Quantification of Brainstem Depigmentation Using Neuromelanin-MRI in Ageing and Parkinson's Disease

Abdul Halim Sapuan, MSc.

Thesis submitted to the University of Nottingham in fulfilment of the Doctor of Philosophy program

May 2022

Table of Contents

Abstract ..... *iv*

Declaration ..... *ix*

Acknowledgments ..... *x*

List of Author's publications and abstracts ..... *xi*

Abbreviations and Definitions ..... *xiii*

1. Introduction *1*

1.1 Clinical aspects of Parkinson disease **2**

1.2 Age as the primary risk factor in Parkinson disease **9**

1.3 The importance of neuromelanin in ageing and PD **9**

1.4 Imaging in Parkinson disease and the need for biomarker **13**

1.5 MRI and NM detectability **16**

1.6 Quantification of NM-rich region **21**

1.7 Measures of diagnostic test performance **23**

2. Thesis aims, sub studies and hypotheses *25*

3. General Methodologies *29*

3.1 Introduction **30**

3.2 Volunteer selection process **30**

3.3 Image acquisition **31**

3.3.1 Preparation before the scanning 31

3.3.2 Imaging protocols 32

3.4 Image processing **33**

3.4.1 Image quality control 33

3.4.2 Image Processing methods 34

3.4.3 Processing method 1: Semiautomated manual drawing method 35

3.4.4 Processing method 2: Automated template-based method 37

4. Optimisation and Test-retest reproducibility of NM-MRI *41*

4.1 Introduction **42**

4.2 Test-retest Reliability of two image processing methods **44**

4.2.1 Materials and Methods 44

4.2.2 Results 48

4.2.3 Discussion 56

4.2.4 Conclusion 60

4.3 Assessment of two NM-weighted MRI sequences **61**

4.3.1 Materials and methods 61

4.3.2 Results 64

4.3.3 Discussion 70

4.3.4 Conclusion 72

5. Neuromelanin pigment changes from adulthood to old age *73*

5.1 Introduction 74

5.2 Materials and Methods 75

5.2.1 Participants 77

5.2.2 Image analysis 77

5.3 Results 78

5.4 Discussion 84

5.5 Conclusions 88

6. Meta-analysis and systematic review of neuromelanin-sensitive MRI in Parkinson disease 89

6.1 Meta-analysis of Application of Neuromelanin Weighted Imaging and Free-water Imaging in Parkinson's Disease 90

6.1.1 Introduction 90

6.1.2 Methods 91

6.1.3 Results 96

6.1.4 Discussion 117

6.1.5 Limitation and future recommendation 120

6.1.6 Conclusions 121

6.2 Systematic Review of NM-weighted MRI in Parkinson disease and parkinsonian disorders. 123

6.2.1 Introduction 123

6.2.2 Methods 123

6.2.3 Results 125

6.2.4 Limitation and future recommendation 135

6.2.5 Conclusions 135

7. A case control single centre sub study of PaMIR of neuromelanin-weighted imaging in detecting Parkinson disease 137

7.1 Introduction 138

7.2 Materials and methods 138

7.2.1 Participants and data acquisition 139

7.2.2 Image analysis 141

7.3 Results 141

7.3.1 PD Disease effect 141

7.3.2 Laterality and sex effects 144

7.3.3 Diagnostic performance 145

7.3.4 Disease progression and disease severity effect 148

7.4 Discussion 150

7.5 Conclusion 152

8. Can NM MRI detect nigral depigmentation in Clinically Uncertain Parkinsonism: a N3iPD sub study 153

8.1 Introduction 154

8.2 Materials and methods 155

8.2.1 Participants 155

8.2.2 Image Acquisition and processing 157

8.2.3 Statistical analysis 159

8.3 Results 160

8.4 Discussion 167
8.5 Conclusion 168
9. General discussion and Conclusion 169
9.1 Introduction 170
9.2 NM intensity detection and estimation method 170
9.3 Understanding risk factor and pathogenesis of Parkinson 172
9.4 Role as a Parkinson disease biomarker 174
9.5 Conclusion 177
References 179
Appendices 197
Appendix 1 197
Appendix 2 197
Appendix 3 197
Appendix 4 198

Abstract

Background

Parkinson's disease (PD) is the most common form of movement disorder, predominantly affecting those above 60 years old with an upward trend of 2.5 million in 1990 to 6.1 million in 2016. PD diagnosis is currently administered by clinicians, whose skill and experience dependent. Neuromelanin-sensitive MRI (NM-MRI), a relatively new non-invasive imaging tool offers great potential as a quantitative biomarker in detecting neuromelanin depigmentation, a cardinal characteristic of PD. However, its clinical utility remains unclear. This study aims to investigate the clinical usefulness of dedicated NM-MRI in quantifying brainstem depigmentation in PD. The specific objectives were (i) to assess and optimise the quantification methods and acquisition protocols, (ii) to assess physiological ageing effects on Substantia nigra pars compacta (SNpc) depigmentation using a control cohort from adulthood-old age, (iii) to determine the diagnostic potential and accuracy of NM-MRI in Parkinson disease, and (iv) to explore whether NM- MRI may be sensitive to detect pathology in cases with uncertain parkinsonism alternative to dopamine transporter single-photon emission computed tomography (DaTSCAN).

Materials and Methods

A total of 120 normal subjects, 65 PD patients, 36 clinically uncertain Parkinsonism patients, and a meta-analysis of 849 healthy controls and 1213

PD patients from various published studies were analysed. This research compared two 3 Tesla NM sensitive MR imaging sequences and two image processing methods. The SNpc contrast-to-background ratio (CBR) was measured using a semiautomated manual drawing and a recently developed automated template-based method. The following specific research questions were addressed:

(i) to assess and optimise the quantification methods and acquisition protocols

The first study investigated the test-retest reliability of the two abovementioned image processing methods. MRI scans from six participants over a few sessions were analysed. Four aspects of reliability were analysed; intrarater reliability, interrater reliability, short-term rescan reliability and medium-term rescan reliability. All four aspects consistently displayed data points lie within the upper and lower boundaries limit of agreement (LOA) in the Bland-Altman (BA) plots. Hence displaying high agreement of the CBR in dorsal and ventral SNpc for both processing methods. Meanwhile, the second study evaluated the reproducibility of two dedicated NM-sensitive MRI sequences processed using methods assessed in the first study. Data from 69 volunteers comprising 35 controls and 34 PD patients were scanned via both NM-MRI protocols within a single session and analysed. A strong positive linear correlation was observed between sequences in control and PD measured using both methods.

(ii) to analyse a control cohort from adulthood-old age to assess physiological aging effects on SNpc depigmentation

The third analysis elucidated the *in-vivo* lifespan trajectory of neuromelanin in adults aged 22 to 83 years old. Analysis of 94 healthy participants via automated and semiautomated methods showed a strong non-linear effect with an inverted U-shaped association beginning with a linear increase up to a plateau at about the mid-5th decade and decayed until old age.

(iii) to determine the diagnostic potential and accuracy in Parkinson's

The fourth study was a meta-analysis conducted by systematic pooling of 19 articles that focused on NM-weighted imaging and recent potential MR free water imaging to evaluate cumulative diagnostic ability and performance. NM-MRI forest plot with over 2000 population depicted a significant large disease effect size (1.41, $P < 0.00001$). The diagnostic performance of seven articles showed high pooled sensitivity (91%) and specificity (83%), suggesting the potential of NM as a diagnostic biomarker. For free-water imaging, the significant positive disease effect size alongside high heