summary of study MRI characteristics and sequences, *?-unclear from publication if meets criteria*

	Magnet Strength (Tesla)	*Scan Sequence*	*Image Resolution <1mm*	*Coronal Plane used*
Ramli et al (2014)	3	T1-weighted 3D FSPGR	X	✓
Hickman et al (2004)	1.5	sTE fFLAIR, fat-saturated dual echo FSE	✓	✓
Hickman et al (2002)	1.5	sTE fFLAIR, fat-saturated dual echo FSE	✓	✓
Lagreze et al (2007)	3	T2-weighted HASTE	✓	✓
Hickman et al (2001)	1.5	sTE fFLAIR, fat-saturated T2-weighted FSE	✓	✓
Oyama et al (2016)	1.5	T1-weighted TSE, T2-weighted TSE	✓	✓
Birkebaek et al (2004)	0.5	T2-weighted TSE, TSE dual	?	✓
Trip et al (2006)	1.5	sTE fFLAIR	✓	✓
Yiannakas et al (2010)	1.5	3D FRFSE-XL , sTE fFLAIR	✓	✓
Hickman et al (2003)	1.5	STIR	X	✓
Yiannakas et al (2013)	1.5	fat-suppressed T2-weighted multislice “single-shot” 2D TSE	✓	✓
Lenhart et al (2014)	1.5 & 3	3D T2-weighted SPACE, 3D T2-weighted CISS	?	✓
Lagreze et al (2009)	3	T2-weighted HASTE	✓	✓

Table 7.3 *Summary of study reliability measures and test methods*

	Intra-observer variance				*Inter-observer variance*				*Scan-rescan*				*Results*
	Reported?	ICC	CoV	Test-retest interval	Reported?	ICC	CoV	Test-retest interval	Reported?	ICC	CoV	Test-retest interval	
Ramli et al (2014)	✓	✓		within 1month	✓	✓		within 1month					Intra=0.997; Inter=0.994
Hickman et al (2004)	✓	✓	✓						✓			2, 4, 8, 12, 26 and 52 weeks	Intra(Controls)=4.0%, 0.9(0.84-0.97); Intra(DiseasedON)=4.8%, 0.96(0.93-0.99); Intra(ContralateralON)=5.3%, 0.84(0.7-0.97)
Hickman et al (2002)	✓		✓						✓			1year	MeanIntra=6.5%±3.7
Lagreze et al (2007)	✓		✓	3months	✓		✓	3months	✓		✓	3months	Inter=4%anterior, 7%posterior; IntraScan-rescan@5mm=4.18%; IntraScan-rescan@10mm=4.56%; IntraScan-rescan@15mm=6.54%
Hickman et al (2001)	✓		✓	3 measures over 7days					✓		✓	3 measures over 7days	MeanIntra=4.8%±3.6; MeanScan-rescan=6.5%±5.5
Oyama et al (2016)	✓	✓											Intra(ODheight/width)= 0.71, 0.85; Intra(OS height/width)= 0.84, 0.88
Birkebaek et al (2004)	✓		✓										Intra(HypoplasticON)=18.2%; Intra(ContralateralON)=16.5%
Trip et al (2006)	✓		✓						✓		✓		Intra(Controls)=4%; Intra(DiseasedON)=6%; Intra (ContralateralON)=5%; Scan-rescan(Control)=4%
Yiannakas et al (2010)	✓		✓	7-14days					✓		✓	7-14days	Intra(9mm-apex)= 3.4%±0.01; Scan-rescan(9mm-apex)= 3.5%±2; Intra(apex-9mm)=2.1±0.01; Scan-rescan(apex-9mm)=5.7%±2
Hickman et al (2003)	✓	✓	✓						✓	✓	✓		Intra (DiseasedON)=6.4%, 0.92(0.88-0.96); Intra (ContralateralON)=6.8%, 0.89(0.84-0.94) ; Scan-rescan(DiseasedON)=4.6%, 0.96(0.95-0.98); Scan-rescan(ContralateralON)=6.4%, 0.91(0.87-0.95)
Yiannakas et al (2013)	✓		✓	14days					✓		✓	14days	Intra(OD)=2.1%; Intra(OS)=1.8%; Scan-rescan(OD)=4.3%; Scan-rescan(OS)=4.4%
Lenhart et al (2014)	✓	✓											Mean Intra=0.98
Lagreze et al (2009)					✓		✓						Inter(5mm)=7%; Inter(10mm)=10%; Inter(15mm)=10%

7.4.4 Synthesis of Results

7.4.4.1 The Type of Reliability Measure(s) Taken

Table 7.3 shows the reliability measures of the investigations, including the test-retest interval incorporated and the reliability index results. All but one incorporated intra-observer variation measures (12 instances, 92%), Lagrèze *et al.* (127) being the publications without. Only three (23%) of the investigations included inter-observer variation measures (118,119,127). Eight of the thirteen publications (62%) included a test-retest design using newly collected scans (scan-rescan), rather than repeatedly measuring the previous MRI images taken, giving a better observation of measurement reproducibility (2,3,113–116,118,137).

To measure the variation of the intra and inter-observer measures, two major tests were incorporated: CV and ICC. Nine instances of CV were found (2,3,113–116,118,127,137) and five instances of ICC were found (114,116,119,125,128). However, no publications explicitly defined the type of ICC used. The lowest CV result was 1.8%, featured by Yiannakas *et al.* (137), for the intra-observer measurements of the right ON in healthy participants. The highest CV given was by Birkebæk *et al.* (123) at 18.2% for the intra-observer measurement of hypoplastic ONs. No confidence intervals (CI) were given with either result. The highest ICC was illustrated by Ramli *et al.* (119) at 0.997, for the mean intra-observer variation of all volume measurements of the ON taken in the study. The lowest ICC offered was 0.71 by Oyama *et al.* (128), for the intra-observer variation in the left ON height measurements in healthy participants. Both results were reported without CI.

7.4.4.2 The Re-Test Interval Used

A surprisingly low 46% explicitly stated their test-retest interval for repeated measures (3,113,116,118,119,137). 23% of the investigations stated a 2 week or less interval (3,113,137), Ramli *et al.* (121) used an interval of within 1 month and Lagrèze *et al.* (118) used an interval of within 3 months. Hickman *et al.* (116) reported a test-retest interval of 2-52 weeks for the scan-rescan measures but did not explicitly report the interval for the intra-observer measurements.

7.4.4.3 The Type of Optic Nerve Measure(s) Taken

The most frequent type of measure used for the ON was CSA (8 instances, 62%) (2,3,113,115,116,123,128,137). While diameter was the second most frequently used ON measure (4 instances, 31%) (118,125,127,128). See Figure 7.6 for details.

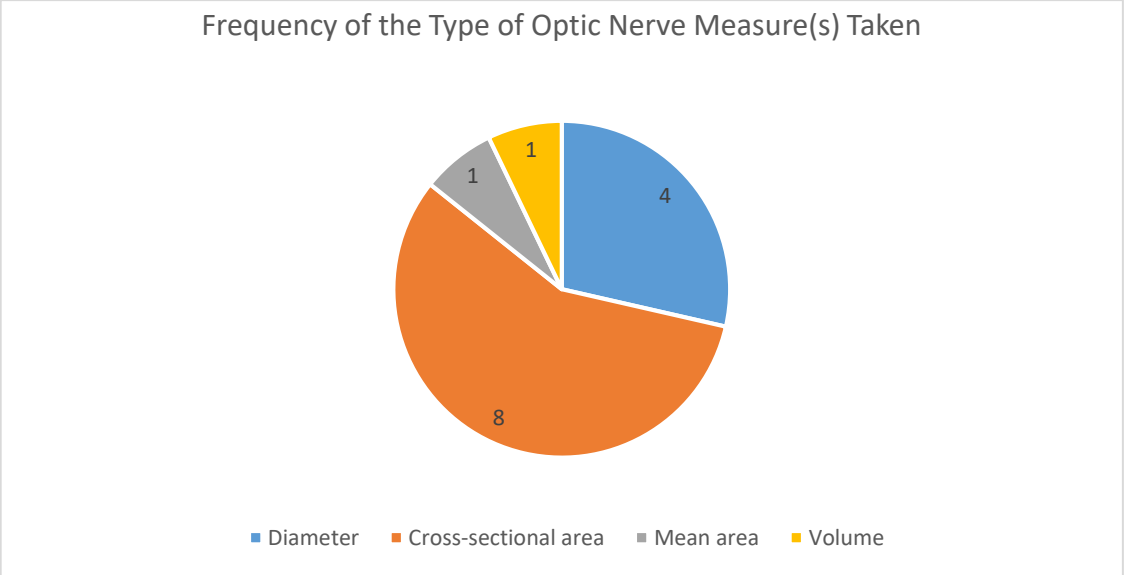

Figure 7.6 *A pie chart illustrating the number of studies that used the different types of optic nerve measurement. It can be seen clearly that cross-sectional area measurement is the most frequently used in the included studies*

7.4.4.4 Where on the Optic Nerve was the Measure(s) Taken

All of the studies featured took the ON measurements from the intraorbital portion of the ON. While some averaged the multiple measurements taken from consecutive MRI slices (38%) (2,113–116), other studies specifically define locations on the intraorbital portion of the ON, such as 5mm, 10mm and 15mm posterior to the globe (118,127) which occurred in 15% of the studies, and 3mm from globe to orbital apex, which was also reported in 15% of the studies (3,137). This data can be seen in Table 7.1 and Figure 7.7.

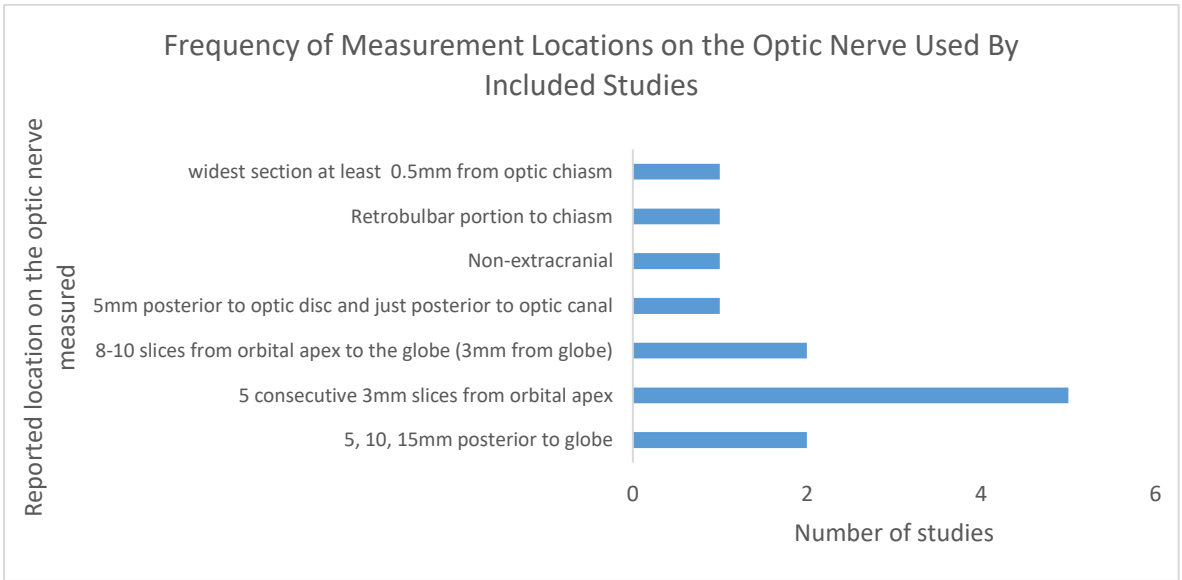

Figure 7.7 *A bar chart illustrating the number of studies that used different locations of optic nerve measurement. Using 5 consecutive 3mm slices is the most frequently used in the included studies*

7.5 Discussion

7.5.1 Reliability Measure(s)

The aim of this review was to identify reliable, manual methods of measuring the ON. Reliability has been defined as encompassing repeatability, measured by intra-observer measures, and reproducibility, measured by inter-observer measures. Moreover, the reviewer wanted to assess which studies also implemented scan-rescan measures. Whereby, not only are the measures repeated by the same observer (repeatability) or a different observer (reproducibility), but also repeats the scanning process with the same participants to subsequently measure all factors of reliability using the stipulated method.

The review illustrates 92% of the publications featured an intra-observer variation measure, thus identifying their methods as being tested for repeatability. However, only 23% of the publications incorporated an inter-observer variation measure, highlighting only a minimal selection of publications tested their method for reproducibility (118,119,127). Therefore, it can be questioned if the methods only tested for repeatability can be defined as reliable in its truest sense. The only paper considered all facets of reliability in testing their method of measurement was Lagrèze *et al.* (120).

To increase the validation of a measurement method using MRI, testing the method again using new scans would further indicate method reliability. This is due to repeat scanning using the same participants and same sequences does not yield identical images, because of possible changes in image orientation and magnetic inhomogeneity (138). With regards to scan-rescan measures, 62% of the featured publications included a test-retest measure using newly collected scans from the same participants. However, only one publication used a scan-rescan method in conjunction with inter-observer variation measures (118).

To ensure reliability was measured, the criteria of publication inclusion of this review incorporated the publications must feature a reputable measure of reliability, either via an ICC or CV index. Other methods of testing reliability are used in the literature such as Bland & Altman plots, paired t-tests and Pearson correlation coefficients. Although, Bland & Altman plots and paired t-tests are designed to be used for analysing agreement rather than identifying error variation, while Pearson correlation coefficients are designed to only measure correlation (129,139). Hence, the previously stated reliability indexes are a more suitable choice for assessing variation

7.5.1.1 Reliability Indexes: Coefficient of Variation

Comparisons of the reliability indexes across the featured studies is difficult due to the differences in participant age range, case type, method type, re-test interval, placement and type of measurement of the ON and if the indexes are averaged across all measurements or given for each individual subset. However, as there are some similarities between studies, some evaluations can be made.

The CV describes the spread of the data and it should only be used for measurements on a ratio scale. Hence, it is suitable for the featured measurements. 77% of the featured publications utilised the CV index of reliability. For example, for the intra-observer CSA measurements of healthy volunteers and controls using a semi-automated method, the CV results ranged from 1.8% (137) averaged from nine measurements across the entire ON, to 6.54% (116), when specifically measuring the nerve at 15mm from the globe. These values indicate a low level of measurement variation, implying the methods have a high rating of repeatability in healthy ON measurements. However, neither study gave CI for their CV reliability index. Thus, the range in which the real answer lies is unknown to the reviewer, and to anyone wanting to replicate the study.

With reference to the intra-observer CSA measurements of people with ON pathology using a semi-automated method, the CV results ranged from 4.8% from Hickman *et al.* (118) to 6.5%±5.5 from Hickman *et al.* (113). Both measures were averaged from measures made in five consecutive 3mm slices from orbital apex in acute optic neuritis patients. Although no CI were given by Hickman *et al.* (118), the low values of reliability index illustrate the methods appear to have a high rating of repeatability in ON measures in volunteers with acute optic neuritis. The ratings are not as low as Yiannakas *et al.* (137) method, but it can be inferred measuring the size of a diseased ON is more challenging than measuring a regularly sized healthy ON and thus measures would illustrate more variation.

Only one study looked at the intra-observer CSA measurements of people with ON pathology using a manual method of measurement. Birkebæk *et al.* (123) reported a reliability index of 18.2% in the diseased nerve in volunteers with hypoplastic ONs and 16.5% in the contralateral nerve. These values were averaged from the multiple measures taken from the entire intraorbital portion of the nerve. Despite no CI being stated, and 18.2% being the highest CV result seen in this review, it is considered a low level of variation. Hence, implying a high rating of repeatability in case ON measurements using a manual method, but a lower level of reliability than illustrated by the semi-automated methods assessed in this review.

When comparing the CV results in the inter-observer measurements, only two publications can be considered: Lagrèze *et al.* 2007 and 2009 (118,127). Both studies investigated manual diameter measures of the ON at 5mm, 10mm and 15mm posterior to globe, and incorporated healthy volunteers. Lagrèze *et al*. (120) reported a CV of 4% anterior to the globe at 5mm, to 7% posterior to the globe, again at 15mm. While Lagrèze *et al.,* (129) reported a CV of 7% at 5mm, 10% at 10mm and 10% at 15mm. Lagrèze *et al's* previous study focussing on a comparison between MRI and conventional sonography appears to have yielded a more reliable result. It can be postulated a different test-retest interval was utilised in the studies, as Lagrèze *et al.* (129) did not explicitly report one. However, both results are low, thus inferring the methods have a high rating of reproducibility in healthy ON diameter measurements using a manual method.

7.5.1.2 Reliability indexes: Intraclass Correlation Coefficient

Reporting an ICC was the second option of repeatability index permitted in this review. Despite the statistical test being a widely used method of testing reliability, only 38% of publications in this review reported it (114,116,119,125,128). It can be speculated finding the CV index is easier as it only requires the mean, the standard deviation and a small amount of division, while ICCs require a statistics package.

In all five instances the ICC was found, not one publication explicitly reported the type of ICC test used. It is assumed as the reliability index is analysed from multiple measures from the same rater (with regards to intra-observer variation measures), a 2-way mixed-effects model should be incorporated with an absolute agreement definition (129). Although, as there are six different types, this hampers the reviewer's ability to deduce if the correct form of test was utilised and thus reduces the reviewer's belief of the reliability of the methods reported.

The highest ICCs were illustrated by Ramli *et al.* (121) at 0.997, for the mean intra-observer variation and 0.994 for the mean inter-observer variation, both reported without CI. These values are averaged across the ON volume measures taken manually from healthy volunteers and both mild and severe sufferers of glaucoma. In contrast to the highest CV value featured in this review which was for a semi-automated method (137), this method is completely manual. Therefore, it cannot be argued a partly automated measure would yield more reliable results. However, it would have been insightful to see the reliability index of the measures taken from each participant group separately by Ramli *et al*. (121), as to illustrate if there is a group difference of variation which is evident in the CV results.

The ICC deduced from intra-observer measures of the CSA of the ON in volunteers with optic neuritis, Hickman *et al*. (118) reports an index of 0.96 (0.93-0.99) in the diseased nerves and 0.84 (0.7-0.97) in the contralateral nerve, while Hickman et al. (114) found an index of 0.92 (0.88-0.96) in the diseased nerves and 0.89 (0.84-0.94) in the contralateral nerve. Both these methods averaged the CSA from five consecutive 3mm slices from orbital apex using a semi-automated method. Both results are very similar, are low and both note the CI, indicating a high rating of repeatability in the methods utilised in volunteers with optic neuritis.

The lowest ICC indicated was from Oyama *et al.* (128), who manually measured intra-observer variation of the height and width of the ON in healthy participants. Oyama *et al.* (130) reported 0.71 for the height and 0.85 for the width averaged across the left nerves measured, and 0.84 and 0.88 respectively for the right nerves measured. No CI were reported so the range in which the real answer lies is unknown to the reviewer. These repeatability indexes are lower, inferring the method does not appear to be as repeatable as the aforementioned studies. However, this may be due to the method measuring the height and width of the ON, as this would depend on the orientation of the image being measured, which can vary.

7.5.2 Re-test Intervals

A secondary aim of this review was to identify what re-test intervals were being used in studies citing reliable measurement methods of the ON. The explicit identification of the time interval chosen for test-retest measures is extremely important when comparing measurement methods, as the time interval can impact the reliability. For example, a short re-test interval may induce carry-over effects in the rater(s) due to learning and practice effects. However, a long re-test interval may result in changes in the participant status (140). Due to these changes, a researcher wanting to replicate a study would require information regarding re-test interval chosen, to control for the same effects. Moreover, to compare similar studies the re-test interval indicated by each study should be evaluated, as differences in reliability may be attributed to this factor.

Only six publications in this review reported their re-test interval (3,113,116,118,119,137).This implies the remaining seven publications could not be effectively replicated unless further guidance was sourced from the research teams. Hickman *et al.* (113) and Yiannakas *et al.* (3,137) featured the shortest re-test intervals of up to 2 weeks, with Hickman *et al.* (115) having the shortest interval of 3 measures within 7 days. It can be implied the rater of this study may have been subjected to learning effects which may have subsequently interfered with the resulting reliability indexes. However, the

short interval for optic neuritis-based investigations is related to the rapid time course of the disease. Lagrèze *et al*., (120) used the longest re-test interval of 3 months between repeated measurements. This interval is speculated to be long enough to ensure immediate learning effects would not impact the measurement reliability. However, there are no gold standard guidelines for test-retest intervals. Thus, it is up to the individual research team to decide on an appropriate re-test interval. Nevertheless, for research of reputable quality, it should be explicitly stated in the methods.

7.5.3 Optic Nerve Measurement Types

A further aim of this study was to identify what types of measures are being taken of the ON in the literature. The ON can be considered as typically ellipsoid in shape, thus different dimensions of the nerve can be measured. The entire ON volume (119), the nerve CSA (2,3,113,115,116,123,128,137), the diameter (118,125,127,128) and mean area (114). In this review, the most frequent type of measure used for the ON was CSA (8 instances, 62%) while diameter was the second most frequently used ON measure (4 instances, 31%).

When the ON is diseased, its structure can change and tortuosity can occur. Hence, some forms of measurement are more advantageous than others. For example, measuring the diameter of the ON implies the nerve is completely circular which is less likely when the nerve is structurally impaired by disease. In an effort to retain a systematic approach when measuring the diameter of the nerve, an observer may choose the same plane and the same orientation to take the measurement, but the nerve may appear wider at a different orientation. Diameter measurements may not be the most suitable method when measuring structurally altered ONs. However, CSA and volume would take these factors into account and reflect changes in ON shape using more dimensions than diameter measures. This postulation is supported by the fact two of the publications utilised diameter measures were measuring only healthy ONs so a regular shape could be assumed (118,128). In addition, Oyama *et al.* (130) measured both diameter and CSA.

7.5.4 Location of Optic Nerve Measurements

The last aim of this review was to assess where on the ON measures were being taken from. The ON is approximately 50mm long (141), thus there is variation in the location and length of the ON measured within the reviewed literature. All of the studies measured the ON from the intraorbital portion, which is the longest segment of the ON with a length of 25mm (141). This part of the nerve

appears to have been chosen due to its relative ease at being identified on MRI, its length and the ease at which the orbital apex can be defined.

Five studies used a semi-automated contouring technique to measure the CSA of the ON on five consecutive scan images (2,113–116). The measurements were averaged across the nerve and the scan slices to result in a single value. However, as it has been found the ON decreases in size as it moves away from the optic bulb, the CSA reported would not be a true representation of the ranging size of the entire intraorbital portion (124). In addition, if the ON indicated any structural abnormalities, this method would not indicate where on the ON is abnormal. Moreover, the method would not indicate if other locations of the intraorbital ON have been impacted by the local abnormality.

Other studies were more specific of their locations of measurement, rather than averaging across the entire nerve. For example, Lagrèze *et al.* (118,127) reported measuring the ON CSA at 5mm, 10mm and 15mm posterior to the globe. This method can illustrate the decrease in size of the ON. For example, Lagrèze *et al*. (120) cites a decrease from 3.23mm (3.14-3.32) at 5mm to 2.67mm (2.57-2.77) at 15mm in healthy volunteers. Moreover, this method would highlight any structural changes in the intraorbital portion of the ON, but may miss them if they exist in the remaining 10mm of the nerve is not measured.

Yiannakas *et al.* (3,137) evade this problem by measuring 3mm from the globe to the orbital apex, hence measuring the entire intraorbital portion of the ON. Both publications confirm Karim *et al.* (124) findings and can identify any abnormalities across the ON portion measured. However, the more measurements are made, the longer the process takes and the more open the method is to human error.

The main methods evident from this literature search are using a semi-automated contouring technique on five consecutive 3mm slices of a sTE fFLAIR sequence to assess mean CSA of the intraorbital portion, manually assessing the diameter of the intraorbital portion on a T2-weighted HASTE sequence at 5mm, 10mm and 15mm locations, using a semi-automatic method to assess the CSA of the intraorbital portion of the ON 3mm from globe to the orbital apex at 3mm locations and other manual measures averaging the CSA or volume of the intraorbital portion of the ON.

7.5.5 Summary

Overall, the reported methods included in this review have a high level of repeatability, but further assessment of the reproducibility using inter-observer measures should be considered to ratify them

as reliable and ultimately viable enough for another research team to repeat the method and find similar results. The included publications have been limited by errors such as not explicitly specifying the type of ICC used, not reporting the re-test interval for repeated measures and not giving CI of reported reliability indexes.

It does appear using a CSA measurement is the better option for measuring diseased and healthy ONS, with separately reported measurements being taken along the nerve to assess changes in size may occur. Moreover, from this review, it appears both intra-observer and inter-observer measures should be both taken with a suitable test-retest measure, along with scan-rescan measures if possible. The reliability should be measured with either an ICC or a CV, preferably both, with CI clearly stated.

To my knowledge, this review is the first systematic literature review of reliable measurement methods for manually assessing ONS using MRI. However, it is limited by incomplete reporting. For example, measurement methods have assessed the reliability of their method using statistical tests other than those stipulated in the inclusion criteria of this review were not reported on (142). Additionally, automated measurement methods were also not reported on (117,143). Moreover, this review is limited as direct comparisons are difficult to make across the studies due to the vast range in scan sequence, measurement type, location and case type. Of what comparisons have been made, it is clear there are numerous ways to conduct ONS measures using MRI of which are deemed reliable.

7.6 Conclusion

I have shown the literature indicates MRI as being a sufficient modality to produce repeatable measures of ONS. Both CSA and diameter are popular choices for ONS measures with it appearing preferable to take multiple measurements along the intraorbital portion of the nerve. With regards to testing reliability, the literature indicates intra-observer variation is reported regularly indicating measure repeatability, but inter-observer is not reported as often, thus reproducibility is not as well established. Furthermore, when reporting reliability tests, it is indicated as invaluable to explicitly report all factors of the analysis, including re-test interval, exact type of reliability index and applicable CI.

Based upon this review's findings, I will now implement a CSA measure in the coronal plane to assess the ONS in children with OPGs and compare this to a linear measure in the axial plane. To thoroughly assess the reliability of the method and to address the deficiencies within the reviewed literature, I will conduct a reliability assessment utilising clearly stated ICCs and CV measures to identify intra and inter-observer measurement variation, described in Section 10. Once the method is quantified as

appropriate for use in a cohort with irregular ONS, it can be used in a normal cohort for comparison, described in Section 11. Following a normal baseline being established, ONS can be assessed for correlation with VA and other measures of structural integrity such as RNFLt, described in Section 12.

8.0 Research Questions

The following question will be addressed in **Section 10.0**

i. What is a reliable measurement method for the assessment of ONS in children with OPGs?

The following questions will be addressed in **Section 11.0**

ii. What is the variation of ONS and VA in healthy children?

iii. Does ONS correlate to VA and/or Age in healthy children?

The following questions will be addressed in **Section 12.0**

iv. Does ONS, VA, and RNFLt correlate in children with OPGs?

v. Does ONS correlate to Age in children with OPGs?

vi. Does ONS and RNFLt predict VA in children with OPGs?

vii. Does ONS and VA found in healthy children significantly differ to ONS and VA found in children with OPGs?

9.0 Hypotheses

i. A CSA measurement method will be a repeatable and reproducible measure of assessing ONS in children with OPGs

ii. ONS will correlate to Age in healthy children and in children with OPGs

iii. ONS will correlate to VA in healthy children

iv. ONS, RNFLt and VA will all correlate in children with OPGs

v. ONS and RNFLt will predict VA

vi. ONS in children with OPGs will be significantly smaller

vii. VA will be significantly worse in children with OPGs compared to healthy children

10.0 Testing a Reproducible Method to Assess Optic Nerve Size in Children with Optic Pathway Gliomas

10.1 Introduction

The following research question was investigated: what is a reliable measurement method for the assessment of ONS in children with OPGs? To achieve this, a manual method of measuring the ON needs to be decided upon and tested with a paediatric cohort of children with OPGs. Following from the review of the literature, two manual method types were decided on: CSA and the diameter of the nerve at 5mm from the optic bulb. CSA was chosen as it was a frequently used measurement within the literature that considers more planes of the ON compared to a singular diameter measurement. A simple linear diameter measurement method was chosen to compare to the CSA measure. The location of the measure at 5mm from the bulb was chosen as it is far enough away from the bulb as to be impacted less by bulb movements related motion artefacts.

10.2 Method

10.2.1 Setting and Study Design

This investigation conducted by BH and CB has a retrospective, observational design, aiming to assess the repeatability and reproducibility of two measurement methods of ONS measurement in a cohort of children with OPG. This study was approved by the University of Nottingham Medical School Ethics Committee and ensured protocols of Good Clinical Practice were adhered to. It was conducted within the Radiological Sciences Group in the School of Medicine, The University of Nottingham. Retrospective scans were used to assess ONS, and consent was collected from parents/guardians.

Inclusion criteria for the study specified children who were diagnosed with an OPG and had appropriate MRI structural scans. To be considered appropriate, the scans must have been conducted at either 1.5T or 3T, using a CISS scan sequence. The CISS scan must have a slice thickness of ≤1mm as to ensure acceptable resolution for ON measurement. Exclusion criteria included any scans that were too degraded by movement. Scans were disregarded if the white ON sheath could not be clearly defined.

The DICOM database was searched by BH using the HERBY workstation and appropriate scans were identified and collected. During assessment, BH and CB were masked to participant identity, NF1 status and OPG severity, as well as their previous measurements during the repeated measurement stage. Participants could only be identified via a numerical label allocated to them.

10.2.2 Cross-Sectional Area Assessment of Optic Nerve Size

OsiriX MD (2013) DICOM viewing and image analysis software was used to measure ONS for both measurement types. To standardise the viewing conditions, a private office was utilised by BH an CB to minimise disturbances and overhead lights were switched off. CSA of both ONs were measured and reported. A strict protocol was used to measure the CSA perpendicular to the long axis of the ON as follows. Individually, each scan was opened in 3D viewer, to illustrate each orientation of the 3D image. CB and BH used the axial plane window (outlined in purple) to find the slice that most clearly illustrates the ON within 2cm of the optic bulb, by aligning the crosshair tool indicating the sagittal plane (yellow line) directly through the ON. Using the sagittal plane window (outlined in yellow), the crosshair tool indicating the axial plane (purple line) was checked to ensure it also ran directly through the centre of the optic bulb in this orientation too.

The linear calliper tool was used to measure both 5mm and 10mm from the optic bulb in the axial plane window, while the zoom tool was used to ensure this was done as accurately as possible. Once the distance was allocated, the centre of the cross where the yellow and blue coloured lines intersect was placed at the specified distance from the optic bulb. Following this, the coronal plane window was opened and the ON was zoomed in on. Using the closed polygon tool the outside of the ON was drawn around, visible by the hypointense signal of the ON dural sheath encircling it. The resulting image can be seen in Figure 10.1a.

10.2.3 Diameter Assessment of Optic Nerve Size

As before, OsiriX MD was used to assess the size of the ON. The same protocol was followed as the previous method. However, instead of using the closed polygon tool in the coronal plane, the line tool was used to measure the diameter of the nerve in the axial plane, perpendicular to the line defining the distance from the optic globe. The image was rotated so the optic globe was at the top of the screen each time, to ensure all diameter measurements were taken at the same orientation. The resulting image can be seen in Figure 10.1b.

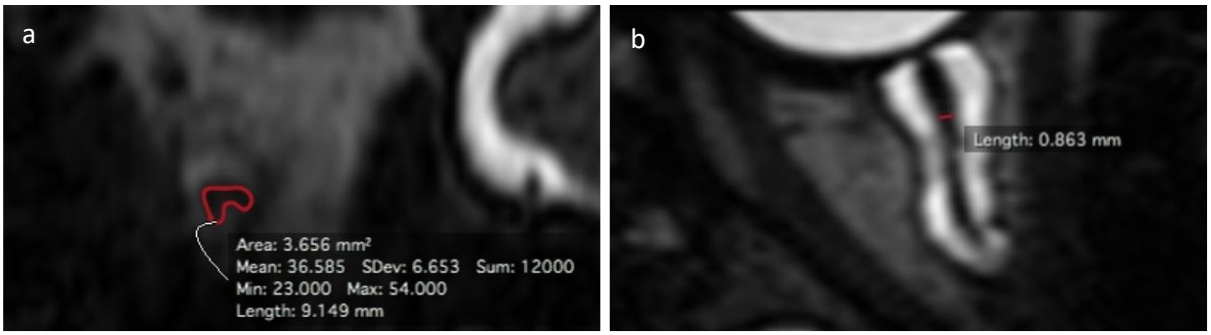

Figure 10.1 *illustrating the two measurement methods of ONS. a is the CSA method in the coronal plane and b is the diameter measurement in the axial plane. Note the tortuosity of the nerve on the left.*

10.2.4 Intra-observer and Inter-observer measurements

All subjects were assessed by BH to measure for intra-observer variance. CB assessed a subset of 5, randomly selected participants to assess for intra and inter-observer variance. BH and CB conducted their assessments separately, on different days. An interval of 2 weeks was taken between BH's repeated measurements of both the CSA and diameter methods. An interval of within 1 week was taken by CB for the repeated measurements of the CSA method. The order of the participants assessed was randomised.

10.2.5 Statistical Analysis

One eye was randomly selected using a randomised list and assessed as the eyes are a paired structure, thus it cannot be assumed that they are statistically independent, which is an assumption of Bland and Altman plots (144). Data for OPG014 was removed as the repeat measures were too varied due to the ON being extremely tortuous.

To assess intra-observer variance of the measures, a 2-way mixed-effects model ICC using absolute agreement was used, as well as the CV and a Bland Altman plot. The Bland Altman graphical model was chosen as it visually illustrates the agreement between two methods (139).

For inter-observer variance, a 2-way random-effects ICC using absolute agreement was used. This was chosen as the reliability results can be applied to any observers who possess the same characteristics as the featured observers (145). In addition, the CV was found.

SPPS statistics package were used to compute ICC, CV and Bland and Altman plots. An ICC value lower than 0.5 would be regarded as not reliable. This limit was chosen with consideration of the small sample size used. To mimic this, the lower limit of the CV was decided to be 50%, again taking into account the small sample size. For the Bland and Altman plots, an acceptable error was decided at ±2mm² for the CSA measures and ±1mm for the diameter measures. These errors were decided with consideration of the ease of the method, and the impact of over/under-estimating the ONS.

10.3 Results

10.3.1 Retrospective MRI Parameters

52 scans were found and 3 scans were excluded as they did not meet the inclusion criteria. 1 scan per participant was randomly selected, giving 22 scans. For 82% of the scans, the resulting MRI parameters included: Phillips 3T Achievia, TR= 2000, TE=200, a SENSE-HEAD 8 multicoil, Flip-Angle=90, Bandwidth=320Hz, matrix 304/208, echo-train=55, slice thickness 1mm, space between slices 0.5mm and pixel spacing 0.302734.

Of the remaining 3 scans completed on a 3T Achievia, the sequences varied on Bandwidth, 2 used 337Hz while the other used 337Hz. The last scan was conducted using a Phillips 1.5T Intera, the sequences had a TR=1500, TE=250, Bandwidth=201Hz, matrix size=400/245, echo train=74 and a pixel spacing=0.322266. Despite these differences, the rest of the parameters stayed the same unless specified, and slice thickness consistently remained 1 as per the inclusion criteria.

10.3.2 Participants

22 participants were included in the study, 13 males and 9 females (2-15years, mean=8years) 12 were NF1-positive thus have syndromic OPGs, while 10 have sporadic OPGs. [see Figure 10.2]. For the subset of participants that additional intra-observer variance and inter-observer variance was measured, 4 participants were randomly selected: 2 males (aged 6 and 6 years) and 2 females (aged 6 and 2 years). 2 were NF1-positive and 2 were not.

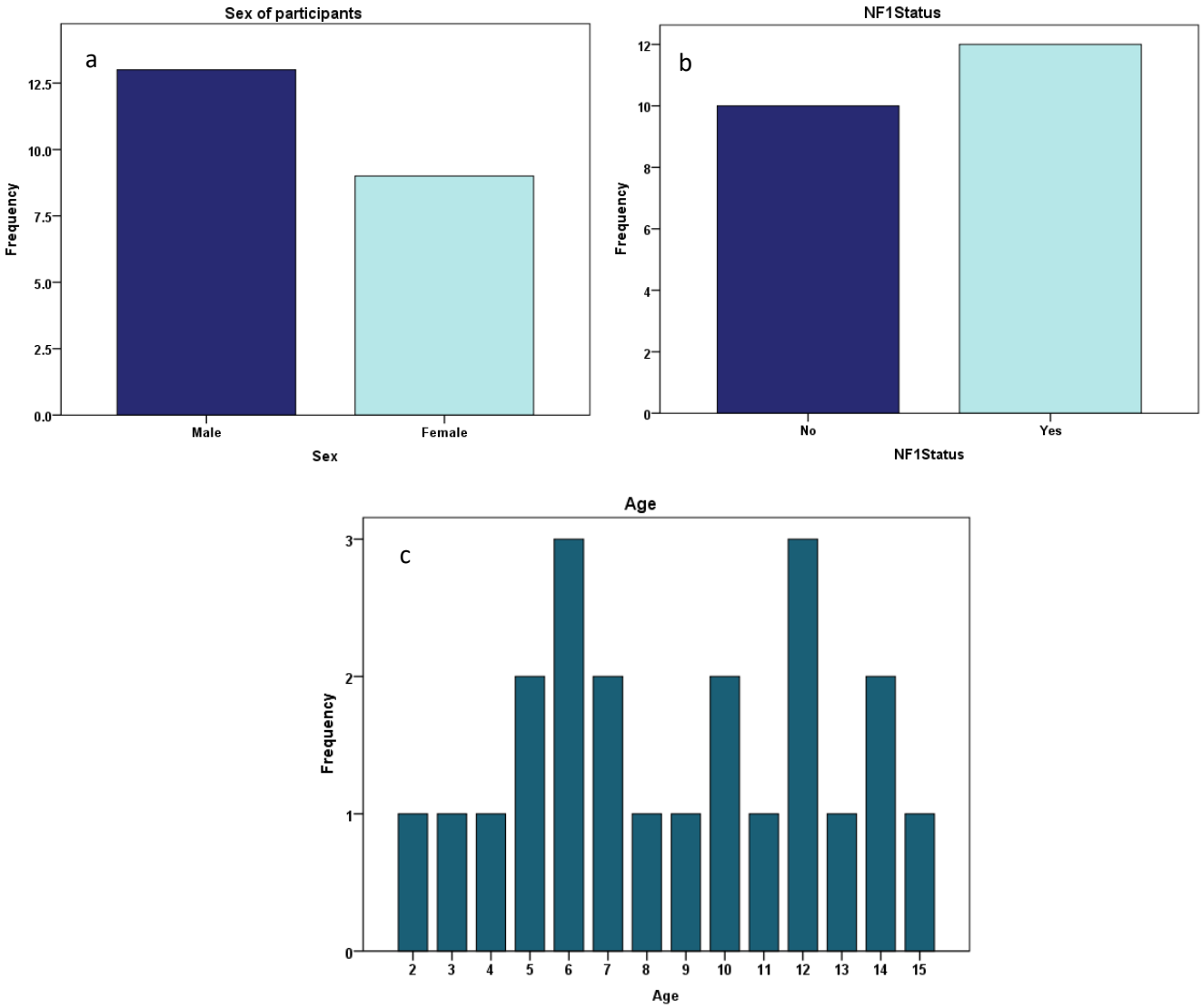

Figure 10.2 *illustrating the frequency of sex (a), NF1 status (b) and age range (c) in the whole cohort*

10.3.3 Normality

All data was assessed for normality by visually inspecting histograms, q-q plots and Shapiro-Wilks test. All data appears either normal or deemed acceptable for inclusion.

10.3.4 Intra and Inter-Observer Variance Results

For the assessment of all 21 participants, the ICC and CV for the CSA measurement were 0.9 (95%CI=0.772-0.958) and 10.89%, respectively. For the subset of participants whom underwent additional intra-observer tests by both BH and CB, the ICCs accrued were 0.98 (0.815-0.999) and

3.72%, and 0.994 (0.93-1) and 2.81%, respectively. For the inter-observer variance of the CSA measurements, the index scored 0.526 (-0.151-0.953) and 27.81%.

For the measurement of variance of the diameter measurement for 21 participants, BH recorded an ICC index of 0.984 (0.961-0.993) and 2.86%. These low values of variation were mimicked by the indexes for the subset of participants: 0.995 (0.929-1) and 2.15% [See Table 10.1].

Table 10.1. *Results table for intra and inter-observer variance, mean ONS and SD*

Measurement Type	n	ICCa (CI)	CoVb	Bias (CI)	Mean ON size	SDc
Cross intra-observer-BH	21	0.9 (0.772-0.958)	10.89%	0.324 (±0.4685)	4.924mm²	0.536
Cross intra-observer-BH	4	0.98 (0.815-0.999)	3.72%	0.216 (±0.6554)	5.162mm²	0.192
Cross intra-observer-CB	4	0.994 (0.930-1)	2.81%	0.0588 (±0.2935)	3.514mm²	0.099
Cross inter-observer	4	0.526 (-0.151 -0.953)	27.71%	1.768 (±1.8963)	4.407mm²	1.221
Diameter intra-observer-BH	21	0.984 (0.961-0.993)	2.86%	0.0118 (±0.0498)	2.097mm	0.060
Diameter intra-observer-BH	4	0.995 (0.929-1)	2.15%	0.0073 (±0.1525)	2.247mm	0.048

[atwo] ICC types were used, intra-observer = 2-way mixed-effects model using absolute agreement, inter-observer = 2-way random-effects model using absolute agreement. [b] Coefficient of Variation expressed as a percentage. [c] 95% of measurements are expected to lie within this range of the true value. (Cross= cross-sectional area, n=number of scans assessed)

10.3.5 Bland and Altman Plot results

Plots were completed to compare the agreement between the intra-observer measures of the entire participant group for both CSA and diameter measures and the inter-observer measures for CSA. [See Figures 10.3-10.5 and Table 10.2]. 95% CI of the limits of agreement and bias were calculated using the Standard Error multiplied by the correct t-distribution (*n-1*), sought from the 0.5 two-tailed column of a significance level chart.

Bland and Altman plots visually illustrate the agreement and differences between the repeated measures. Figure 10.3 indicates the differences between BH's repeated CSA measures for the whole cohort. With regards to data spread, all but one point on both graphs lie between the limits of agreement (LoA). While one point lies outside the defined acceptable error of ±2mm². The CI range are small at ±0.389 and the data indicates a weak positive trend, inferring that the larger the ONs appear to be, the more variation in the results. The bias was 0.324 (±0.4685) and the LoA were 2.3415 and -1.6935 (± 0.389). The average discrepancy between the measures indicate a very small

difference, thus can be regarded as acceptable. However, the LoA range appears quite large when compared to the average CSA of ON of 4.924mm².

Figure 10.4 indicates the differences between BH's repeated Diameter measures for whole participant cohort. The scale on this plot is different to the previous as diameter measurement are smaller, thus result in smaller means and differences. With regards to data spread, all but one point lie between the LoA and hover around the bias. All the points are within the specified error limit. There is no trend or indication of consistent variation on the plot. The bias was -0.0118 (±0.0498) and the LoA were 0.2027 and -0.2263 (± 0.0863). The average discrepancy between the measures indicate a minuscule, negative difference and the LoA range reflects this, not surpassing 0.25. The CI of all values are also very small, at most equalling ±0.09.

Figure 10.5 indicates the differences between BH and CB's CSA measures for 4 participants. With regards to data spread, all points lie between the LoA with two hovering around the bias. There is no trend or indication of consistent variation on the plot. The bias was 1.768 (±1.8963) and the LoA were 4.063 and -0.6095 (± 3.2846). As there are only four points, it is difficult to describe a data trend, but the LoA range is wide, implying the results are ambiguous. Moreover, the CI of all values are also larger than expected.

Table 10.2. *Results table for Bland Altman plots*

Measurement Type	n	Bias	Upper LoA	Lower LoA
Cross intra-observer-BH	21	0.324 (±0.4685)	2.3415 (± 0.389)	-1.6935 (± 0.389)
Diameter intra-observer-BH	21	0.0118 (±0.0498)	0.2027 (± 0.0863)	-0.2263 (± 0.863)
Cross inter-observer	4	1.768 (±1.8963)	4.063 (± 3.2846)	-0.6095 (± 3.2846)

95% confidence intervals shown in brackets, (Cross= cross-sectional area, n=number of scans assessed, LoA=limits of agreement)

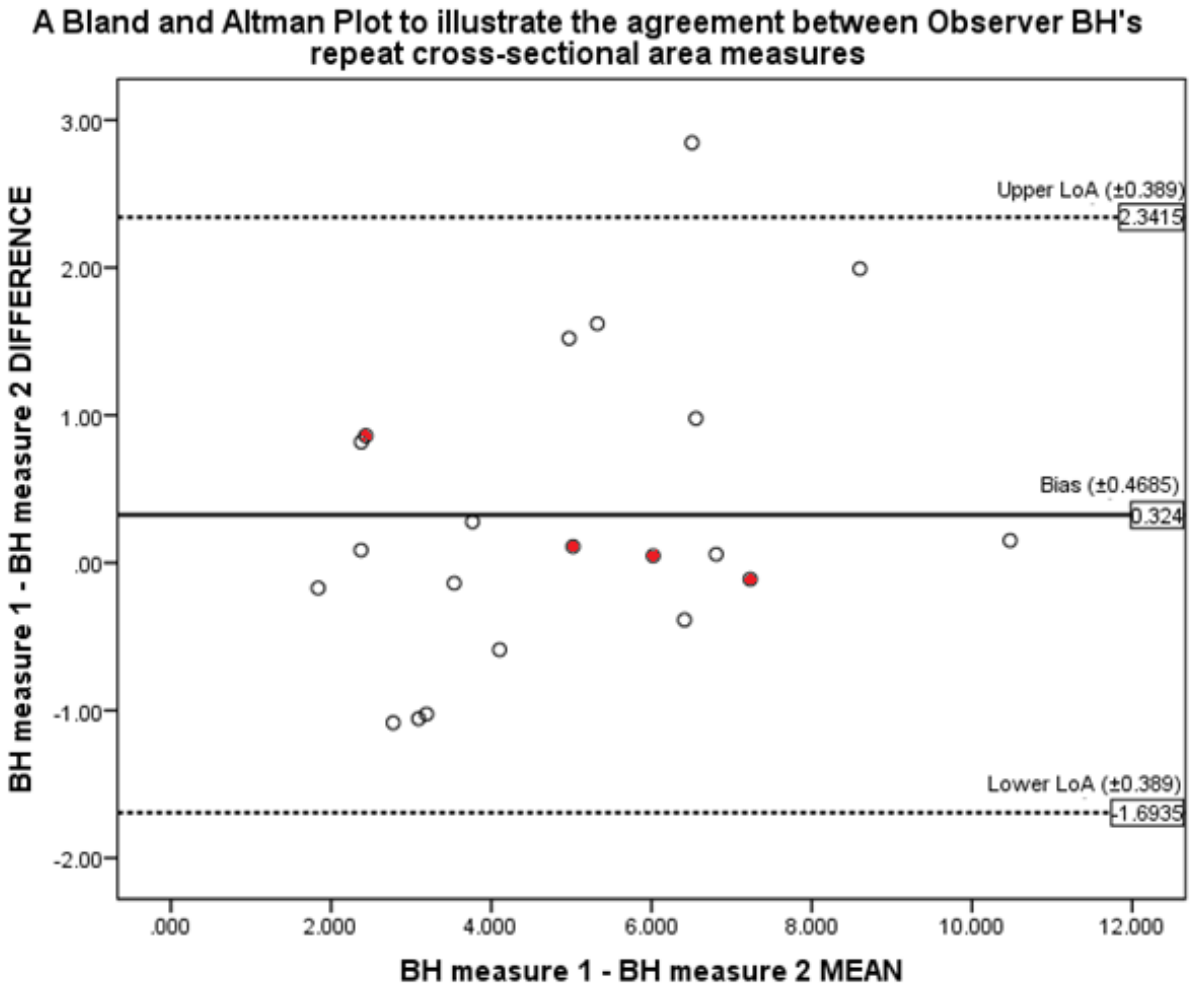

Figure 10.3 *A Bland and Altman plot illustrating the agreement between Observer BH's repeat CSA measurements. The data points for the 4 participants used in the subset analysis are highlighted in red.*

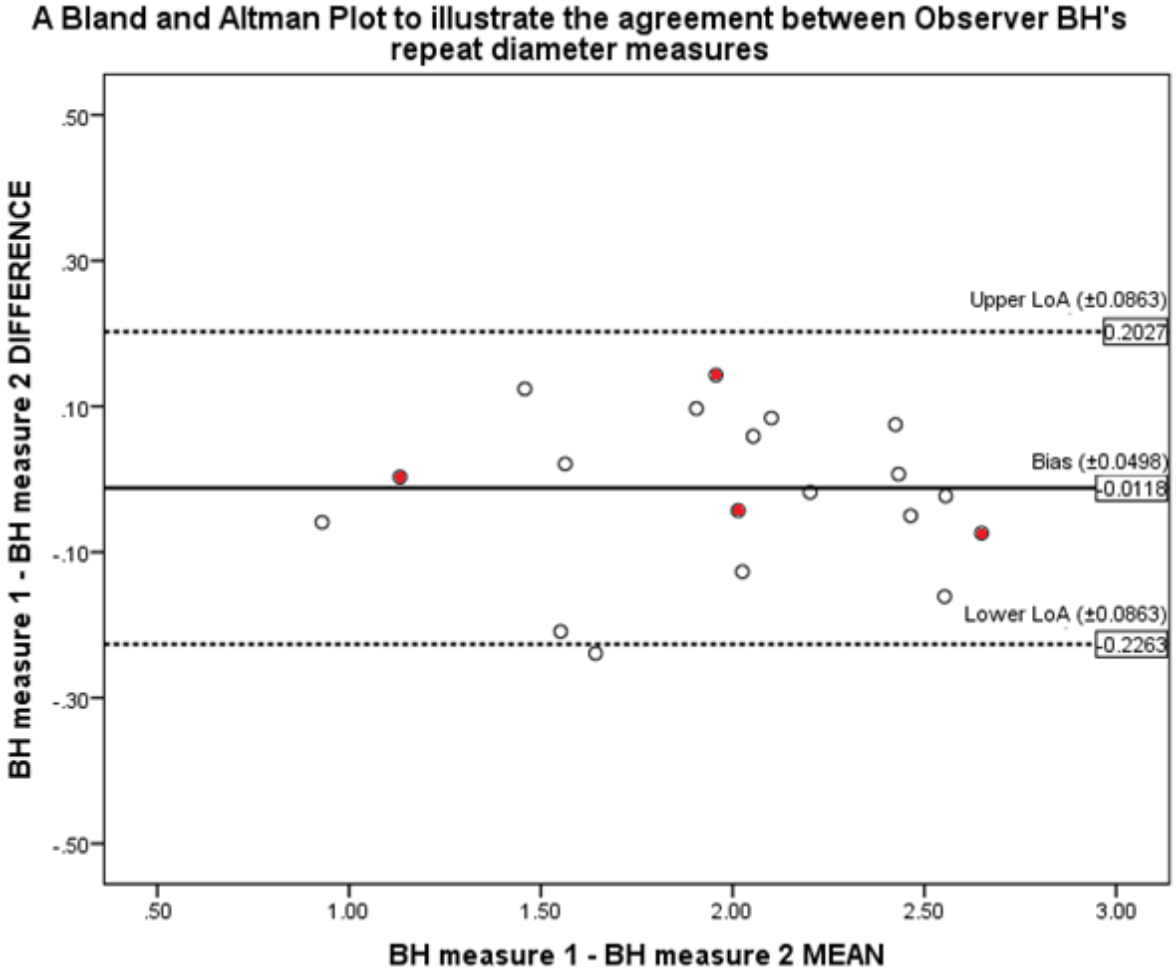

Figure 10.4 *A Bland and Altman plot illustrating the agreement between Observer BH's repeat diameter measurements. The data points for the 4 participants used in the subset analysis are highlighted in red.*

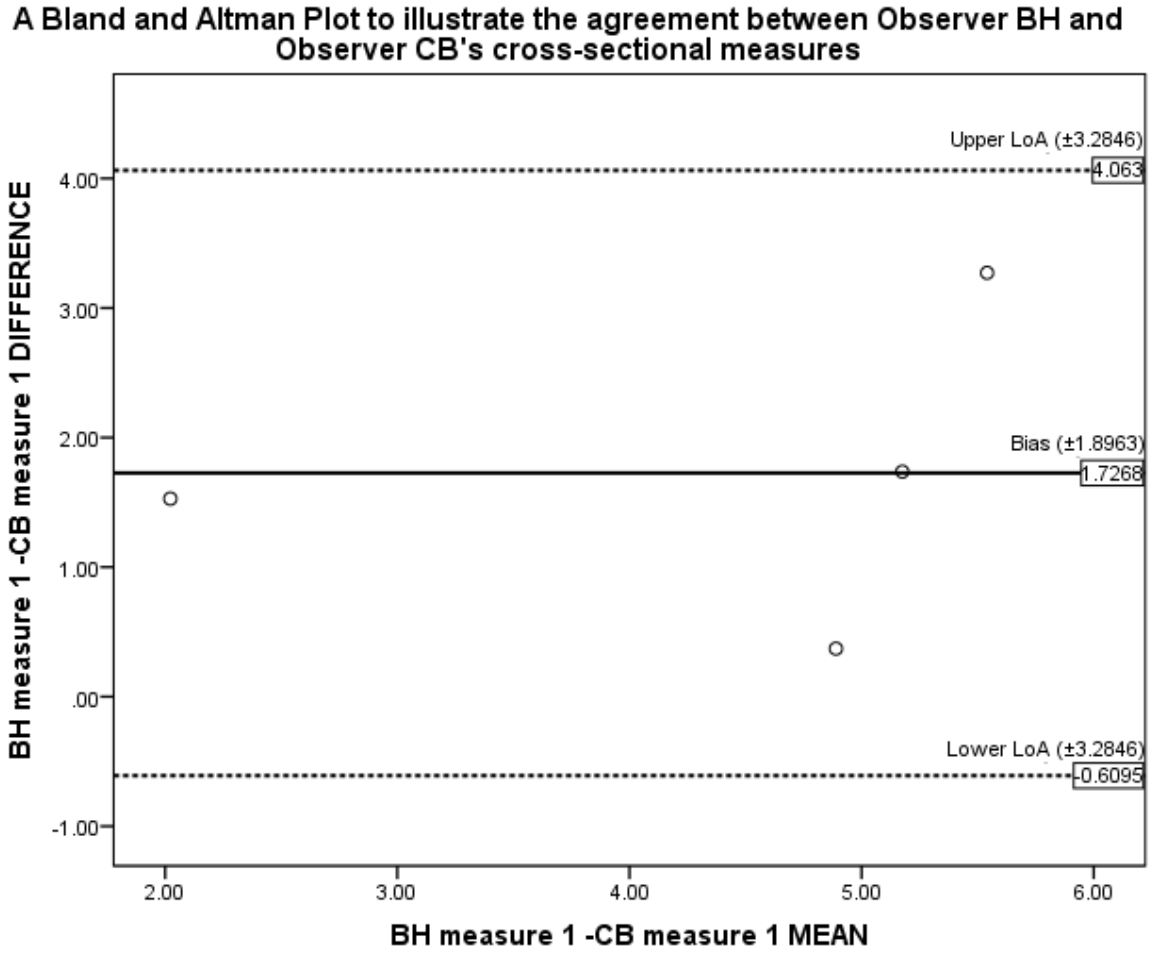

Figure 10.5 *A Bland and Altman plot illustrating the agreement between Observer BH and Observer CB's cross-sectional measurements.*

10.4 Discussion

10.4.1 Intra-Observer Variance of Cross-Sectional Area Measurements

For the assessment of all 21 participants, the ICC has the widest CI compared to the remaining intra-observer variance measures (0.772-0.958), inferring this ICC to be the least precise of the values. However, the ICC value is high and acceptable (0.9), coupled with an acceptable CV value (10.89%) and a low bias value (0.324), thus it can be inferred there's a good level of repeatability within the repeated measures of BH. It can be inferred that the reliability index illustrated more variation as the smaller sample had less variability by chance and that the larger re-test interval adopted by BH may have led to this variance. Figure 10.2 illustrates the relationship between BH's measures and identifies that the subset participant's data points in comparison to the rest of the cohort, demonstrating that these points do not indicate any unexpected variance.

For the subset of participants, the ICC values are high implying a high level of repeatability for both observers, with CB's index being slightly higher and more precise (0.994), with smaller CI (0.93-1), compared to BH's index (0.98 (0.815-0.999)), but overall both results are very similar, showing good repeatability. BH and CB also accrued very low, acceptable CV values (3.72%, and 2.81%), thus overall repeatability appears good. CB's mean difference was lower than BH's bias (0.0588 and 0.216, respectively). The minor difference between them could be explained by the larger re-test interval used by BH. BH also appears to estimate the ONS as larger compared to CB, indicated by the Mean ONS in Table 10.1. This variation can be attributed to the small sample size of 4 participants.

When compared to the existing literature, the reliability indexes appear consistent. For example, Hickman *et al*. (114,116) examined the CSA measure of diseased and contralateral nerves in optic neuritis sufferers. Hickman *et al*. (118) reported indexes of 0.96 (0.93-0.99) and 4.8% in diseased nerves and 0.84 (0.7-0.97) and 5.6% in contralateral nerves. While Hickman *et al.* (116) reported indexes of 0.92 (0.88-0.96) and 6.4% in diseased nerves and 0.89 (0.84-0.94) and 6.8% in contralateral nerves. Hence, the reliability indexes reported are consistent with the literature. Moreover, Hickman *et al*'s CSA method was conducted semi-automatically, hence the manual method exhibited in this report should show more variation due to user-error, further emphasising its reliability.

10.4.2 Inter-observer Variance of Cross-Sectional Area Measurements

This is the lowest ICC index (0.526), the highest CV (27.81%) and the widest CI reported (-0.151-0.953), indicating variation in the Observer's measurements thus implying the CSA measure appears to have a low level of reproducibility. This is reinforced by the ICC value being on the borderline of unacceptable variation, as the limit was defined as 0.5. This is also supported by the differences in the Observer's individual mean ONS results, with a difference of 1.648mm² and the highest bias recorded (1.768). Hence, either one or both CB and BH are not measuring the ONS accurately. Figure 10.5 illustrates the relationship between the two observer's measures and indicates a consistent positive bias as all the data points are above 0, demonstrating that one observer is overestimating the ONS consistently. This may be due to a segmentation difference.

Although, the low level of intra-observer variance indicates that the Observers are able to replicate their own findings thus are measuring consistently. However, as the aim of this report is to validate a reliable measurement method of the ON, with respect to both repeatability and reproducibility, it is clear more research is required. It should be noted that this variation was only measured from 4 participants, thus with more participants the reliability of the study may be indicated as better than indicated by this result. No CSA methods have been used on diseased ONs and had inter-observer variance measured within the literature, hence it cannot be validated.

10.4.3 Intra-observer Variance of Diameter Measurements

The inclusion of the diameter measurement was to compare the simpler measure to the CSA measurement. As it is a more straightforward procedure, it should indicate less variation. Despite the method accruing low reliability indexes (0.984 (0.961-0.993) and 2.86% for the whole cohort and 0.995 (0.929-1) and 2.15% for the subset) and the lowest standard deviations (0.060 and 0.048, respectively), they were not drastically different to the more informative CSA method's results. Therefore, it is more advisable to utilise the more complex method as it reflects the integrity of the ON as a whole.

With regards to the literature, the reported method appears to produce higher reliability index when compared Oyama *et al.* (128), who manually measured intra-observer variation of the height and width of the ON in healthy participants. Oyama *et al.* (130) reported 0.85 for the width averaged across the left nerves measured and 0.88 respectively for the right nerves. However, different patient groups were utilised for these studies so they are not directly comparable.

10.4.4 Bland and Altman Plots for Intra-Observer Variance of Cross-Sectional Area and Diameter

Table 10.2 and Figure 10.3 demonstrate the minimal differences between BH's repeated CSA measures for the 21 participants and indicate that the measurement method is reproducible by the same observer. There is no indication of a systematic bias evident in the graph. The CI range are small at ±0.389, indicating that the results appear to be precise. However, the LoA range appears quite large when compared to the average CSA of ON of 4.924mm². Although, this may be attributed to the larger sized nerves resulting in more variation. This is self-explanatory as a larger shape is open to more variation in its measurement. Moreover, it should be noted that CSA measures are inherently larger than diameter measures, which can explain some of the larger differences. Additionally, as the ON of children with OPGs are irregular, the process of measuring the CSA is user-dependent. Furthermore, the sample size of 21 is relatively small. To reliably ascertain if the results are ambiguous, as indicated by the wide LoA, a repeat of the study with a larger sample size should be completed.

Figure 10.4 and Table 10.2 indicate the differences between BH's repeated Diameter measures for whole participant cohort. The scale on this plot is different to the previous plot as diameter measurements are smaller, thus result in smaller means and differences. With regards to data spread, all but one point lie between the limits of agreement and hover around the bias. All the points are within the specified error limit and there is no trend or indication of consistent variation on the plot, hence systematic bias does not appear to be evident. The average discrepancy between the measures indicate a minuscule, negative difference and the LoA range reflects this, not surpassing 0.25. The CI of all values are also very small, at most equalling ±0.09. Hence, the repeated measures appear to be essentially equivalent. This supports that diameter measures appear to be a reliable measurement method. However, it still remains true it but may be unsuitable for measuring potentially tortuous ONs of OPG patients.

Figure 10.5 and Table 10.2 indicate the differences between BH's and CB's CSA measures for a subset of 4 participants. The Bland Altman plot indicates a consistent positive bias of 1.768 with CI of ±1.8963. The four data points are above 0 on the y axis, demonstrating that one observer appears to over or underestimate the size of the ONS consistently. If we compare this to the mean ONS, BH measures the ONS of the four participants as larger than CB, by 1.648mm². This shows there's a systematic difference in the measurements taken by the two observers which may be due to a segmentation difference.

10.5 Summary

To summarise, both the CSA and diameter measurement methods are reliable, but the CSA method offers more in measuring ON integrity as the diameter method does not consider the entire shape of the ON which can be malformed in children with OPGs, thus making it unsuitable for measuring this particular cohort. However, this study has been limited by the small sample size chosen to measure the inter-observer variance of the CSA method, resulting in data that may not be truly representative of the repeated measure agreement. Hence, if this study were to be repeated, a larger sample size should be used and both observers should individually assess the whole cohort, opposed to a subgroup.

11.0 Normal Variation of Optic Nerve Size and Visual Acuity in Healthy Children

11.1 Introduction

The following research questions are addressed: what is the variation of ONS and VA in healthy children and does ONS correlate with to VA, height and/or age? The ONS of a cohort of healthy children was assessed using the evaluated CSA method tested previously. To correlate ONS and VA, the participant's VA was assessed using the Sheridan and Gardiner (SG) test on the same day as an MRI scan incorporating a 3D CISS sequence was completed. These values were then correlated.

11.2 Method

11.2.1 Setting and Study Design

This investigation has an observational study design, aiming to assess the correlation between ONS, age, height and VA in healthy children, and to evaluate the variation in ONS in children aged 5-16years. This study used data from the Normal Variability Evaluation for MRI in children and young people (NOVEL-MRI) study, Chief Investigator Dr RA Dineen, funded by The A-T Children's Project, University of Nottingham Medical School REC L14112016. The conduct of the study ensured protocols of Good Clinical Practice were adhered to. It was conducted within the Radiological Sciences Group in the School of Medicine, The University of Nottingham.

11.2.2 Participants and Recruitment

Recruitment was through posters and social media postings to local community groups in accordance with the REC approved protocol [Appendix B]. The inclusion criteria included healthy children aged 5-18 years, without visual deficits (disregarding refractive error) or cognitive disabilities. If children wore glasses, they were asked to bring them. Consent was collected on the day from parents/guardians or from the young person if aged ≥16.

31 participants were recruited by BH: 17 females and 14 males (5-16years, mean=10years). 8 suffered from refractive error but wore corrections. Participants were allocated a numerical label and all collected data was affiliated to this code.

11.2.3 MRI protocol

Imaging was conducted by a registered radiographer on a 3T GE MRI scanner, based in the Sir Peter Mansfield Imaging Centre, Nottingham Medical School, Nottingham. A 3D CISS sequence was chosen as it is a fast, 3D structural sequence that gives a high resolution of the anterior VP within the small FOV. The parameters included: TR= 6.351ms, TE=2.288ms, a 32Ch head-coil, Flip-Angle=35, Bandwidth=162.754Hz, matrix 512/352, echo-train=1, slice thickness 1.2mm, slice overlap -0.6mm and isotropic pixel size=0.166. Scan time for the sequence was 12 minutes 15 seconds.

Before the child entered the scan room, each had the opportunity to practice in a mock scanner. The child was informed how MRI worked and the importance of not moving. Once comfortable, the child entered the scanner. MRI suitable corrections were supplied and the child's weight and height were recorded. The child was able to watch a film-clip of their choosing as a projector was erected in front of the bed, this could be watched via a small viewing-mirror reducing eye-movement, while rendering the child ideally preoccupied thus motionless. As a reward, the child was gifted a picture of their brain following the session. If the resulting scan was deemed unusable due to motion artefacts, the child's data was excluded.

11.2.4 Cross-Sectional Area Assessment of Optic Nerve Size

The method was conducted, as previously described in section 10.1.2. As the participants had healthy, corrected vision and as eyes are a paired structure thus values would not be statistically independent, one eye was randomly selected from each control to measure (144).

11.2.5 Visual Acuity Test Protocol

Immediately before the scanning session, the children's VA was assessed by BH using the SG distance and near test. This test type was chosen as it allows the testing of both pre-literate and literate children without changing test types – ideal for this study. The distance test was conducted at 6m and started with the brown booklet with optotypes sized 6/60-6/6 in Snellen letters. Depending on how well the child performed, a second booklet was chosen, either: green 6/6 to 6/3, red 6/60-6/18 or yellow 6/18-6/6. The child was supplied with a lap card, thus enabling pre-literate children to point to the optotype. The near SG test was conducted using the blue booklet, held at a comfortable reading distance for the child, as directed by test guidelines.

The eyes were assessed separately, as to collect values from each eye. This was achieved by asking the child to manually cover their contralateral eye. For younger children, their parent/guardian was asked to do this. During testing, the investigator ensured to give praise to the child to encourage cooperation. However, if the child was uncooperative or was unable to complete the SG test, their results were excluded. Data from the same eye selected to measure the ONS was recorded and changed from Snellen to logMAR.

11.2.6 Statistical Tests

SPPS and STATA statistics packages were used. VA indicates distance visual acuity, while VANear indicates Near visual acuity. Pearson's bivariate correlations were conducted to identify if a correlation exists between the variables. Scatterplots were used to illustrate the relations between variables. 95% CI were calculated for the correlations, using Fisher's z transformation.

11.3 Results

11.3.1 Participants

16 females and 14 males were included in the VA section of this study (5-16years, mean=10years), [see Figure 11.1]. 1 exclusion from distance VA and 2 exclusions from Near VA assessments were due to lack of cooperation. 17 left and 13 right eye values were randomly selected from these participants.

16 participants were excluded from the MRI section of this study due to unusable scans following excessive motion artefacts. 15 healthy participants were included in the study: 6 males and 9 females aged (6-16years, mean=11years). 7 left and 8 right eyes were randomly selected from these participants.

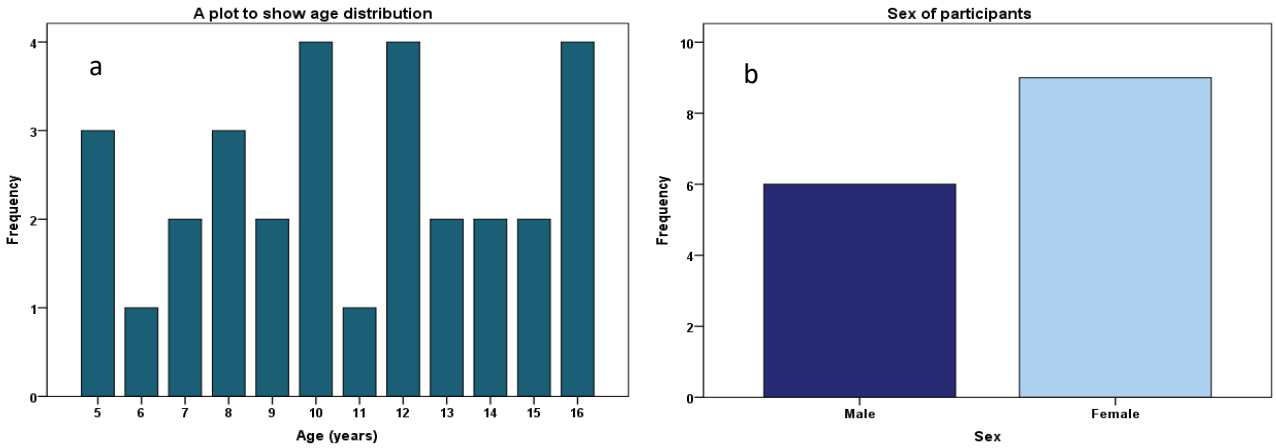

Figure 11.1 *illustrating the frequency of age range (a) and sex (b) in the whole cohort.*

11.3.2 Normality

All data was assessed for normality by visually inspecting histograms, q-q plots and Shapiro-Wilks test. All data appears either normal or deemed acceptable for inclusion.

11.3.3 Descriptive statistics

As evident from Table 11.1, the mean ONS of the participants was 8.798mm² (5.99-13.85mm², SD=2.17). The mean distance VA of the whole cohort was -0.27logMAR (-0.3 to -0.1logMAR, SD=0.05) and for NearVA was 0.01logMAR (-0.2 to 0.3logMAR, SD=0.05). While the mean distance VA of the subset was -0.29logMAR (-0.3 to -0.2logMAR, SD=0.13).

Table 11.1. *Table of descriptive statistics: sample size, range, mean and SD*

	N	Minimum	Maximum	Mean	Std. Deviation
ON Size	15	5.99	13.85	8.7977	2.17035
VA (subset)	15	-0.30	-0.20	-0.2867	0.03519
VA (whole cohort)	30	-0.30	-0.1	-0.27	0.0535
VANear (whole cohort)	29	-0.2	0.3	0.01	0.1319
Weight	15	19.95	61.55	38.9567	14.37286
Height	15	123	181	149.19	17.475

11.3.4 Variation of Visual Acuity

There was a lack of variation in the distance VA data, as one participant scored -0.1logMAR, while the remaining 29 participants acrued either -0.3logMAR or -0.2logMAR. Hence, correlational analysis between distance VA and age was not possible, but a scatterplot was produced to illustrate this lack of variance [See Figure 11.2]. There was no significant correlation between Near Visual Acuity and Age: r=-0.095 (-0.446 to 0.282) n=29 p=0.625 (2-tailed). [See Figure 11.3].

11.3.5 Variation of Optic Nerve Size

A significant, negative correlation was observed between age and ONS: r=-0.629 (-0.863 to -0.172) n=15 p=0.012 (2-tailed) with R^2=0.396, indicating moderate relation strength [See Figure 11.4]. The correlation between age and ONS was not significant when height was included as a covariate: r=-0.082 (-0.570 to 0.449) n=15 p=0.772 (2-tailed) [See Figure 11.5].

A significant negative correlation was also observed between Height and ONS: r=-0627 (-0.862 to -0.169) n=15 p=0.012 (2-tailed) R^2=0.393, indicating moderate relation strength [See Figure 11.6].

11.3.6 Distance Visual Acuity vs Optic Nerve Size

A Pearson's correlation between ONS and VA was not attempted due to the lack of variation in the data [See Figure 11.7].

Figure 11.2 *Scatterplot illustrating the significant correlation between Age and VA*

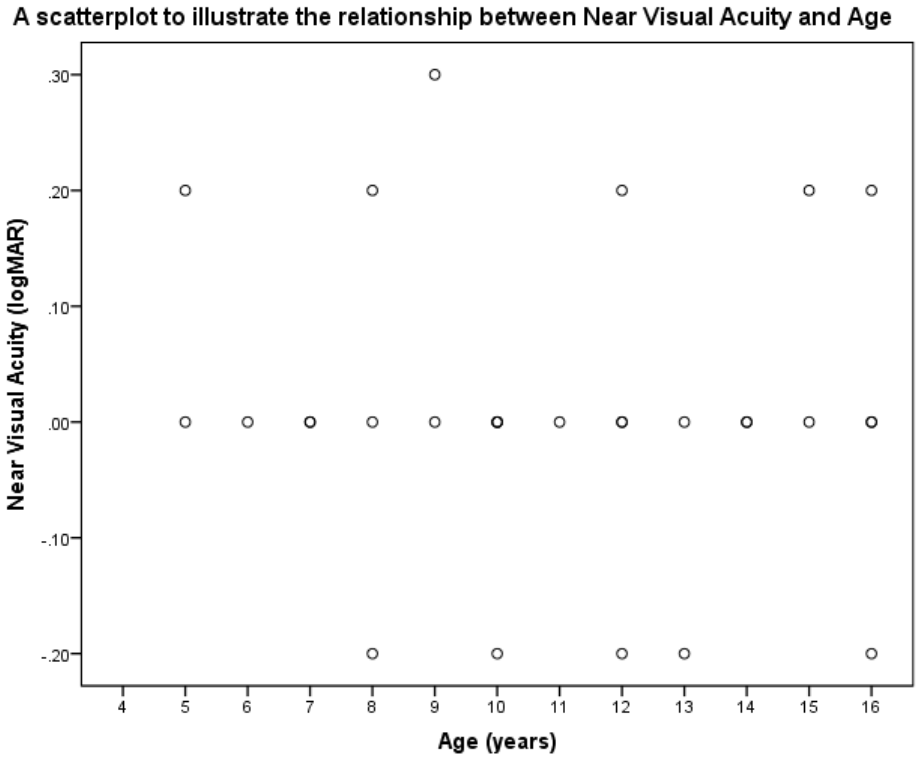

Figure 11.3 *Scatterplot illustrating the relation between Age and NearVA*

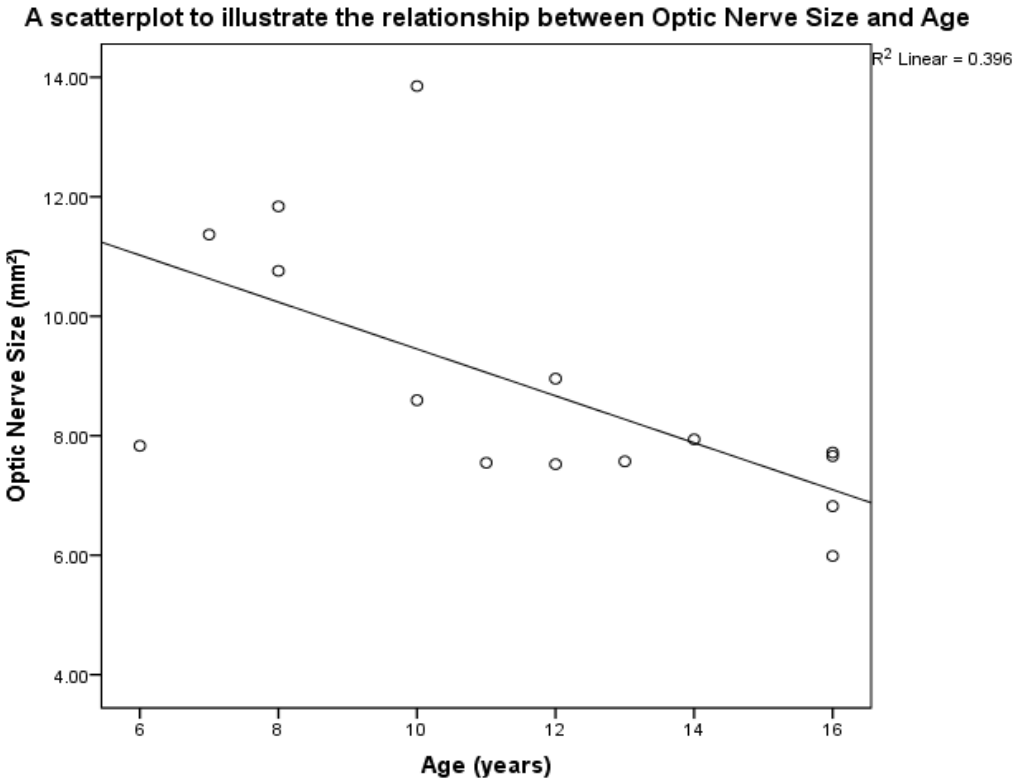

Figure 11.4 *Scatterplot illustrating the relation between ONS and Age*

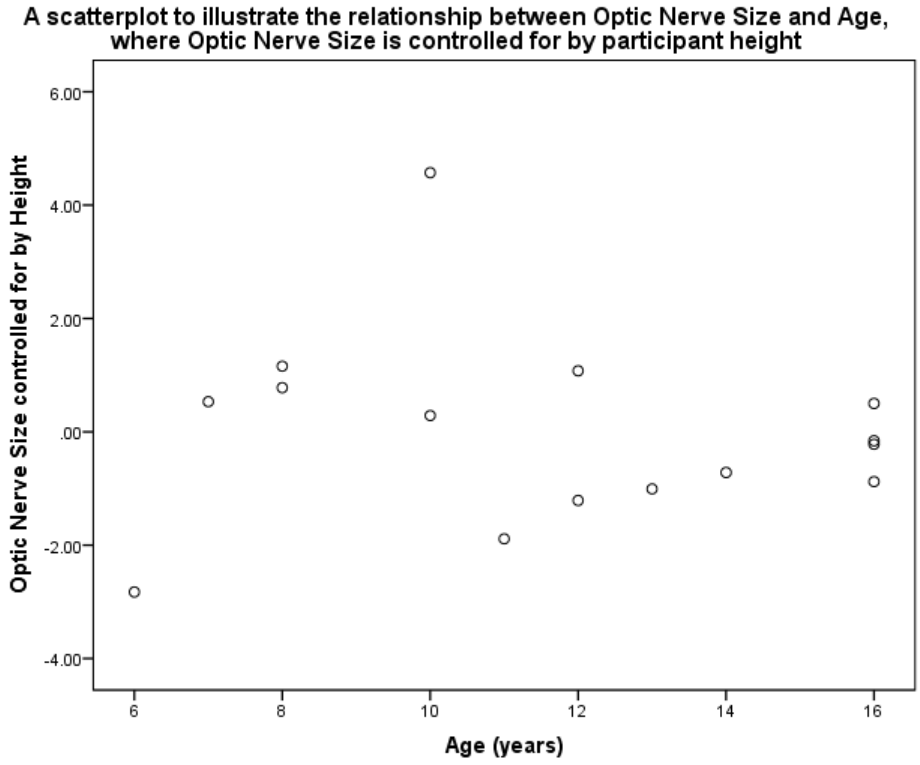

Figure 11.5 *Scatterplot illustrating the relation between ONS controlled for by Height, and Age*

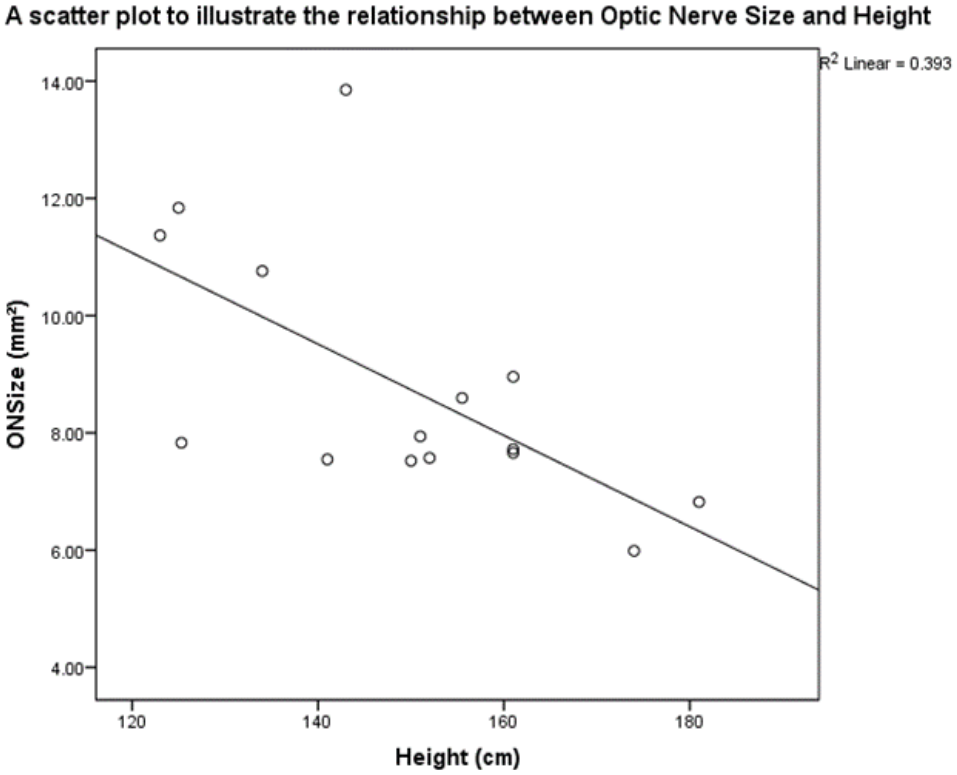

Figure 11.6 *Scatterplot illustrating the negative correlation between ONS and Height*

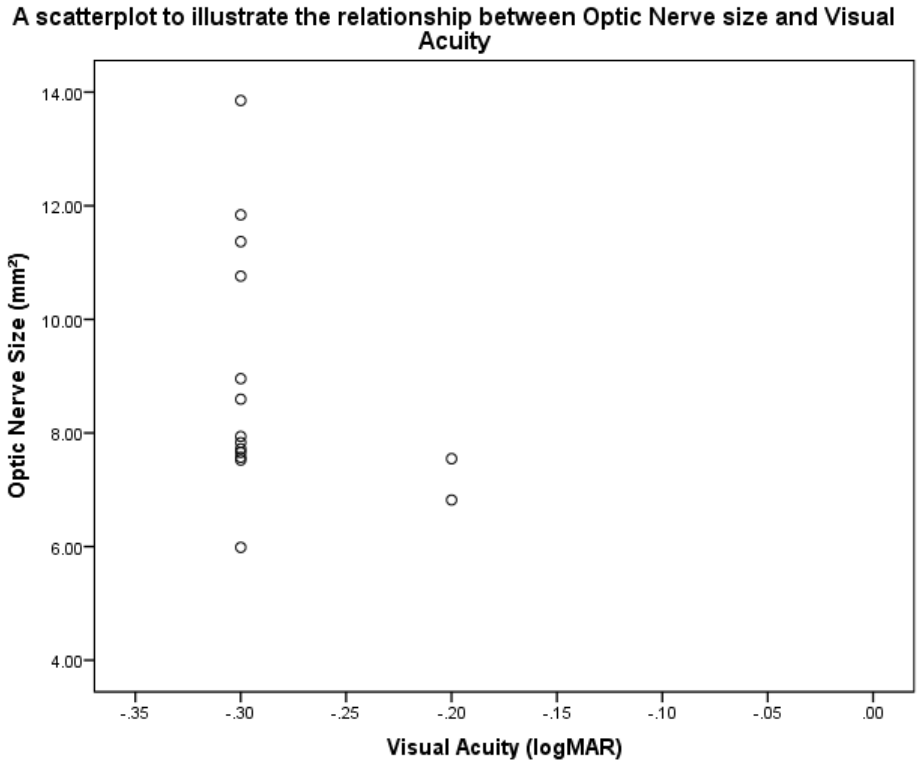

Figure 11.7 *Scatterplot illustrating the relation between ONS and distance VA*

11.4 Discussion

11.4.1 Variation of Visual Acuity

Lack of variation in the distance VA data meant that correlational analysis could not be conducted. However, Figure 11.2 clearly illustrates this lack of variation, which indicates that a consistent, good level of acuity was evident across the participant cohort. Moreover, the poorest VA score of -0.1logMAR was accrued by the youngest aged participant: aged 5years. This is supported by the literature, in that a healthy child's VA improves until the age of approximately 6-8 years, after which it plateaus (10). Although, as the nature of the SG test measures in 0.1 logMAR units, it is possible that using a VA test measuring in smaller units such as an adult VA test may illustrate a linear correlation between distance VA and Age.

A significant correlation is not seen between VANear and Age, although this may be due to the small sample size used. However, more variation is evident in this VA dataset compared to the distance VA data. The data spread on Figure 11.3 indicates this greater range in values, from -0.2 to 0.3logMAR. These scores may vary more due to difficulty in retaining a standard viewing distance between children, as directions for NearVA assessment indicate that the test should be conducted at a comfortable reading distance, despite this varying between-subjects.

Moreover, the literature indicates that in 8-13 year olds, NearVA tests utilising letter acuity, like the type used in this study, yield scores 0.1-0.2logMAR worse than picture acuity (146). Thus, inferring that if a picture-type NearVA test was used such as a Lea symbols test, a correlation may have been evident between VA and Age. However, this is contradicted by the fact that letter charts are still advised for testing NearVA in older age-groups as it is reflective of functional vision.

11.4.2 Variation of Optic Nerve Size

The mean ONS of the cohort was 8.798mm², ranging from 5.99mm² to 13.85mm². These values are comparative to the literature on the CSA of healthy ON, despite the studies featuring older age-groups. For example, Yiannakas *et al.* (137) measured the CSA of the ON of 31 healthy participants with a mean age of 31years (29-33years) and found that on the section of the ON 0-3mm from the bulb, the average size was 7.8mm², and this decreased to 6.2mm² at 3-6mm from the optic bulb. The measurements in this report were taken at 5mm from the optic bulb, hence the results score here appear higher, but not disproportionate (137).

Trip *et al.* (2) used a semi-automated method from Hickman *et al.* (113) to measure the mean CSA of the whole intra-orbital portion of ONs in 15 healthy controls with a mean age of 36years (30-56years). Hence the results would have been subject to the same small sample size limitations as this report. The mean CSA size was found to be 12.7mm² (SD=2.4) which again is not hugely different to the results exhibited in this paediatric cohort. The differences between the studies can be explained by the fact that there has been found to be variation in healthy ONS in different age cohorts (113–116).

11.4.3 Optic Nerve Size and Age

A significant, moderate (R^2=0.396) negative relation was observed between Age and ONS. A correlation between the two variables was hypothesised, but a positive correlation was expected as Age would invariably reflect visual system maturity and as the children got older and grew bigger, the ONS would be expected to mirror this. However, in this data set the opposite is portrayed. Figure 11.4 illustrates the relation whereby, two data points do not the follow the trend. This correlation may be explained by the natural variation of ONS in a healthy cohort (113), represented in children. However, as a negative correlation was also found between ONS and height, this appears to be a genuine finding which could be explained by nerve loss as a function of childhood maturation (147). This is supported by Patel (2014) who demonstrated age-related changes in RNFL and ON head parameters in a healthy cohort aged 19-76years (148). Although, this finding would be the first demonstration of this phenomena in a paediatric cohort.

However, when the correlation was tested again between Age and ONS controlled for by Height, to ascertain if the overall size of the child made an impact, the significant correlation was lost [See Figure 11.5]. The plot illustrates variance between the data points, and no trend is evident. This would therefore imply that when overall size of the child is accounted for, age per se does not appear to have a direct relationship to the resulting ONS.

11.4.4 Optic Nerve Size and Height

Another surprising, significant moderate (R^2=0.393) negative relation was observed between Height and ONS, mirroring that of the ONS and Age correlation. It was expected that, as with Age, the maturation of the anterior VP would correlate with an increase in ONS. However, this data illustrates the reverse, illustrated by Figure 11.6. The scatterplot indicates that most of the data follows this negative trend excluding one data point, hence identifying consistency within the data. As

previously, this data may possibly be explained by natural variation in ONS in healthy cohorts but it is likely that it indicates a realistic correlation between these factors not seen before. Further investigation is required to validate this.

11.4.5 Optic Nerve Size and Visual Acuity

It was hypothesised that there would be a correlation between ONS and VA. Although, due to the lack of variation in the data illustrated in Figure 11.6, correlational analysis would have been inappropriate to perform and ultimately uninformative. Therefore, analysis and subsequent interpretation of the data from this normative healthy cohort cannot be conducted following the limitation of the lack of variation in VA scores.

11.4.6 MRI and the paediatric cohort

The success rate of acquiring useable MRI scans from the young cohort proved challenging. Despite efforts to familiarise the children with the scanner using a mock scanner and incorporating the use of visual stimulation of the child's choosing in an effort to increase cooperation during the scan, only 50% of the scans were useable, due to motion artefacts. This may have been due to a number of factors. By watching the videos, the child's eyes may have been moving. Other studies have incorporated a fixation cross to reduce this (143). However, it was felt that the small viewing-mirror restricted eye movement to a minimum and that using a cross in a young cohort may have forfeited cooperation.

Moreover, the entire scan length was 46minutes and 29seconds and the CISS scan for this study was at 16minutes 27seconds into the session. It is possible that (particularly the young) children grew tired and uncomfortable, thus exhibited motion. To improve upon this study, a CISS sequence should be conducted earlier into the scanning session with possible use of a fixation cross for the duration of the sequence to limit motion effects.

Although 'failure' to acquire usable ON data in around half of participants has limited this analysis and is a frustration, this result is in itself extremely useful; it tells us that use of this sequence in awake child participants is likely to have at best a moderate success rate, and faster or more motion resistant sequences may be required for application in clinical populations undergoing awake scanning for detailed ON measurements.

11.5 Summary

In this study, the ONS was found to vary in the paediatric cohort and thus reflected the literature for adult ONS. ONS was expected to correlate with Age, which was found to be the case, but in a negative fashion opposed to the expected, positive. This trend was mirrored by the relation between ONS and Height. As both of these correlations are negative, this implies that these are genuine findings that may be explained by age-related changes due to maturation, a phenomenon seen in mature cohorts but not paediatric. Hence, more investigation is needed to validify if this finding is reflective of the general paediatric population.

A correlational analysis could not be conducted between distance VA data and the remaining variables as there was little to no variation between the results. However, the data clearly identifies that a cohort with consistent, good vision was assessed. Further investigation with a larger sample size should be conducted to assess the correlation between distance VA and Age and ONS. A correlational analysis was conducted between NearVA and Age, but a significant correlation was not found. Again, this may be due to the small sample size used.

Lastly, the MRI sequence used was degraded by motion in 50% of cases which reduced the sample size available for correlational analysis. Hence a shorter or more motion-resistant sequence will be required for further clinical studies in awake children.

12.0 Does Optic Nerve Size and Retinal Nerve Fibre Layer Thickness Correlate with Visual Acuity in Children with Optic Pathway Gliomas?

12.1 Introduction

The aim of this section was to investigate the following research questions: Does ONS, RNFLt and VA correlate in children with OPGs? In what way does ONS vary in relation to Age in children with OPGs? Does ONS and RNFLt predict VA in children with OPGs? And does ONS and VA in healthy children significantly differ to ONS and VA in children with OPGs? To do this, the ONS of children with OPGs measured in Section 10 was compared to the ONS of healthy children collected in Section 11. Moreover, retrospective visual examination data including OCT results of children with OPGs were assessed and compared to VA results from healthy children.

12.2 Method

12.2.1 Setting and Study Design

This investigation has a retrospective, observational, case and case-control study design, aiming to assess the above research questions. Ethical approvals for the OPG and healthy volunteer studies are as described in sections 10.2.1 and 11.2.1 respectively.

Inclusion criteria for the study for healthy participants is detailed in section 11.2.2. Inclusion criteria for children with OPGs comprised of a diagnosis of a low-grade optic pathway glioma and having either appropriate MRI scans (see section 10.2.1), VA assessment results with the test-type defined and/or OCT assessment results.

12.2.2 Previous Data Collection

Details of measuring CSA of ON, see Section 10.2.2. Information regarding the visual assessment of the healthy cohort is detailed in Section 11.2.5.

12.2.3 MRI Parameters

See Section 10.2.1 for parameters of retrospective scans collected, and Section 11.2.3 for parameters of control scans conducted

12.2.4 Retrospective Visual Data Collection

NHS databases were searched by BH for visual assessment data. Records of VA assessments were recorded from UNITY and NOTIS. Average RNFLt data was sourced from OCT results, accessed via the Heidelberg Eye Explorer system. All OCT data was collected by BH using either an OCT SPECTRALIS or a HRA OCT SPECTRALIS. The average RNFLt data was selected as it infers the overall structure of the entire RNFL. All data was anonymised and participants could only be identified via the numerical label allocated to them.

12.2.5 Optic Pathway Glioma Location

The location of the OPGs was quantified through use of the MDC, issued by a single consultant neuroradiologists working with the Children's Brain Tumour Research Centre, Nottingham. This data was collected as part of a Service Evaluation of Imaging Outcomes of Paediatric Visual Pathway Glioma.

12.2.6 Statistical Tests

SPSS and STATA statistics packages were used. VA indicates distance visual acuity. Pearson's bivariate and partial correlations and Spearman's correlations were conducted to identify if a correlation exists between the variables. Scatterplots were used to illustrate the relations between variables. 95% CI were calculated for the correlations. Multiple linear regression was used to identify if ONS and RNFLt are significant predictors of VA. An independent-samples t-test was used to assess if there was a significant difference between ONS and VA of participants with OPGs and healthy controls.

The dates of the visual assessments and MRI scans included in the study are within 6 months of each other, to ensure reasonable temporal correspondence. If this was not possible, e.g. a child has only one retrospective MRI scan conducted in 2017 and visual assessments dated prior to 2016, the participant was removed from that comparison. If a participant had an OCT conducted, the corresponding VA data from the same day was used. If a participant did not have an appropriate MRI scan, the most recent visual assessment data was used. If the VA test-type was not specified, the score was excluded, and the next most recent assessment was used. The following numerical scores were recorded for low vision notations: NLP=2.48logMAR, HM=2.1logMAR, PL=2.0logMAR and CF=1.7logMAR [see Table 6.1].

As eyes are a paired structure and their individual values may not be statistically independent, only one eye is used for control participants as their eyes are equivalent. For details of which eyes were selected for the control group, see Section 11.2.1. The eyes of the OPG cohort may not be equivalent due to tumour location. As tumour location will not be considered during statistical examination but rather used as a reference, correlational analysis will be conducted on both eyes together. However, the individual eye correlations will be illustrated separately on corresponding scatterplots to identify any differences. Additionally, ONS will not be controlled for by height as detailed in Section 11, as details of height at the time of the retrospective scans was not collected.

12.3 Results

12.3.1 Participants

The entire cohort of children with OPGs consisted of 27 children, 17 males and 10 females (2-15years, mean=8years), 12 were NF1-positive. Together, the entire cohort amounted to 42 children [See Figure 12.1]. For details regarding the healthy control participants, see Section 11.3.1. [Figure 11.1].

For the ON and RNFLt layer comparison, there were 10 participants: 7 males and 3 females (6-15years, mean=9years), 5 were NF1-positive. For the ON and VA comparison, there were 15 participants: 9 males and 6 females (2-15years, mean=8years), 6 were NF1-positive. For the VA and RNFLt layer comparison, there were 18 participants: 12 males and 6 females (6-15years, mean=9years), 12 were NF1-positive. For the multiple regression of all three variables, there were 10 participants: 7 males and 3 females (6-15years, mean=9years), 5 were NF1-positive.

For the comparison of VA scores from both the OPG cohort and the controls, 30 healthy controls were included: 16 females and 14 males (5-16years, mean=10years) and 25 children with OPGs: 16 males and 9 females (2-15years, mean=8years). For the comparison of ONSs, 15 healthy controls were included: 9 females and 6 males (6-16years, mean=11years) and 21 children with OPGs: 13 males and 8 females (2-15years, mean=8years)

OPG011, OPG012 and OPG014 were excluded from comparative statistical analysis of ONS. This was due to inconsistent ONS measurements for the left eyes because of extremely tortuous ONs which may be a result of tumour involvement. OPG003, OPG010 and OPG022 were removed from comparative statistical analysis as appropriately dated visual assessment data was not available.

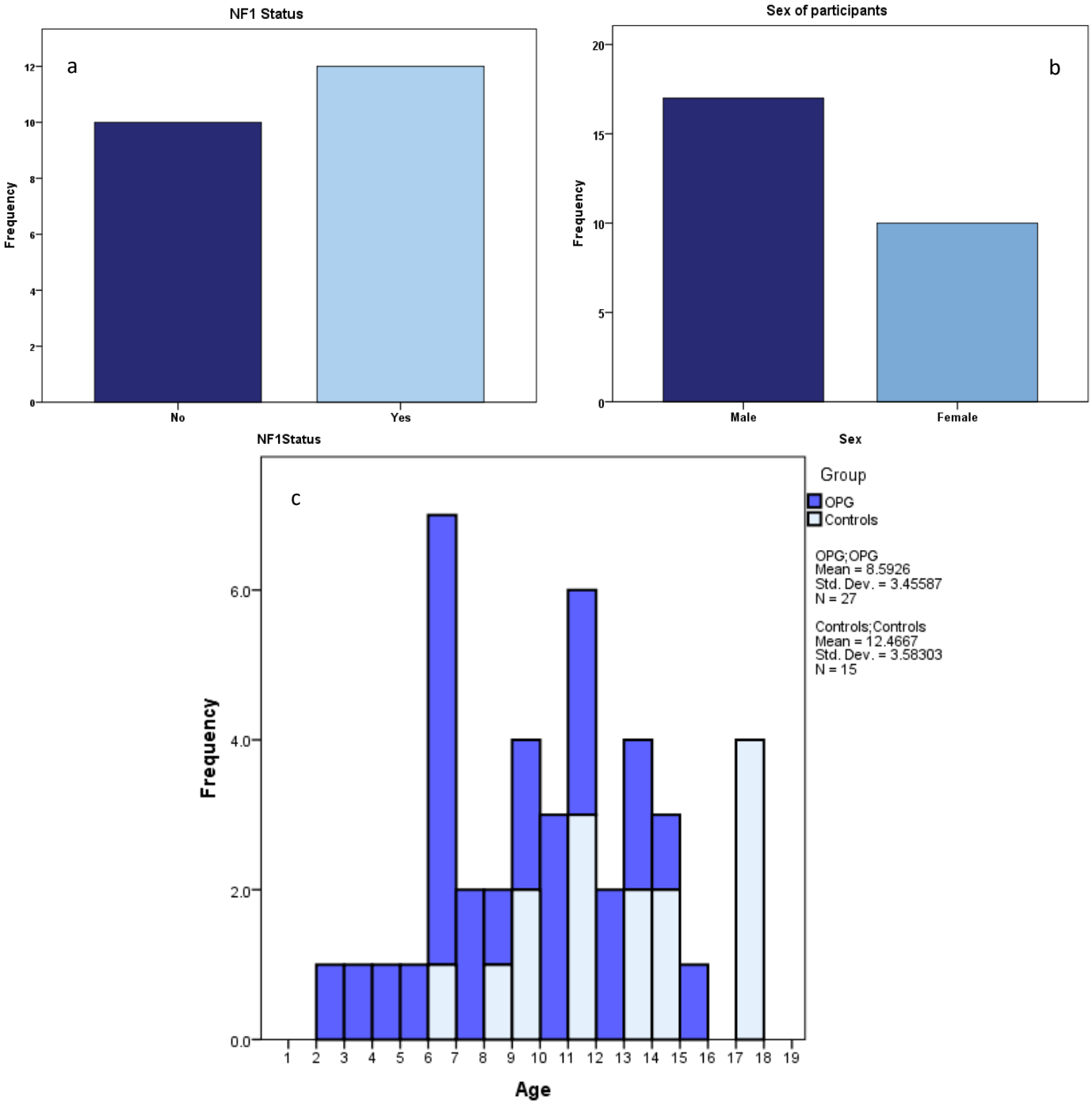

Figure 12.1 *illustrating the frequency of NF1 status (a) sex (b) and age range (c) of both participant cohorts*

12.3.2 Normality

All data was assessed for normality by visually inspecting histograms, q-q plots and Shapiro-Wilks test. All data appears either normal or deemed acceptable for inclusion, except the left eye RNFL data. Hence, this data set and its comparisons were assessed using Spearman's correlation.

12.3.3 Visual test-types

There were 3 test-types in the featured 18 retrospective visual assessments: 8 Bailey Lovie tests, 2 Crowded Kay tests and 8 Crowded Keeler tests. Making 16 letter-based tests and 2 picture-based tests.

12.3.4 Optic Pathway Glioma data vs Control data

The independent-samples t-test indicated that when comparing ONS, there was a significant difference between the participants with OPGs (M=5.63, SD=2.926) and controls (M=8.8, SD=2.17), t(52)=-3.797, p=<0.005 (2-tailed). Moreover, when comparing VA scores, there was a significant difference between the participants with OPGs (M=0.81, SD=0.816) and controls (M=-0.27, SD=0.054), t(78)=-7.201, p=<0.005 (2-tailed)

Figures 12.2 illustrates a five-number summary of the ONS data. For the OPG group, the minimum=1.18mm², Q1=3.7mm², median=5.4mm², Q3=7mm², maximum=10.6mm² and there were two outliers: OPG007 and OPG018. For the control group, the minimum=6mm², Q1=7.6mm², median=7.8mm², Q3=9.8mm², maximum=11.8mm² and there was one outlier: Control19.

Figures 12.3 illustrates a five-number summary of the VA data. For the OPG group, the minimum=-0.2logMAR, Q1=0.1logMAR, median=0.55logMAR, Q3=1.2logMAR, maximum=2.48logMAR. For the control group, the minimum=-0.3logMAR, Q1=-0.3logMAR, median=-0.3logMAR, Q3=-0.2logMAR, maximum=-0.1logMAR.

12.3.5 Optic Nerve Size and Age

ONS of both cohorts were included in the correlation with age and there was no significant correlation between them: r=0.111, (-0.161 to 0.368) n=54 p=0.423 (2-tailed). When ONS and age from the OPG cohort were assessed separately, no significant correlation was found: r=0.06, (-0.261 to 0.368) n=39, p=0.719 (2-tailed) [See Figure 12.4]. Details of the significant correlation between ONS and Age in the control group is in Section 11.2.4.

12.3.6 Optic Nerve Size vs Visual Acuity

No significant relation was observed between ONS and VA: r=-0.273 (-0.576 to 0.097), n=30 p=0.145 (2-sided) when data from both eyes was included in the correlation. When the two eyes were analysed

separately, Figure 12.5 illustrates a significant, moderate negative correlation (R^2=0.360) between the left eye VA and ONS that is not evident in the right eyes.

12.3.7 Retinal Nerve Fibre Layer Thickness vs Visual Acuity

A strong, negative correlation was observed between RNFLt and VA: rho=-0.600 (-0.776 to -0.338), n=36 p=<0.0005 (2-sided) Figure 12.6 illustrates a significant, moderate negative correlation of both the left (R^2=0.380) and right (R^2=0.394) eye between VA and RNFLt.

12.3.8 Retinal Nerve Fibre Layer Thickness vs Optic Nerve Size

No significant relation was observed between RNFLt and ONS: rho=0.290 (-0.175 to -0.649), n=20 p=0.215 (2-sided) when both eyes were included in the analysis. When analysed separately, Figure 12.7 indicates a significant, strong positive correlation (R^2=0.533) between the left eye RNFLt and ONS that is not evident in the right eyes.

12.3.9 Does Retinal Nerve Fibre Layer Thickness and Optic Nerve Size predict Visual Acuity?

RNFLt and ONS statistically significantly predicted VA: F(2, 17) =5.712, p=0.013, R^2 = 0.402, demonstrating a moderate relationship strength. However, while RNFLt had a significance level of p=0.016, ONS did not contribute significantly to the predictive model: p=0.372. However, it should be noted that RNFLt data from left eye was not normally distributed.

12.3.10 Optic Pathway Glioma Location

This data has been tabulated to aid exploratory analysis and reference, to help explain data within the results. The data has been compiled into Table 12.1.

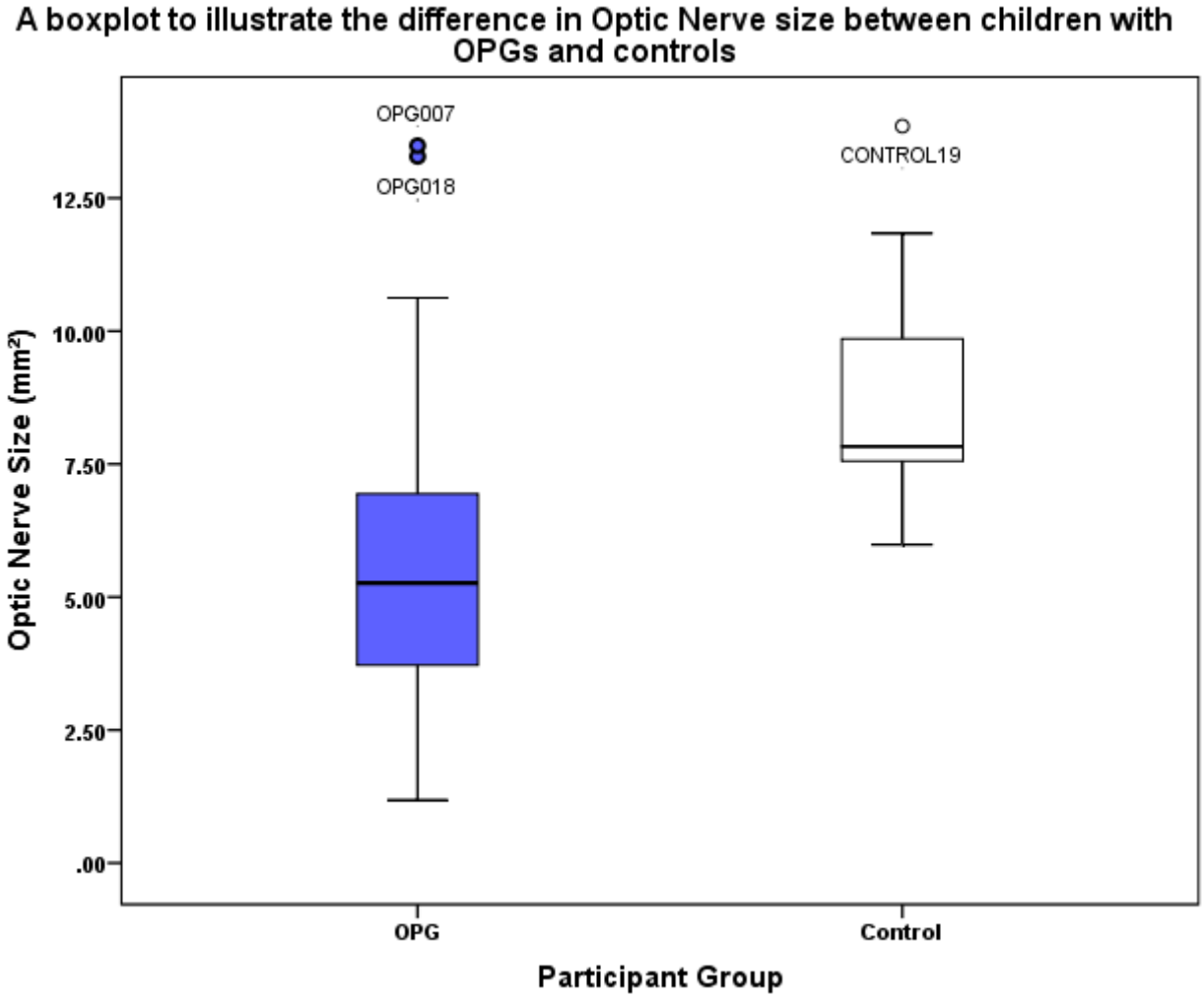

Figure 12.2 *A boxplot portraying the ONS distribution of children with OPGs and healthy controls*

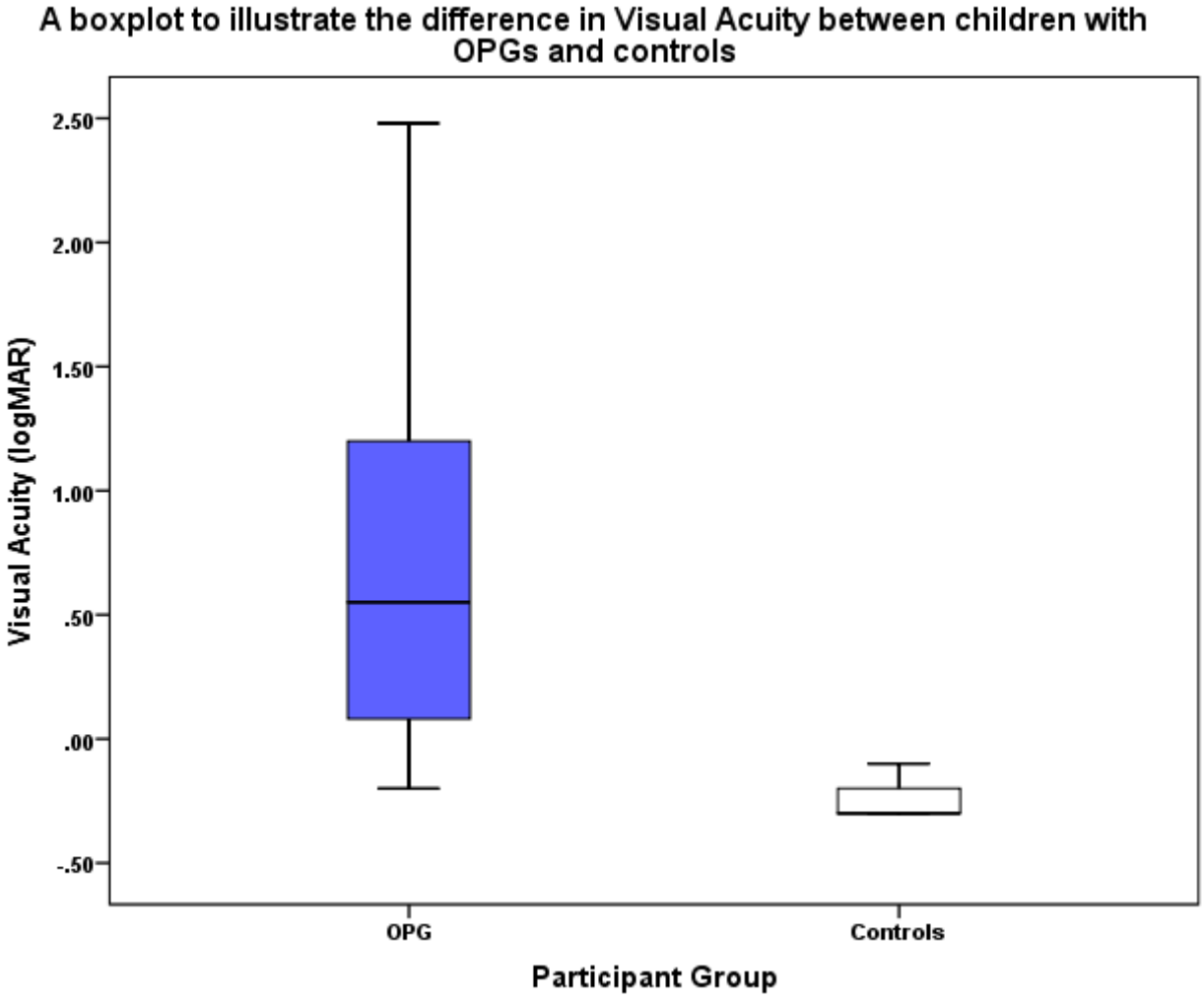

Figure 12.3 *A boxplot portraying the VA distribution of children with OPGs and healthy controls*

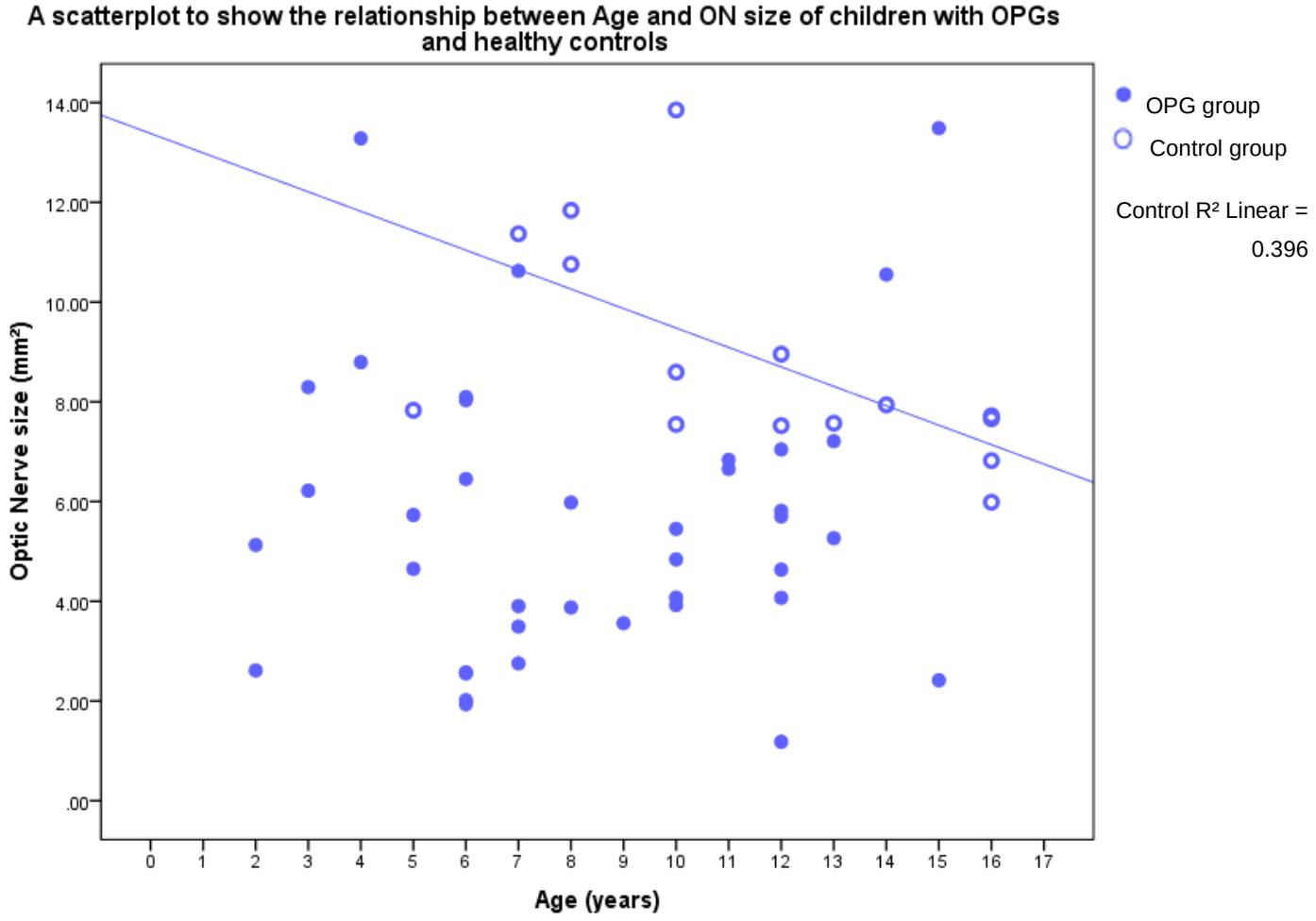

Figure 12.4 *A scatterplot portraying the relationship between ONS and Age of children with OPGs and healthy controls*

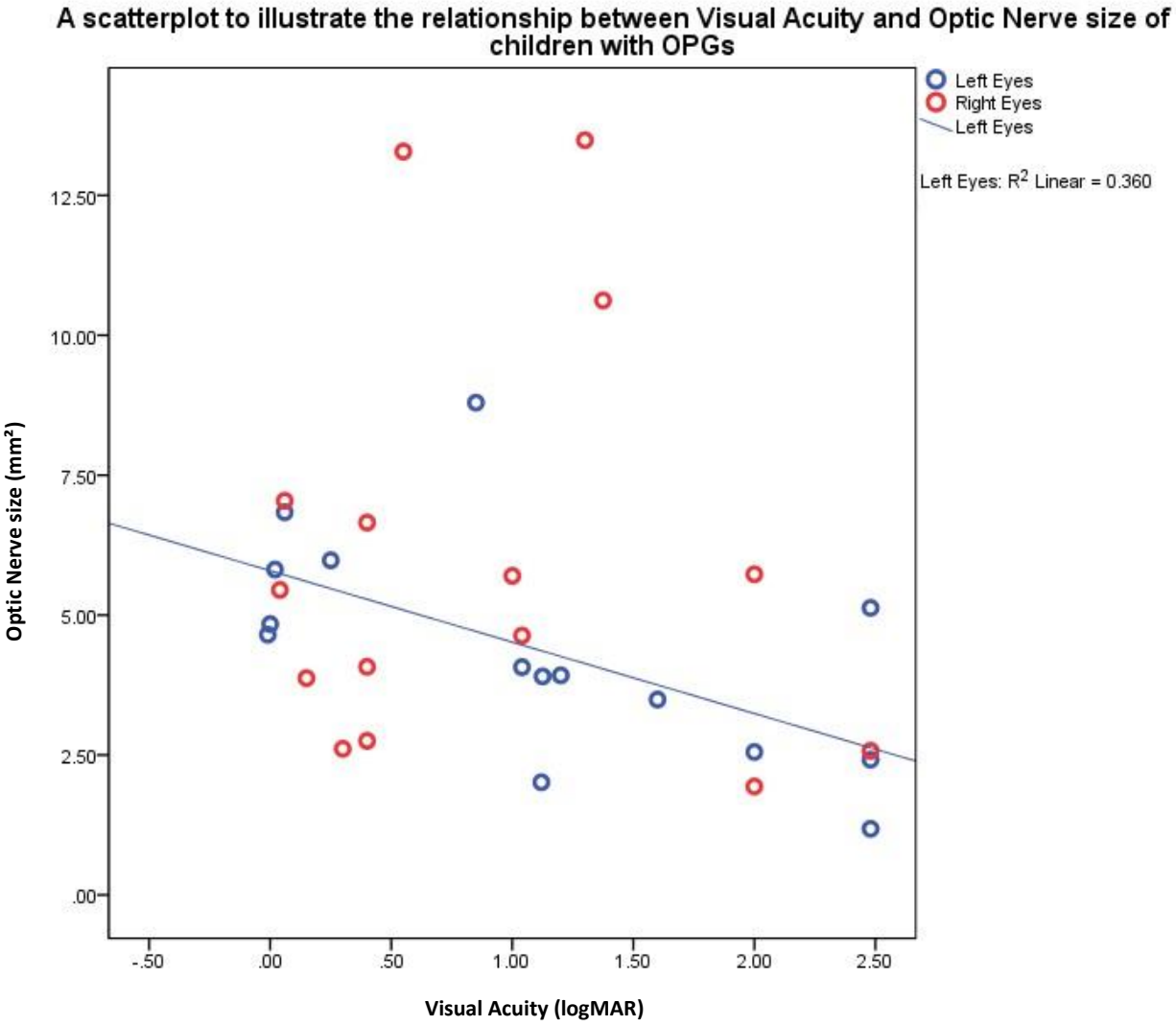

Figure 12.5 *A scatterplot portraying the relation between ONS and VA*

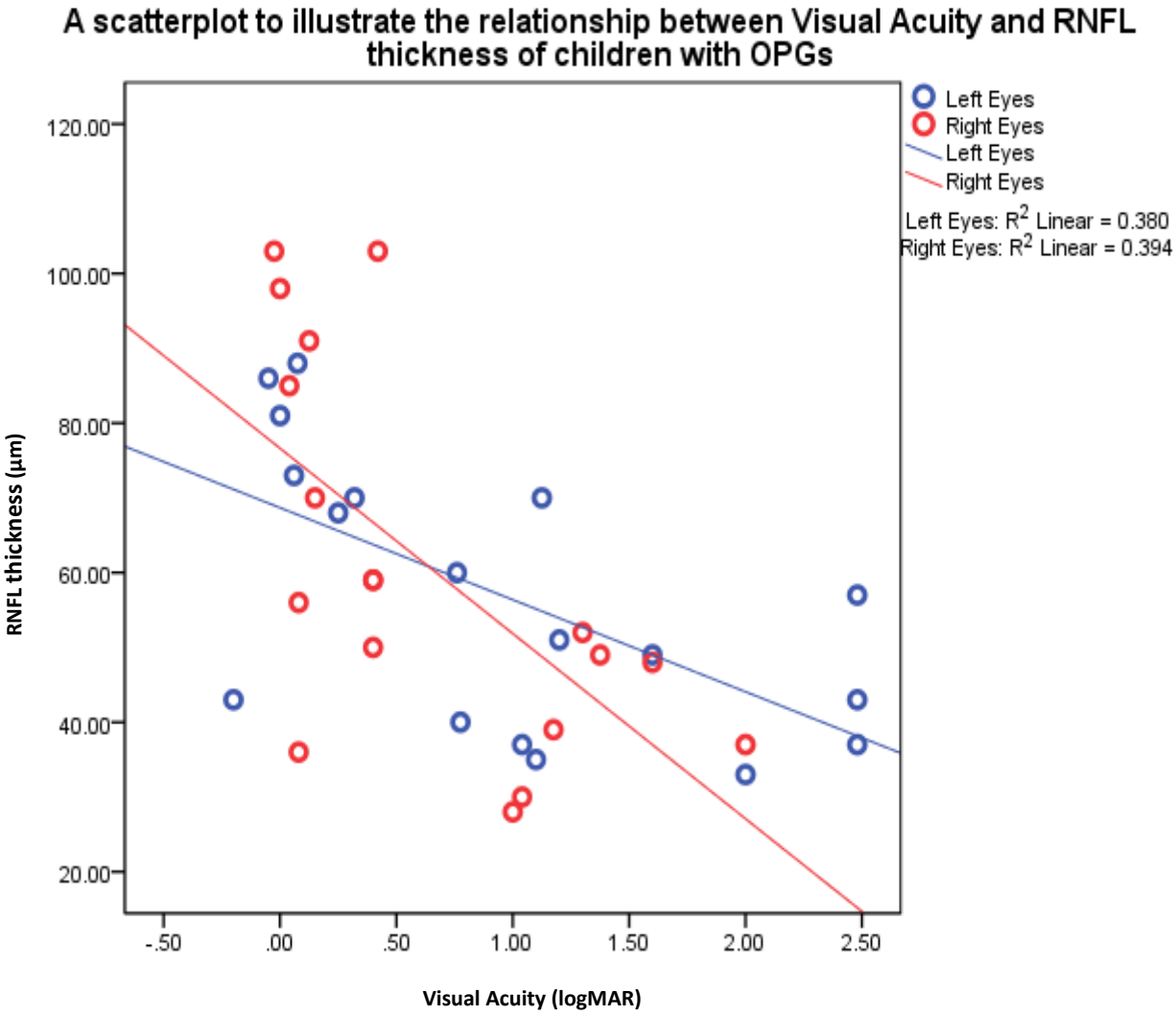

Figure 12.6 *A scatterplot portraying the significant correlation between RNFLt and VA*

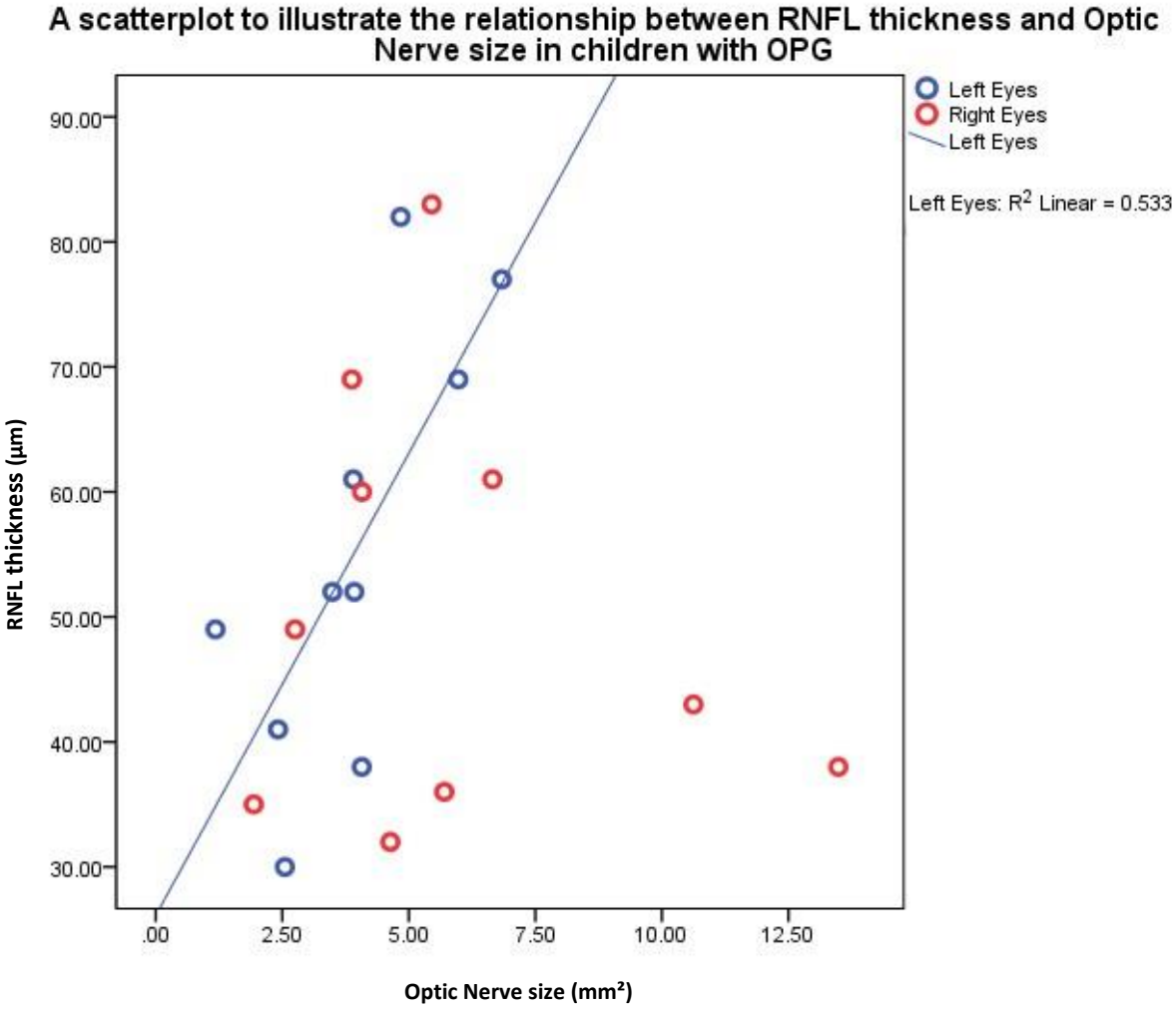

Figure 12.7 *A scatterplot portraying the relation between RNFLt and ONS*

Table 12.1 *A summary of the anatomical location of the OPG tumours according to MDC*

		Participants with OPG (OPGn)																																
Location, mDODGE		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**31**	**33**	**35**	**36**	**37**
left single ON	**1aL**	x														x										x	x	x	x	x	x	x		
right single ON	**1aR**										x		x																					
left>right bilateral ON	**1bL**							x																										
left<right bilateral ON	**1bR**											x			x							x		x										x
cisternal segment ON L	**1cL**	x								x											x										x	x		
cisternal segment ON R	**1cR**			x		x														x		x			x									
cisternal segment ON bilateral	**1cB**		x		x			x						x		x		x															x	
cisternal segment ON bilateral left>right	**1cbL**						x																						x					
cisternal segment ON bilateral left<right	**1cbR**								x								x		x					x										
central chiasmatic	**2a**		x								x		x		x		x	x	x				x	x	x	x			x			x	
asymmetric chiasmatic left>right	**2bL**	x								x																								
asymmetric chiasmatic left<right	**2bR**				x	x		x	x					x		x				x		x												
asymmetric chiasmatic left	**2cL**																																	
asymmetric chiasmatic right	**2cR**																																	
optic tracts left	**3L**																					x												
optic tracts right	**3R**				x									x												x								
optic tracts bilateral	**3B**							x		x							x						x	x	x				x				x	
asymmetric tracts left>right	**3bL**	x														x		x																
asymmetric tracts left<right	**3bR**		x						x											x														
diffuse posterior tracts left	**4L**																	x	x															
diffuse posterior tracts right	**4R**				x				x					x	x	x										x			x					
diffuse posterior tracts bilateral	**4B**							x																x									x	
asymmetric posterior tracts left>right	**4bL**																																	
asymmetric posterior tracts left<right	**4bR**																																	
hypothalamic involvment	**H+/-**		x	x		x		x	x		x		x		x	x	x	x		x		x	x			x			x				x	
leptomeningeal dissemination	**LM+/-**																																	
neurofibromatosis	**NF1+**		x	x	x	x		x			x	x			x	x		x	x			x		x	x	x	x	x	x	x	x	x	x	x

12.4 Discussion

12.4.1 Optic Pathway Glioma Data vs Control Data

There was a significant difference between the participants with OPGs and controls in both VA scores and ONS. As evident in the boxplots illustrated in Figures 12.2 and 12.3, data from the participants with OPGs exhibited a higher degree of variation illustrated by the wide interquartile range, while data from the control group was uniformly lower and more consistent, with a narrow interquartile range. The participants with OPGs consistently had smaller ONS and poorer VA scores compared to controls. Moreover, there was more variation in ONS than VA scores in the control group, which is consistent with the literature that healthy ONS are variable (113).

The results regarding poor VA scores in the OPG cohort are consistent with the literature (10,36). In addition, the locations of the OPGs illustrated in Table 12.1 indicates that tumour infiltration in the anterior VP appears to induce visual impairment (31). The worst recorded VA scores were 2.48logMAR (indicating no light perception) for OPG001, OPG007, OPG029, OPG021 (left eye) and OPG008 (right eye). Referring to Table 12.1, it appears that tumour infiltration before the chiasm on the left side of the VP associates with poorer VA scores of the left eye, and vice versa for OPG008. Hence, tumour location appears to be linked to resulting visual impairment in this dataset (15).

However, there were examples of OPG participants exhibiting similar results to the control group, such as OPG007 and OPG018, whose right ONS greatly varied from the rest of the data from the cohort (13.49mm² and 13.28mm² respectively), while their left ONs were very small in comparison (2.42 mm² and 8.8 mm²). When compared to the location of the tumours, they are not exclusively on the left which could have explained the relatively normal-sized right ON. Additionally, the VA scores from these participants do not indicate any reason as to why the right ONS appears normal. Further investigation is required, however, the differences in the sizes of their ON illustrates the tumour's impact on the anterior VP structure.

12.4.2 Optic Nerve Size and Age

The two variables did not significantly correlate in the OPG cohort, as they did in the healthy cohort in Section 11.3.4, in which smaller ONS were associated with older children. Additionally, when correlational analysis included data from both groups in the analysis, there was no correlation. The most likely explanation is that the glioma introduces profound changes in ONS independently of age. This is evident from the number of small ONS in the OPG cohort (e.g. majority of normal ONS >7mm

whereas most ONS in glioma are <7mm). Also there are a number of large ONS (>12mm) in glioma that may fall outside normal ranges. In both correlations, the CI were relatively narrow (-0.161 to 0.368 for both groups and -0.261 to 0.368 for the OPG group alone), despite the small sample size. This data appears to indicate that the ONS has been significantly changed by the presence of an OPG and has seemingly prevented it from developing normally, if the trend evident in the control data is representative of the general paediatric population. However, more investigation into children with OPGs and healthy paediatric cohort ONS is required to ascertain if this truly is the case.

12.4.3 Optic Nerve Size and Visual Acuity

For the comparison of ONS and VA, it was hypothesised that the two would be correlated in the OPG cohort. No significant correlation was found between the two variables and the CI accrued were relatively narrow (-0.576 to 0.097), indicating that results appears accurate, but not as accurate as the other correlations reported. Parsa (2012) would argue this is due to the lack of correlation between tumour size (hence, ONS) and VA. For example, OPGs formed from a uniform overgrowth of glial cells around the axons may not necessarily obstruct axoplasmic flow. Moreover, OPGs showing signs of regression can disrupt the structure of the ON and result in VA changes (149).

A significant, moderate correlation (R^2=0.360) was found in the left eyes [See Figure 12.5]. The scatterplot illustrates a negative trend in the data, implying that as the ONS decreased, the VA scores became worse. Although, as this is not evident across the entire data set, such as the right eyes of OPG007 and OPG018, a repeated study is required to investigate this correlation further, as it may be due to small sample size.

12.4.4 Retinal Nerve Fibre Layer Thickness and Visual Acuity

A significant, moderate correlation was found for both the left (R^2=0.380) and right (R^2=0.394) eye between VA and RNFLt. The CI accrued were relatively narrow (-0.776 to -0.338), indicating that the results appear accurate. Figure 12.6 illustrates variation in the data points, with an overall weak, negative trend in both plots, with data for the right eyes portraying a steeper, negative gradient. These results are consistent with the literature, in that thinning of the RNFL links to visual acuity impairments. For example, Avery *et al.* (2015) found that in children with OPGs experiencing vision loss, a loss in RNFLt was also experienced (108). Hence, despite the plots illustrating a weak trend, a larger sample size may indicate a stronger correlation in the OPG participant cohort to further reflect the literature.

The reason a negative correlation has been demonstrated between RNFLt and VA and not ONS and VA may be due to the fact that OCT outputs have a higher resolution compared to MRI, meaning OCT produces more image details, yielding more accurate data. An alternative factor may be that the RNFL measurement reflects axonal loss away from the OPG, compared to ONS which reflects axonal changes with regards to the tumour itself.

12.4.5 Retinal Nerve Fibre Layer and Optic Nerve Size

No significant correlation was found between ONS and RNFLt. However, the data for the left eyes alone indicated a strong positive correlation (R^2=0.533), inferring that as ONS increases, so does RNFLt. The CI are narrow (-0.175 to -0.649), despite being the largest evident in this overall analysis, indicating that this result was the least accurate. This may be attributed to the RNFLt data for the left eyes not being normally distributed. The scatterplot [Figure 12.7] illustrates that the right eye does not portray a strong positive trend and exhibits more variation than data for the left eye. The results for the left eyes are consistent with the literature. For example, optic neuritis and MS are associated with ON atrophy, and have been shown to correlate with RNFL thinning (150). As this is the smallest sample size from this section, it is unsurprising that the data does not show an overall strong trend, but a stronger trend may be evident if the study was repeated with a larger sample size of children with OPGs.

12.4.6 Does Retinal Nerve Fibre Layer and Optic Nerve Size predict Visual Acuity?

Exploratory analysis was conducted to investigate if RNFLt and ONS could be used to predict VA. Multiple regression was used, and found that for both eyes, RNFLt and ONS significantly predicted VA. Thus, inferring that these structural measurements of the anterior VP can be utilised in the prediction of visual function. However, ONS did not contribute significantly to the predictive model and as the data for the left RNFLt was not normally distributed which is an assumption of multiple linear regression, this assumption cannot be satisfactorily concluded. Hence, further study is required to identify if this model can be replicated in a larger sample size.

12.5 Summary

This study aimed to answer if ONS, VA and RNFLt correlated and if ONS and RNFLt could be used to predict VA. Significant correlations for the data as a whole was only exhibited between RNFLt and VA,

which is reflective of the literature (108). It has been speculated that the reason for a lack of correlation between ONS and VA may be due to the minimal correlation between tumour size and visual acuity, reported by Parsa (2012) (149) and that RNFLt is more reflective of axonal loss while ONS demonstrates tumour size and nearby axonal changes. However, the data for the left eyes did exhibit a significant, negative correlation between ONS and VA and a significant, positive correlation between ONS and RNFLt. Lack of strong correlations or strong trends in the data can be attributed to the small sample size used, and to difficulties revolving around testing vision in children and paediatric neuroimaging.

This study also investigated if ONS and Age correlated in the same way as seen in the healthy cohort. It was found this was not the case, and it appears that the presence of an OPG has disrupted the normal growth pattern of the ON. Moreover, a secondary aim of this study was to confirm that there was a significant difference between ONS and VA in healthy children and children with OPGs. The data concludes that there is a significant difference between ONS and VA in the two cohorts and that tumour location appears to play a role in resulting visual impairment. Although, further investigation to validate these findings is required.

13.0 Conclusion

The aim of this thesis was to establish if structural aspects of the anterior VP such as the size of the ON and RNFLt correlates to visual function, in an effort to distinguish a structural biomarker of visual impairment from MR images. The results of this thesis identify a reliable manual CSA technique of measuring ONS in both control populations and cohorts with irregular sized and shaped ON. A diameter-based technique was also studied, but the CSA method is preferable as it considers the entire shape of the ON in children with OPGs

This thesis has illustrated the technique's repeatability via the assessment of intra-observer variance and has outlined the measurement procedure as viable for clinical use. Inter-observer variance of the methods was also assessed, and the CSA method was deemed to have an acceptable level of reproducibility. This adds to the literature on using CSA techniques to measure the ONS of clinical groups, which has already been completed for optic neuritis (2,113–116) and ON hypoplasia (123,125). However, to the author's knowledge this is the first instance that this technique has been applied to measuring the ONS of children with OPGs.

The ONS of a healthy paediatric cohort was assessed using the CSA method and the mean ONS was reported as 8.798mm^2. To the author's knowledge, this is the first use of a manual CSA technique applied to this group. Therefore, this result adds to the literature of normative ONS variation (3,118,127,128,137), and reflects the existing findings in adult populations, reflecting ONS variation seen in adult cohorts (113–116). ONS of the healthy cohort was compared to ONS of the OPG group, and it was found that ONS was more variable and, as predicted, was significantly smaller in the children with OPGs. This infers that children with OPGs exhibit atrophic ON, linked to the presence of their benign tumour. As ONS has not been explored in an OPG cohort before, this data adds to the literature regarding structural components of the VP in this case group (104).

The relationships between ONS and height and ONS and age were explored and both were found to be negatively correlated, thus implying that as the healthy children matured, their ONS decreased. This result was unexpected but can be explained by axonal loss due to maturation. Although this has been demonstrated in cohorts aged 19years and above (148), this is the first it has been demonstrated in a paediatric cohort. Although this correlation was not evident in the OPG cohort, this can be explained by the warped ONS due to tumour location. As no prior investigation into normative variation of ONS in paediatric cohorts has been conducted, the results here cannot be validated against existing literature.

Correlational analysis could not be conducted on distance VA due to the lack of variation in the data. However, a consistently good level of VA was observed, and supported the literature on VA in healthy paediatric cohorts (4,10). Analysis was successfully conducted on NearVA and Age. However, no significant correlation was found between the two. The results indicated the overall good level of VA across the cohort. Moreover, the more variation evident in the NearVA data compared to the distance VA scores can be linked to the difficulty retaining a standardised reading-distance. This adds to the literature on using a SG test to assess VA in a paediatric cohort with wide-ranging ages (5-16years).

The OPG cohort was found to have worse VA than the control group, as predicted. However, it should be considered that VA was assessed using different tests between the groups. This adds to the literature on VA impairments associated with OPGs (5,10,34,69,80,108,151–155). Moreover, the location of the tumour appears to link to the location of visual deficits to an extent in this data, thus agreeing with the literature regarding the correlation between OPG location and associated manifestations (24).

It was hypothesised that ONS and RNFLt would predict VA. A significant prediction model was found between the factors although, ONS was not found to be significant within the model. This suggests that RNFLt can be used to predict VA, with ONS having less of an impact. It is inferred that this is due to OCT outputs having a higher resolution and RNFLt reflects axonal loss while ONS reflects tumour size and axonal loss. This is the first time this correlation has been attempted, and with a small sample size. Therefore, further study is required to confirm this lack of relationship.

A significant, negative correlation was exhibited between RNFLt and VA in the OPG group, which supports the literature regarding thinning of the RNFL linking to visual acuity impairments in children with OPG (108). This responds to Avery *et al's* (110) call for an understanding of the correlation between visual function decline and structural malformations in OPG cohorts and adds to the literature regarding the structure of the VP and the impact of OPGs (109).

However, the predicted significant correlations were not evidenced between ONS and RNFLt, and ONS and VA. The lack of correlation between ONS and VA may be due to the minimal correlation between tumour size and resulting visual function (149). While the lack of correlation between ONS and RNFLt may be due to OCT measures having a higher resolution than MRI. However, the data for the left eyes did exhibit a significant negative correlation between ONS and VA and a significant positive correlation between ONS and RNFLt. This either indicates that the data from the left eyes were not reflective of the OPG population and the presence of the outliers in ONS induced this correlation, or that the data of the right eye was not illustrative of the OPG population and resulted in the lack of correlation. However, due to the small sample size, this study would need to be

repeated to validate the result. No study regarding ONS has been conducted in the OPG cohort, to the author's knowledge. Therefore, this study has opened up the opportunity for further research into this structural component of the anterior VP and validated a method to complete it.

Lack of significant correlations in the data can be attributed to the small sample size used within these studies, and to difficulties revolving around testing vision in children and paediatric neuroimaging. Despite the use of a mock scanner, the success rate for the MRI scans was 50%. This contradicts De Bie et al's (62) report, that the use of a protocol incorporating a mock scanner would result in a more successful yield of useable scans. Although this factor was frustrating and resulted in smaller sample sizes than expected, lessons have been learned regarding the imaging of awake children, and this information can be considered in future paediatric imaging studies.

To move forward with this area of research, further investigations should be conducted to reveal the relationship between the wider structure of the anterior VP and resulting visual ability in OPG groups. Moreover, treatment application such as chemotherapy, radiotherapy and surgery could be assessed and their effect on anterior VP structure. Other factors of vision could be assessed such as VF and sweep VEP to see if a relationship exists between all factors of visual function and anterior VP structure. Furthermore, other structural elements of the VP could be investigated such as GCL-IPL or ON head volume. Lastly, more complex techniques such as DTT and DTI could be utilised to further identify ON structure.

14.0 References

1. Kuyk T, Elliott JL. Visual factors and mobility in persons with age-related macular degeneration. J Rehabil Res Dev [Internet]. 1999 Oct [cited 2017 May 8];36(4):303–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10678453

2. Trip SA, Schlottmann PG, Jones SJ, Li W-Y, Garway-Heath DF, Thompson AJ, et al. Optic nerve atrophy and retinal nerve fibre layer thinning following optic neuritis: Evidence that axonal loss is a substrate of MRI-detected atrophy. Neuroimage [Internet]. 2006 May [cited 2016 Nov 15];31(1):286–93. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1053811905024791

3. Yiannakas MC, Wheeler-Kingshott CAM, Berry AM, Chappell K, Henderson A, Kolappan M, et al. A method for measuring the cross sectional area of the anterior portion of the optic nerve in vivo using a fast 3D MRI sequence. J Magn Reson Imaging [Internet]. 2010 May 24 [cited 2017 Aug 29];31(6):1486–91. Available from: http://doi.wiley.com/10.1002/jmri.22202

4. Anstice NS, Thompson B. The measurement of visual acuity in children: An evidence-based update. Clin Exp Optom [Internet]. 2014 Jan 1 [cited 2017 Mar 10];97(1):3–11. Available from: http://doi.wiley.com/10.1111/cxo.12086

5. Parrozzani R, Clementi M, Kotsafti O, Miglionico G, Trevisson E, Orlando G, et al. Optical coherence tomography in the diagnosis of optic pathway gliomas. Invest Ophthalmol Vis Sci [Internet]. 2013 Dec 17 [cited 2017 Sep 14];54(13):8112–8. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.13-13093

6. Avery RA, Hardy KK. Vision specific quality of life in children with optic pathway gliomas. J Neurooncol [Internet]. 2014;116(2):341–7. Available from: http://link.springer.com/article/10.1007/s11060-013-1300-6

7. Mandelstam SA. Challenges of the anatomy and diffusion tensor tractography of the Meyer loop. AJNR Am J Neuroradiol [Internet]. 2012 Aug 1 [cited 2017 Nov 5];33(7):1204–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22422189

8. Garrity J, Whitney M, Betty M. The Optic Pathway - Eye Disorders - MSD Manual Professional Edition [Internet]. The Optic Pathway: Review. 2016 [cited 2017 Nov 5]. p. 1. Available from: http://www.msdmanuals.com/en-gb/professional/eye-disorders/optic-nerve-disorders/the-optic-pathway

9. Ridgway JP. Cardiovascular magnetic resonance physics for clinicians: part I. J Cardiovasc Magn Reson [Internet]. 2010 Nov 30 [cited 2017 Nov 5];12(1):71. Available from: http://jcmr-online.biomedcentral.com/articles/10.1186/1532-429X-12-71

10. Listernick R, Ferner RE, Liu GT, Gutmann DH. Optic pathway gliomas in neurofibromatosis-1: Controversies and recommendations. Ann Neurol [Internet]. 2007 Mar [cited 2016 Oct 18];61(3):189–98. Available from: http://doi.wiley.com/10.1002/ana.21107

11. Norgett and Siderov. Crowding in children's visual acuity tests--effect of test design and age. Optom Vis Sci [Internet]. 2011 Aug [cited 2017 Oct 12];88(8):920–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21532515

12. Jones D, Westall C, Averbeck K, Abdolell M. Visual acuity assessment: A comparison of two tests for measuring children's vision. Ophthalmic Physiol Opt [Internet]. 2003 Nov [cited 2017 Oct 12];23(6):541–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14622358

13. Ostrom QT, de Blank PM, Kruchko C, Petersen CM, Liao P, Finlay JL, et al. Alex's Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007-2011. Neuro Oncol [Internet]. 2015 Jan 1 [cited 2017 May 6];16(suppl 10):x1–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25542864

14. Avery RA, Liu GT, Fisher MJ, Quinn GE, Belasco JB, Phillips PC, et al. Retinal nerve fiber layer thickness in children with optic pathway gliomas. Am J Ophthalmol [Internet]. 2011 Mar [cited 2017 Mar 14];151(3):542–549.e2. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0002939410006896

15. Thomas RP, Gibbs IC, Xu LW, Recht L. Treatment Options for Optic Pathway Gliomas. Curr Treat Options Neurol [Internet]. 2015 Feb 27 [cited 2016 Jul 12];17(2):2. Available from: http://link.springer.com/10.1007/s11940-014-0333-2

16. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. Vol. 114, Acta Neuropathologica. 2007. p. 97–109.

17. Nicolin G, Parkin P, Mabbott D, Hargrave D, Bartels U, Tabori U, et al. Natural history and outcome of optic pathway gliomas in children. Pediatr Blood Cancer. 2009;53(7):1231–7.

18. Piccirilli M, Lenzi J, Delfinis C, Trasimeni G, Salvati M, Raco A. Spontaneous regression of optic pathways gliomas in three patients with neurofibromatosis type I and critical review of the literature. Vol. 22, Child's Nervous System. 2006. p. 1332–7.

19. Helfferich J, Nijmeijer R, Brouwer OF, Boon M, Fock A, Hoving EW, et al. Neurofibromatosis type 1 associated low grade gliomas: A comparison with sporadic low grade gliomas. Vol. 104, Critical Reviews in Oncology/Hematology. 2016. p. 30–41.

20. Chen YH, Gutmann DH. The molecular and cell biology of pediatric low-grade gliomas. Oncogene [Internet]. 2014;33(16):2019–26. Available from: http://www.nature.com/onc/journal/v33/n16/pdf/onc2013148a.pdf

21. Bar EE, Lin A, Tihan T, Burger PC, Eberhart CG. Frequent gains at chromosome 7q34 involving BRAF in pilocytic astrocytoma. J Neuropathol Exp Neurol. 2008;67(9):878–87.

22. Jones DTW, Hutter B, Jäger N, Korshunov A, Kool M, Warnatz H-J, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat Genet [Internet]. 2013 Jun 30 [cited 2017 May 6];45(8):927–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23817572

23. Yu J, Deshmukh H, Gutmann RJ, Emnett RJ, Rodriguez FJ, Watson MA, et al. Alterations of BRAF and HIPK2 loci predominate in sporadic pilocytic astrocytoma. Neurology. 2009;73(19):1526–31.

24. Belirgen M, Berrak SG, Ozdag H, Bozkurt SU, Eksioglu-Demiralp E, Ozek MM. Biologic tumor behavior in pilocytic astrocytomas. Child's Nerv Syst. 2012;28(3):375–89.

25. Wimmer K, Eckart M, Meyer-Puttlitz B, Fonatsch C, Pietsch T. Mutational and expression analysis of the NF1 gene argues against a role as tumor suppressor in sporadic pilocytic astrocytomas. J Neuropathol Exp Neurol [Internet]. 2002 Oct [cited 2017 May 6];61(10):896–902. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12387455

26. Dasgupta B, Dugan LL, Gutmann DH. The neurofibromatosis 1 gene product neurofibromin regulates pituitary adenylate cyclase-activating polypeptide-mediated signaling in astrocytes. J Neurosci [Internet]. 2003;23(26):8949–54. Available from: http://dx.doi.org/%5Cnhttp://www.jneurosci.org/content/23/26/8949.full.pdf

27. Bajenaru ML, Zhu Y, Hedrick NM, Donahoe J, Parada LF, Gutmann DH. Astrocyte-specific inactivation of the neurofibromatosis 1 gene (NF1) is insufficient for astrocytoma formation. Mol Cell Biol [Internet]. 2002;22(14):5100–13. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12077339%5Cnhttp://mcb.asm.org/content/22/14/5100.full.pdf

28. Weiss L, Sagerman RH, King GA, Chung CT, Dubowy RL. Controversy in the management of optic nerve glioma. Cancer [Internet]. 1987 Mar 1 [cited 2017 May 6];59(5):1000–4. Available from: http://doi.wiley.com/10.1002/1097-0142%252819870301%252959%253A5%253C1000%253A%253AAID-CNCR2820590525%253E3.0.CO%253B2-N

29. Binning MJ, Liu JK, Kestle JRW, Brockmeyer DL, Walker ML. Optic pathway gliomas: a review. Neurosurg Focus [Internet]. 2007 Nov [cited 2016 Oct 18];23(5):E2. Available from: http://thejns.org/doi/abs/10.3171/FOC-07/11/E2

30. Steinbok P, Hentschel S, Almqvist P, Cochrane DD, Poskitt K. Management of optic chiasmatic/hypothalamic astrocytomas in children. Can J Neurol Sci [Internet]. 2002 [cited 2017 May 6];29(2):132–8. Available from: https://www.ncbi.nlm.nih.gov/labs/articles/12035834/

31. El Beltagy MA, Reda M, Enayet A, Zaghloul MS, Awad M, Zekri W, et al. Treatment and Outcome in 65 Children with Optic Pathway Gliomas. World Neurosurg. 2016;89:525–34.

32. Walrath JD, Engelbert M, Kazim M. Magnetic Resonance Imaging Evidence of Optic Nerve Glioma Progression Into and Beyond the Optic Chiasm. Ophthalmic Plast Reconstr Surg [Internet]. 2008 Nov [cited 2017 May 6];24(6):473–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19033845

33. Shamji MF, Benoit BG. Syndromic and sporadic pediatric optic pathway gliomas: review of clinical and histopathological differences and treatment implications. Neurosurg Focus [Internet]. 2007 Nov [cited 2017 Apr 11];23(5):E3. Available from: http://thejns.org/doi/10.3171/FOC-07/11/E3

34. Avery RA, Fisher MJ, Liu GT. Optic pathway gliomas. J Neuroophthalmol [Internet]. 2011 [cited 2016 Oct 18];31:269–78. Available from: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/jamaophthalmol.2013.1652

35. Packer R, Bilaniuk L, Cohen B. Intracranial Visual Pathway Gliomas in Children With Neurofibromatosis - Journals - NCBI. Neurofibromatosis [Internet]. 1988 [cited 2017 May 6];1(4):212–22. Available from: https://www.ncbi.nlm.nih.gov/labs/articles/3152473/

36. Fisher MJ, Loguidice M, Gutmann DH, Listernick R, Ferner RE, Ullrich NJ, et al. Visual outcomes in children with neurofibromatosis type 1-associated optic pathway glioma following chemotherapy: A multicenter retrospective analysis. Neuro Oncol [Internet]. 2012 Jun 1 [cited 2017 Jun 26];14(6):790–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22474213

37. Balcer LJ, Liu GT, Heller G, Bilaniuk L, Volpe NJ, Galetta SL, et al. Visual loss in children with neurofibromatosis type 1 and optic pathway gliomas. Am J Ophthalmol [Internet]. 2001 Apr [cited 2017 Mar 14];131(4):442–5. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0002939400008527

38. Diggs-Andrews KA, Brown JA, Gianino SM, Rubin JB, Wozniak DF, Gutmann DH. Sex Is a major determinant of neuronal dysfunction in neurofibromatosis type 1. Ann Neurol [Internet]. 2014 Feb [cited 2017 May 8];75(2):309–16. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/24375753

39. Thiagalingam S, Flaherty M, Billson F, North K. Neurofibromatosis type 1 and optic pathway gliomas: Follow-up of 54 patients. Ophthalmology [Internet]. 2004 Mar [cited 2017 Mar 14];111(3):568–77. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0161642003014209

40. Campagna M, Opocher E, Viscardi E, Calderone M, Severino SM, Cermakova I, et al. Optic pathway glioma: Long-term visual outcome in children without neurofibromatosis type-1. Pediatr Blood Cancer [Internet]. 2010 Dec 1 [cited 2017 May 8];55(6):1083–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20979170

41. Chateil J-F, Soussotte C, Pédespan J-M, Brun M, Le Manh C, Diard F. MRI and clinical differences between optic pathway tumours in children with and without neurofibromatosis. Br J Radiol [Internet]. 2001 Jan [cited 2017 May 8];74(877):24–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11227773

42. Shapey J, Danesh-Meyer HV, Kaye AH. Diagnosis and management of optic nerve glioma. J Clin Neurosci. 2011;18(12):1585–91.

43. Lavery MA, O'Neill JF, Chu FC, Martyn LJ. Acquired nystagmus in early childhood: a presenting sign of intracranial tumor. Ophthalmology [Internet]. 1984 May [cited 2017 May 8];91(5):425–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6739045

44. Robinson KE, Kuttesch JF, Champion JE, Andreotti CF, Hipp DW, Bettis A, et al. A quantitative meta-analysis of neurocognitive sequelae in survivors of pediatric brain tumors. Pediatr Blood Cancer [Internet]. 2010 May 24 [cited 2017 May 8];55(3):525–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20658625

45. Listernick R, Darling C, Greenwald M, Strauss L, Charrow J. Optic pathway tumors in children: The effect of neurofibromatosis type 1 on clinical manifestations and natural history. J Pediatr [Internet]. 1995 Nov [cited 2017 Mar 14];127(5):718–22. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0022347695701591

46. Winter MD. The Basics of Musculoskeletal Magnetic Resonance Imaging. Vet Clin North Am Equine Pract [Internet]. 2012 Dec [cited 2017 May 8];28(3):599–616. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23177134

47. Bydder GM, Young IR. MR imaging: clinical use of the inversion recovery sequence. [Internet]. Vol. 9, Journal of computer assisted tomography. 1985 [cited 2017 May 8]. p. 659–75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2991345

48. Wan MJ. Optic Pathway Gliomas in Children [Internet]. American Academy of Ophthalmology. 2015. Available from: http://www.aao.org/pediatric-center-detail/neuro-ophthalmology-optic-pathway-gliomas-in-child

49. Dhermain FG, Hau P, Lanfermann H, Jacobs AH, van den Bent MJ. Advanced MRI and PET imaging for assessment of treatment response in patients with gliomas. Lancet Neurol [Internet]. 2010 Sep [cited 2017 Jun 3];9(9):906–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20705518

50. Yamamoto A, Miki Y, Urayama S, Fushimi Y, Okada T, Hanakawa T, et al. Diffusion tensor fiber tractography of the optic radiation: analysis with 6-, 12-, 40-, and 81-directional motion-probing gradients, a preliminary study. AJNR Am J Neuroradiol [Internet]. 2007 Jan [cited 2017 Jun 5];28(1):92–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17213432

51. Nickerson JP, Salmela MB, Koski CJ, Andrews T, Filippi CG. Diffusion tensor imaging of the

pediatric optic nerve: Intrinsic and extrinsic pathology compared to normal controls. J Magn Reson Imaging [Internet]. 2010 Jun 23 [cited 2017 Jun 5];32(1):76–81. Available from: http://doi.wiley.com/10.1002/jmri.22228

52. Astrup J. Natural history and clinical management of optic pathway glioma. Br J Neurosurg. 2003;17(4):327–35.

53. Lober RM, Guzman R, Cheshier SH, Fredrick DR, Edwards MSB, Yeom KW. Application of diffusion tensor tractography in pediatric optic pathway glioma. J Neurosurg Pediatr [Internet]. 2012;10(4):273–80. Available from: http://thejns.org/doi/full/10.3171/2012.7.PEDS1270

54. Weizman L, Ben Sira L, Joskowicz L, Constantini S, Precel R, Shofty B, et al. Automatic segmentation, internal classification, and follow-up of optic pathway gliomas in MRI. Med Image Anal. 2012;16(1):177–88.

55. Lee C-H, Schmidt M, Murtha A. Segmenting brain tumors with conditional random fields and support vector machines. Comput Vis … [Internet]. 2005 [cited 2017 May 20];469–78. Available from: https://pdfs.semanticscholar.org/740a/8d65203a6d2a029bf57ad32b02ddb0158434.pdf

56. Kaus MR, Warfield SK, Nabavi A, Black PM, Jolesz FA, Kikinis R. Automated Segmentation of MR Images of Brain Tumors. Radiology [Internet]. 2001;218(2):586–91. Available from: http://pubs.rsna.org/doi/10.1148/radiology.218.2.r01fe44586

57. Dodge J, H.W., Love JG, Craig WM, Dockerty MB, Kearns TP, et al. Gliomas of the optic nerves. AMA Arch Neurol Psychiatry [Internet]. 1958 Jun 1 [cited 2017 May 23];79(6):607–21. Available from: http://archneurpsyc.jamanetwork.com/article.aspx?doi=10.1001/archneurpsyc.1958.02340060003001

58. Taylor T, Jaspan T, Milano G, Gregson R, Parker T, Ritzmann T, et al. Radiological classification of optic pathway gliomas: Experience of a modified functional classification system. Br J Radiol [Internet]. 2008 Oct [cited 2017 May 29];81(970):761–6. Available from: http://www.birpublications.org/doi/10.1259/bjr/65246351

59. Walker DA, Liu J, Kieran M, Jabado N, Picton S, Packer R, et al. A multi-disciplinary consensus statement concerning surgical approaches to low-grade, high-grade astrocytomas and diffuse intrinsic pontine gliomas in childhood (CPN Paris 2011) using the Delphi method. Neuro Oncol [Internet]. 2013 Apr [cited 2017 May 29];15(4):462–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23502427

60. Edwards AD, Arthurs OJ. Paediatric MRI under sedation: Is it necessary? What is the evidence for the alternatives? [Internet]. Vol. 41, Pediatric Radiology. 2011. p. 1353–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21678113

61. Shields CH, Johnson S, Knoll J, Chess C, Goldberg D, Creamer K. Sleep Deprivation for Pediatric Sedated Procedures: Not Worth the Effort. Pediatrics [Internet]. 2004 [cited 2017 May 29];113(5). Available from: http://pediatrics.aappublications.org/content/113/5/1204

62. de Bie HMA, Boersma M, Wattjes MP, Adriaanse S, Vermeulen RJ, Oostrom KJ, et al. Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans. Eur J Pediatr [Internet]. 2010 Sep [cited 2017 Jun 5];169(9):1079–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20225122

63. Fisher M, Balcer L, Gutmann D, Listernick R, Ferner R, Packer R, et al. NEUROFIBROMATOSIS

TYPE 1 ASSOCIATED OPTIC GLIOMA VISUAL OUTCOMES FOLLOWING CHEMOTHERAPY: AN INTERNATIONAL MULTI-CENTER RETROSPECTIVE ANALYSIS. NEURO-ONCOLOGY (pp II19 - II20) OXFORD UNIV Press INC [Internet]. 2010 [cited 2017 Jun 26]; Available from: http://discovery.ucl.ac.uk/1326767/

64. Listernick R, Louis DN, Packer RJ, Gutmann DH. Optic pathway gliomas in children with neurofibromatosis 1: Consensus statement from the nf1 optic pathway glioma task force. Ann Neurol [Internet]. 1997 Feb [cited 2016 Oct 18];41(2):143–9. Available from: http://doi.wiley.com/10.1002/ana.410410204

65. Ferner RE, Huson SM, Thomas N, Moss C, Willshaw H, Evans DG, et al. Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. J Med Genet [Internet]. 2007 [cited 2017 May 8];44(2):81–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17105749

66. Liu GT. Optic gliomas of the anterior visual pathway. Curr Opin Ophthalmol. 2006;17:427–31.

67. Mahoney DH, Cohen ME, Friedman HS, Kepner JL, Gemer L, Langston JW, et al. Carboplatin is effective therapy for young children with progressive optic pathway tumors: a Pediatric Oncology Group phase II study. Neuro Oncol [Internet]. 2000;2(4):213–20. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1920597&tool=pmcentrez&rendertype=abstract

68. Janss AJ, Grundy R, Cnaan A, Savino PJ, Packer RJ, Zackai EH, et al. Optic pathway and hypothalamic/chiasmatic gliomas in children younger than age 5 years with a 6-year follow-up. Cancer [Internet]. 1995 Feb 15 [cited 2017 Jul 16];75(4):1051–9. Available from: http://doi.wiley.com/10.1002/1097-0142%2819950215%2975%3A4%3C1051%3A%3AAID-CNCR2820750423%3E3.0.CO%3B2-S

69. Kelly JP, Leary S, Khanna P, Weiss AH. Longitudinal measures of visual function, tumor volume, and prediction of visual outcomes after treatment of optic pathway gliomas. Ophthalmology [Internet]. 2012;119(6):1231–7. Available from: c:%5CUsers%5Clpzccb%5CAppData%5CRoaming%5CMozilla%5CFirefox%5CProfiles%5Crbi2rvfy.default%5Czotero%5Cstorage%5C3EM3E6HP%5CS0161642011012383.html

70. Moreno L, Bautista F, Ashley S, Duncan C, Zacharoulis S. Does chemotherapy affect the visual outcome in children with optic pathway glioma? A systematic review of the evidence. Eur J Cancer [Internet]. 2010 Aug [cited 2017 Mar 14];46(12):2253–9. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0959804910002595

71. Horwich A, Bloom HJGJ. Optic gliomas: Radiation therapy and prognosis. Int J Radiat Oncol [Internet]. 1985 Jun [cited 2017 Mar 14];11(6):1067–79. Available from: http://linkinghub.elsevier.com/retrieve/pii/0360301685900525

72. Combs SE, Schulz-Ertner D, Moschos D, Thilmann C, Huber PE, Debus J. Fractionated stereotactic radiotherapy of optic pathway gliomas: Tolerance and long-term outcome. Int J Radiat Oncol Biol Phys. 2005;62(3):814–9.

73. Wong JYC, Uhl V, Wara WM, Sheline GE. Optic gliomas: A reanalysis of the university of California, San Francisco experience. Cancer. 1987;60(8):1847–55.

74. Grabenbauer GG, Schuchardt U, Buchfelder M, Rödel CM, Gusek G, Marx M, et al. Radiation therapy of optico-hypothalamic gliomas (OHG) - Radiographic response, vision and late toxicity. Radiother Oncol. 2000;54(3):239–45.

75. Grill J, Couanet D, Cappelli C, Habrand JL, Rodriguez D, Sainte-Rose C, et al. Radiation-induced

cerebral vasculopathy in children with neurofibromatosis and optic pathway glioma. Ann Neurol. 1999;45(3):393–6.

76. Lacaze E, Kieffer V, Streri A, Lorenzi C, Gentaz E, Habrand J-L, et al. Neuropsychological outcome in children with optic pathway tumours when first-line treatment is chemotherapy. Br J Cancer [Internet]. 2003 Dec 1 [cited 2017 Jul 16];89(11):2038–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14647135

77. Blanchard G, Lafforgue M-P, Lion-François L, Kemlin I, Rodriguez D, Castelnau P, et al. Systematic MRI in NF1 children under six years of age for the diagnosis of optic pathway gliomas. Study and outcome of a French cohort. Eur J Paediatr Neurol. 2016;20(2):275–81.

78. Sigorini M, Zuccoli G, Ferrozzi F, Bacchini E, Street ME, Piazza P, et al. Magnetic resonance findings and ophthalmologic abnormalities are correlated in patients with neurofibromatosis type 1 (NF1). Am J Med Genet. 2000;93(4):269–72.

79. Chen SI, Chandna A, Norcia AM, Pettet M, Stone D. The repeatability of best corrected acuity in normal and amblyopic children 4 to 12 years of age. Investig Ophthalmol Vis Sci [Internet]. 2006 Feb 1 [cited 2017 Oct 12];47(2):614–9. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.05-0610

80. Avery RA, Ferner RE, Listernick R, Fisher MJ, Gutmann DH, Liu GT. Visual acuity in children with low grade gliomas of the visual pathway: Implications for patient care and clinical research [Internet]. Vol. 110, Journal of Neuro-Oncology. Springer US; 2012 [cited 2016 Oct 17]. p. 1–7. Available from: http://link.springer.com/10.1007/s11060-012-0944-y

81. Volpe N, Liu G, Galetta S. Neuro-ophthalmology: Diagnosis and Management, Second Edition. [Internet]. W.B. Saunders Company; 2010 [cited 2017 Jul 18]. Available from: https://www.scholars.northwestern.edu/en/publications/neuro-ophthalmology-diagnosis-and-management-second-edition-2

82. Beck RW, Maguire MG, Bressler NM, Glassman AR, Lindblad AS, Ferris FL. Visual Acuity as an Outcome Measure in Clinical Trials of Retinal Diseases. Ophthalmology. 2007;114(10).

83. Virgili G, Acosta R, Grover LL, Bentley SA, Giacomelli G. Reading aids for adults with low vision. In: Virgili G, editor. Cochrane Database of Systematic Reviews [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2013 [cited 2017 Oct 12]. p. CD003303. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24154864

84. Dobson V, Quinn GE, Biglan AW, Tung B, Flynn JT, Palmer EA. Acuity card assessment of visual function in the cryotherapy for retinopathy of prematurity trial. Investig Ophthalmol Vis Sci. 1990;31(9):1702–8.

85. Kushner BJ, Lucchese NJ, Morton G V. Grating Visual Acuity With Teller Cards Compared With Snellen Visual Acuity in Literate Patients. Arch Ophthalmol [Internet]. 1995 Apr 1 [cited 2017 Jul 18];113(4):485. Available from: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/archopht.1995.01100040107035

86. Dobson V, Quinn GE, Tung B, Palmer EA, Reynolds JD. Comparison of recognition and grating acuities in very-low-birth-weight children with and without retinal residua of retinopathy of prematurity. Investig Ophthalmol Vis Sci. 1995;36(3):692–702.

87. Dobson V, Quinn GE, Siatkowski RM, Baker JD, Hardy RJ, Reynolds JD, et al. Agreement between grating acuity at age 1 year and Snellen acuity at age 5.5 years in the preterm child. Investig Ophthalmol Vis Sci. 1999;40(2):496–503.

88. Mayer DL, Beiser AS, Warner AF, Pratt EM, Raye KN, Lang JM. Monocular acuity norms for the Teller Acuity Cards between ages one month and four years. Invest Ophthalmol Vis Sci [Internet]. 1995 Mar [cited 2017 Oct 12];36(3):671–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7890497

89. Hargadon DD, Wood J, Twelker JD, Harvey EM, Dobson V. Recognition Acuity, Grating Acuity, Contrast Sensitivity, and Visual Fields in 6-Year-Old Children. Arch Ophthalmol [Internet]. 2010 Jan 1 [cited 2017 Oct 12];128(1):70. Available from: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/archophthalmol.2009.343

90. Becker R, Hübsch S, Gräf MH, Kaufmann H. Examination of young children with Lea symbols. Br J Ophthalmol [Internet]. 2002 May 1 [cited 2017 Oct 12];86(5):513–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11973243

91. Simmers AJ, Gray LS, Spowart K. Screening for amblyopia: a comparison of paediatric letter tests. Br J Ophthalmol [Internet]. 1997 Jun 1 [cited 2017 Oct 12];81(6):465–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9274410

92. Drover JR, Felius J, Cheng CS, Morale SE, Wyatt L, Birch EE. Normative pediatric visual acuity using single surrounded HOTV optotypes on the Electronic Visual Acuity Tester following the Amblyopia Treatment Study protocol. J Am Assoc Pediatr Ophthalmol Strabismus [Internet]. 2008 Apr 1 [cited 2017 Oct 12];12(2):145–9. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1091853107004946

93. Pan Y, Tarczy-Hornoch K, Cotter SA, Wen G, Borchert MS, Azen SP, et al. Visual acuity norms in pre-school children: the Multi-Ethnic Pediatric Eye Disease Study. Optom Vis Sci [Internet]. 2009 Jun [cited 2017 Mar 27];86(6):607–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19430325

94. Birch EE, Strauber SF, Beck RW, Holmes JM. Comparison of the Amblyopia Treatment Study HOTV and the Electronic-Early Treatment of Diabetic Retinopathy Study visual acuity protocols in amblyopic children aged 5 to 11 years. J Am Assoc Pediatr Ophthalmol Strabismus [Internet]. 2009 Feb 1 [cited 2017 Oct 12];13(1):75–8. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1091853108002814

95. Chang BCM, Mirabella G, Yagev R, Banh M, Mezer E, Parkin PC, et al. Screening and Diagnosis of Optic Pathway Gliomas in Children with Neurofibromatosis Type 1 by Using Sweep Visual Evoked Potentials. Investig Opthalmology Vis Sci [Internet]. 2007 Jun 1 [cited 2016 Oct 18];48(6):2895. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.06-0429

96. Angeles-Han ST, Griffin KW, Harrison MJ, Lehman TJA, Leong T, Robb RR, et al. Development of a vision-related quality of life instrument for children ages 8-18 years for use in juvenile idiopathic arthritis-associated uveitis. Arthritis Care Res (Hoboken) [Internet]. 2011 Sep [cited 2017 May 8];63(9):1254–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21678564

97. Okamoto F, Okamoto Y, Fukuda S, Hiraoka T, Oshika T, L M. Vision-Related Quality of Life and Visual Function after Vitrectomy for Various Vitreoretinal Disorders. Investig Opthalmology Vis Sci [Internet]. 2010 Feb 1 [cited 2017 May 8];51(2):744. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.09-3992

98. Cubbidge RP. Visual fields [Internet]. Elsevier Butterworth-Heinemann; 2005 [cited 2017 Jul 23]. 119 p. Available from: https://books.google.com.mt/books?id=q2yzHg1oph4C&printsec=frontcover&dq=visual+field&hl=en&sa=X&ved=0ahUKEwjWt5DugJ_VAhVFiRoKHcchATkQ6AEIIjAA#v=onepage&q=visua

l field&f=true

99. Karcioglu ZA. Orbital tumors : diagnosis and treatment, 2nd edition [Internet]. 2014 [cited 2017 Jul 23]. pages cm. Available from: https://books.google.co.uk/books?id=48TlBAAAQBAJ&pg=PA58&dq=goldmann+test&hl=en&sa=X&ved=0ahUKEwjxv7Soup_VAhWLfRoKHfGwDj4Q6AEIKDAA#v=onepage&q=goldmann test&f=false

100. Getz LM, Dobson V, Luna B, Mash C. Interobserver reliability of the Teller Acuity Card procedure in pediatric patients. Investig Ophthalmol Vis Sci [Internet]. 1996;37(1):180–7. Available from: file:///C:/Users/bethh/Downloads/180.pdf

101. Moke PS, Turpin AH, Beck RW, Holmes JM, Repka MX, Birch EE, et al. Computerized method of visual acuity testing: Adaptation of the Amblyopia Treatment Study visual acuity testing protocol. Am J Ophthalmol. 2001;132(6):903–9.

102. Ferner RE. Neurofibromatosis 1 and neurofibromatosis 2: a twenty first century perspective. Vol. 6, Lancet Neurology. 2007. p. 340–51.

103. Gu S, Glaug N, Cnaan A, Packer RJ, Avery RA. Ganglion Cell Layer–Inner Plexiform Layer Thickness and Vision Loss in Young Children With Optic Pathway Gliomas. Investig Opthalmology Vis Sci [Internet]. 2014 Mar 10 [cited 2017 Jun 14];55(3):1402. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3954001/

104. Fisher JB, Jacobs DA, Markowitz CE, Galetta SL, Volpe NJ, Nano-Schiavi ML, et al. Relation of visual function to retinal nerve fiber layer thickness in multiple sclerosis. Ophthalmology. 2006;113(2):324–32.

105. Danesh-Meyer H V., Carroll SC, Foroozan R, Savino PJ, Fan J, Jiang Y, et al. Relationship between retinal nerve fiber layer and visual field sensitivity as measured by optical coherence tomography in chiasmal compression. Investig Ophthalmol Vis Sci. 2006;47(11):4827–35.

106. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. Vol. 26, Progress in Retinal and Eye Research. 2007. p. 688–710.

107. Phillips PH. Is optical coherence tomography indicated for the evaluation of patients with neurofibromatosis type 1? J Am Assoc Pediatr Ophthalmol Strabismus [Internet]. 2010 Dec [cited 2017 Jul 18];14(6):467–8. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1091853110004799

108. Avery RA, Cnaan A, Schuman JS, Trimboli-Heidler C, Chen CL, Packer RJ, et al. Longitudinal change of circumpapillary retinal nerve fiber layer thickness in children with optic pathway gliomas. Am J Ophthalmol. 2015;160(5):944–952.e1.

109. Pinto LM, Costa EF, Melo LAS, Gross PB, Sato ET, Almeida AP, et al. Structure-function correlations in glaucoma using matrix and standard automated perimetry versus time-domain and spectral-domain OCT devices. Investig Ophthalmol Vis Sci. 2014;55(5):3074–80.

110. Levin MH, Armstrong GT, Broad JH, Zimmerman R, Bilaniuk LT, Feygin T, et al. Risk of optic pathway glioma in children with neurofibromatosis type 1 and optic nerve tortuosity or nerve sheath thickening. Br J Ophthalmol [Internet]. 2015 Apr [cited 2016 Sep 28];100(4):515–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26294105

111. Ji J, Shimony J, Gao F, McKinstry RC, Gutmann DH. Optic nerve tortuosity in children with neurofibromatosis type 1. Pediatr Radiol [Internet]. 2013 Oct 2 [cited 2017 Oct 13];43(10):1336–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23636538

112. Avery RA, Mansoor A, Idrees R, Trimboli-Heidler C, Ishikawa H, Packer RJ, et al. Optic pathway glioma volume predicts retinal axon degeneration in neurofibromatosis type 1: Table. Neurology [Internet]. 2016 Dec 6 [cited 2017 Sep 16];87(23):2403–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27815398

113. Hickman SJ, Brex PA, Brierley CMH, Silver NC, Barker GJ, Scolding NJ, et al. Detection of optic nerve atrophy following a single episode of unilateral optic neuritis by MRI using a fat-saturated short-echo fast FLAIR sequence. Neuroradiology. 2001;43(2):123–8.

114. Hickman SJ, Kapoor R, Jones SJ, Altmann DR, Plant GT, Miller DH. Corticosteroids do not prevent optic nerve atrophy following optic neuritis. J Neurol Neurosurg Psychiatry [Internet]. 2003 Aug [cited 2017 Aug 21];74(8):1139–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12876255

115. Hickman SJ, Brierley CMH, Brex P a, MacManus DG, Scolding NJ, Compston D a S, et al. Continuing optic nerve atrophy following optic neuritis: a serial MRI study. Mult Scler [Internet]. 2002 [cited 2017 Aug 21];8(4):339–42. Available from: http://journals.sagepub.com/doi/pdf/10.1191/1352458502ms809oa

116. Hickman SJ, Toosy AT, Jones SJ, Altmann DR, Miszkiel KA, MacManus DG, et al. A serial MRI study following optic nerve mean area in acute optic neuritis. Brain [Internet]. 2004 [cited 2016 Nov 15];127(11):2498–505. Available from: http://brain.oxfordjournals.org/

117. Harrigan RL, Smith AK, Mawn LA, Smith SA, Landman BA. Short Term Reproducibility of a High Contrast 3-D Isotropic Optic Nerve Imaging Sequence in Healthy Controls. Proc SPIE--the Int Soc Opt Eng [Internet]. 2016 Feb 27 [cited 2016 Nov 10];9783. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27175048

118. Lagrèze WA, Lazzaro A, Weigel M, Hansen HC, Hennig J, Bley TA. Morphometry of the retrobulbar human optic nerve: Comparison between conventional sonography and ultrafast magnetic resonance sequences. Investig Ophthalmol Vis Sci [Internet]. 2007 May 1 [cited 2017 Aug 17];48(5):1913–7. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.06-1075

119. Ramli NNM, Sidek S, Rahman FA, Peyman M, Zahari M, Rahmat K, et al. Novel use of 3T MRI in assessment of optic nerve volume in glaucoma. Graefe's Arch Clin Exp Ophthalmol [Internet]. 2014 Jun 27 [cited 2016 Nov 15];252(6):995–1000. Available from: http://link.springer.com/10.1007/s00417-014-2622-6

120. Zhang YQ, Li J, Xu L, Zhang L, Wang ZC, Yang H, et al. Anterior visual pathway assessment by magnetic resonance imaging in normal-pressure glaucoma. Acta Ophthalmol [Internet]. 2012 Jun 1 [cited 2016 Dec 6];90(4):e295–302. Available from: http://doi.wiley.com/10.1111/j.1755-3768.2011.02346.x

121. Chen WW, Wang N, Cai S, Fang Z, Yu M, Wu Q, et al. Structural Brain Abnormalities in Patients with Primary Open-Angle Glaucoma: A Study with 3T MR Imaging. Investig Opthalmology Vis Sci [Internet]. 2013 Jan 17 [cited 2017 Aug 21];54(1):545. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23258150

122. Miller DH, Barkhof F, Frank JA, Parker GJM, Thompson AJ. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain [Internet]. 2002;125(Pt 8):1676–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12135961

123. Birkebæk NH, Patel L, Wright NB, Grigg JR, Sinha S, Hall CM, et al. Optic nerve size evaluated by magnetic resonance imaging in children with optic nerve hypoplasia, multiple pituitary

hormone deficiency, isolated growth hormone deficiency, and idiopathic short stature. J Pediatr [Internet]. 2004 Oct [cited 2017 Aug 21];145(4):536–41. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0022347604005414

124. Karim S, Clark RA, Poukens V, Demer JL. Demonstration of Systematic Variation in Human Intraorbital Optic Nerve Size by Quantitative Magnetic Resonance Imaging and Histology. Investig Opthalmology Vis Sci [Internet]. 2004 Apr 1 [cited 2017 Feb 24];45(4):1047. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.03-1246

125. Lenhart PD, Desai NK, Bruce BB, Hutchinson AK, Lambert SR. The role of magnetic resonance imaging in diagnosing optic nerve hypoplasia. Am J Ophthalmol [Internet]. 2014 Dec [cited 2017 Aug 21];158(6):1164–1171.e2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25128595

126. Mafee MF, Rapoport M, Karimi A, Ansari SA, Shah J. Orbital and Ocular Imaging Using 3- and 1.5-T MR Imaging Systems. Neuroimaging Clin N Am [Internet]. 2005 Feb [cited 2017 Sep 11];15(1):1–21. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1052514905000110

127. Lagrèze WA, Gaggl M, Weigel M, Schulte-Mönting J, Bühler A, Bach M, et al. Retrobulbar optic nerve diameter measured by high-speed magnetic resonance imaging as a biomarker for axonal loss in glaucomatous optic atrophy. Investig Ophthalmol Vis Sci [Internet]. 2009 Sep 1 [cited 2017 Aug 29];50(9):4223–8. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.08-2683

128. Oyama J, Mori K, Imamura M, Mizushima Y, Tateishi U. Size of the intracranial optic nerve and optic tract in neonates at term-equivalent age at magnetic resonance imaging. Pediatr Radiol [Internet]. 2016 Apr 8 [cited 2017 Aug 29];46(4):527–33. Available from: http://link.springer.com/10.1007/s00247-015-3495-5

129. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med [Internet]. 2016 Jun [cited 2017 Sep 7];15(2):155–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27330520

130. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol [Internet]. 2008 Apr 1 [cited 2017 Aug 7];31(4):466–75. Available from: http://doi.wiley.com/10.1002/uog.5256

131. Higgins J, Green S. Chapter 22: Overview of reviews. Cochrane handbook for systematic reviews of interventions. Cochrane Database Syst Rev [Internet]. 2008 [cited 2017 Jul 6];187–235. Available from: http://www.mri.gov.lk/assets/Uploads/Research/Cochrane-Hand-booktext.pdf

132. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J, Glasziou P, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol [Internet]. 2003 [cited 2017 Aug 15];3(1):25. Available from: https://bmcmedresmethodol.biomedcentral.com/track/pdf/10.1186/1471-2288-3-25?site=bmcmedresmethodol.biomedcentral.com

133. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies [Internet]. Vol. 155, Annals of Internal Medicine. American College of Physicians; 2011 [cited 2017 Aug 15]. p. 529–36. Available from: http://annals.org/article.aspx?doi=10.7326/0003-4819-155-8-201110180-00009

134. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane

Collaboration's tool for assessing risk of bias in randomised trials. BMJ [Internet]. 2011 [cited 2017 Aug 3];343. Available from: http://www.bmj.com/content/343/bmj.d5928

135. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. Biostatistics [Internet]. 2001 Dec 1 [cited 2017 Aug 4];2(4):463–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12933636

136. Weigel M, Lagrèze WA, Lazzaro A, Hennig J, Bley TA. Fast and quantitative high-resolution magnetic resonance imaging of the optic nerve at 3.0 tesla. Invest Radiol [Internet]. 2006 Feb [cited 2017 Aug 29];41(2):83–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16428977

137. Yiannakas MC, Toosy AT, Raftopoulos RE, Kapoor R, Miller DH, Wheeler-Kingshott CAM. MRI Acquisition and Analysis Protocol for In Vivo Intraorbital Optic Nerve Segmentation at 3T. Investig Opthalmology Vis Sci [Internet]. 2013 Jun 21 [cited 2017 Aug 29];54(6):4235. Available from: http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.13-12357

138. Morey RA, Selgrade ES, Wagner HR, Huettel SA, Wang L, McCarthy G, et al. Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. Hum Brain Mapp [Internet]. 2010 Nov [cited 2017 Sep 7];31(11):1751–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20162602

139. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England) [Internet]. 1986 Feb 8 [cited 2017 Sep 7];1(8476):307–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2868172

140. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol [Internet]. 2003 Aug [cited 2017 Sep 10];56(8):730–5. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0895435603000842

141. Gala F. Magnetic resonance imaging of optic nerve. Indian J Radiol Imaging [Internet]. 2015 [cited 2017 Sep 10];25(4):421–38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26752822

142. Geeraerts T, Newcombe VFJ, Coles JP, Abate MG, Perkes IE, Hutchinson PJA, et al. Use of T2-weighted magnetic resonance imaging of the optic nerve sheath to detect raised intracranial pressure. Crit Care [Internet]. 2008 [cited 2017 Aug 17];12(5):R114. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18786243

143. Harrigan RL, Plassard AJ, Mawn LA, Galloway RL, Smith SA, Landman BA. Constructing a statistical atlas of the radii of the optic nerve and cerebrospinal fluid sheath in young healthy adults. Proc SPIE--the Int Soc Opt Eng [Internet]. 2015 Mar 20 [cited 2016 Nov 10];9413. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25914505

144. Ederer F. Shall we count numbers of eyes or numbers of subjects? Arch Ophthalmol (Chicago, Ill 1960) [Internet]. 1973 Jan [cited 2017 Oct 17];89(1):1–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4684894

145. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.

146. Huurneman B, Boonstra FN. Assessment of near visual acuity in 0-13 year olds with normal and low vision: a systematic review. BMC Ophthalmol [Internet]. 2016 Dec 8 [cited 2017 Oct 21];16(1):215. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27931205

147. Patel A, Purohit R, Lee H, Sheth V, Maconachie G, Papageorgiou E, et al. Optic Nerve Head

Development in Healthy Infants and Children Using Handheld Spectral-Domain Optical Coherence Tomography. Ophthalmology [Internet]. 2016 Oct 1 [cited 2018 May 27];123(10):2147–57. Available from: https://www.sciencedirect.com/science/article/pii/S0161642016305899

148. Patel NB, Lim M, Gajjar A, Evans KB, Harwerth RS. Age-associated changes in the retinal nerve fiber layer and optic nerve head. Invest Ophthalmol Vis Sci [Internet]. 2014 Jul 22 [cited 2018 May 27];55(8):5134–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25052998

149. Parsa CF. Why Visual Function Does Not Correlate With Optic Glioma Size or Growth. Arch Ophthalmol [Internet]. 2012 Apr 1 [cited 2018 May 27];130(4):521. Available from: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/archophthalmol.2011.1432

150. Saxena R, Bandyopadhyay G, Singh D, Singh S, Sharma P, Menon V. Evaluation of changes in retinal nerve fiber layer thickness and visual functions in cases of optic neuritis and multiple sclerosis. Indian J Ophthalmol [Internet]. 2013 Oct [cited 2017 Nov 1];61(377):562–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24212307

151. Avery RA, Hwang EI, Ishikawa H, Acosta MT, Hutcheson KA, Santos D, et al. Handheld Optical Coherence Tomography During Sedation in Young Children With Optic Pathway Gliomas. JAMA Ophthalmol [Internet]. 2014 Mar 1 [cited 2016 Nov 18];132(3):265. Available from: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/jamaophthalmol.2013.7649

152. Fisher MJ, Avery RA, Allen JC, Ardern-Holmes SL, Bilaniuk LT, Ferner RE, et al. Functional outcome measures for NF1-associated optic pathway glioma clinical trials. Neurology [Internet]. 2013 [cited 2017 Mar 14];81(21 Suppl 1):S15-24. Available from: www.neurology.org

153. Chang L, El-Dairi MA, Frempong TA, Burner EL, Bhatti MT, Young TL, et al. Optical coherence tomography in the evaluation of neurofibromatosis type-1 subjects with optic pathway gliomas. J AAPOS. 2010;14(6):511–7.

154. de Blank PMK, Berman JI, Liu GT, Roberts TPL, Fisher MJ. Fractional anisotropy of the optic radiations is associated with visual acuity loss in optic pathway gliomas of neurofibromatosis type 1. Neuro Oncol [Internet]. 2013;15(8):1088–95. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3714157/

155. Wan MJ, Ullrich NJ, Manley PE, Kieran MW, Goumnerova LC, Heidary G. Long-term visual outcomes of optic pathway gliomas in pediatric patients without neurofibromatosis type 1. J Neurooncol. 2016;129(1).

156. Grover S, Fishman GA, Anderson RJ, Tozatti MS., Heckenlively JR, Weleber RG, et al. Visual acuity impairment in patients with retinitis pigmentosa at age 45 years or older. Ophthalmology [Internet]. 1999 Sep [cited 2017 Oct 12];106(9):1780–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10485550

157. Schulze-Bonsel K, Feltgen N, Burau H, Hansen L, Bach M. Visual Acuities " Hand Motion " and " Counting Fingers " Can Be Quantified with the Freiburg Visual Acuity Test. Invest Ophthalmol Vis Sci [Internet]. 2006 [cited 2017 Oct 12];47:1236–40. Available from: http://www.michaelbach.de/fract/media/lit/SchulzeBonsels2006IOVS_FrACT.pdf

15.0 Appendices

15.1 Appendix A

Assessment of Research Quality

A.1. Quality of Reporting

	Quality of Reporting							
	Clear hypothesis and objectives	Clear type of study design stated	Elligibility of participants: e.g. participant characterisitcs described	Exclusions noted and explained	Appropriate graphs utilised	Study limitations considered		
Scale	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**		
Oyama et al (2016)	2	1	2	2	2	2		
Ramli et al (2014)	2	2	2	2	1	2		
Weigel et al (2006)	2	1	1	2	0	1		
Lagreze et al (2007)	2	1	1	2	0	1		
Yiannakas et al (2010	2	1	1	2	0	2		
Yiannakas et al (2013	2	1	1	2	0	1		
Hickman et al (2001)	2	1	2	2	2	2		
Hickman et al (2004)	2	1	2	2	2	1		
Hickman et al (2002)	2	1	2	2	2	2		
Hickman et al (2003)	2	1	1	2	0	2		
Birkebaek et al (2004)	2	2	2	2	2	2		
Lenhart et al (2014)	2	2	2	2	0	2		
Lagreze et al (2009)	2	1	2	2	2	2		
Trip et al (2006)	2	1	2	2	2	1		
Karim et al (2004)	2	1	1	2	2	1		
	2=explicit	2=explicit; 1=can infer	2=explicit, own section; 1=screening	2=explicit	2=yes, well designed, easily understood; 1=yes but unclear	2=explicit		

A.2. Quality of Methodology

	Quality of methodology								
	All participants subjected to the same method of measurement(s) using the same facilities	Conditions on test subjects were unlikely to introduce bias	Evidence of blinding observers to participant identity and to previous measures made	Appropriate intra-observer test(s) conducted	Appropriate inter-observer test(s) conducted	Appropriate test-retest(s) conducted	Appropriate test-retest interval incorporated	Appropriate reliability statistical test (e.g. ICC, CoV)	Appropriate imaging parameters (Image resolution <1mm, coronal plane)
Scale	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**	**1-2**
Oyama et al (2016)	2	2	1	2	0	1	0	2	2
Ramli et al (2014)	2	2	2	2	2	1	2	2	0
Weigel et al (2006)	2	2	1	1	0	1	0	0	1
Lagreze et al (2007)	2	2	2	2	2	2	2	2	2
Yiannakas et al (2010	2	2	0	2	0	2	2	2	2
Yiannakas et al (2013	2	2	0	2	0	2	2	2	2
Hickman et al (2001)	2	2	1	2	0	2	2	2	2
Hickman et al (2004)	2	2	2	2	0	2	0	2	2
Hickman et al (2002)	2	2	1	2	0	2	0	2	2
Hickman et al (2003)	2	2	2	2	0	2	0	2	0
Birkebaek et al (2004)	2	2	2	2	0	1	0	2	1
Lenhart et al (2014)	1	2	2	2	0	1	0	2	1
Lagreze et al (2009)	2	1	0	0	2	1	0	2	2
Trip et al (2006)	2	2	1	2	0	2	0	2	2
Karim et al (2004)	1	1	0	1	0	1	0	1	0
	2=explicit; 1=inferred	2=explicit; 1=inferred	2=details on blinding method e.g. to their measures, to participant status; 1=state blinding occurred	2=explicit; 1=inferred	2=explicit; 1=inferred	1 - repeated measures, 2 - explicitly says new scans taken, w/ stated interval	2= >2 weeks	1=SD and means	1 = unclear, 2= for yes

A.3. Quality of Research Results

	Quality of reporting score	Quality of methodology score	TOTAL
Scale	**(highest=12)**	**(highest=18)**	**(30 highest)**
Oyama et al (2016)	11	12	23
Ramli et al (2014)	11	15	26
Weigel et al (2006)	7	8	15
Lagreze et al (2007)	7	18	25
Yiannakas et al (2010)	8	14	22
Yiannakas et al (2013)	7	14	21
Hickman et al (2001)	11	15	26
Hickman et al (2004)	10	14	24
Hickman et al (2002)	11	13	24
Hickman et al (2003)	8	12	20
Birkebaek et al (2004)	12	12	24
Lenhart et al (2014)	10	11	21
Lagreze et al (2009)	11	10	21
Trip et al (2006)	10	13	23
Karim et al (2004)	9	5	14

15.2 Appendix B

Recruitment leaflet for NOVEL MRI study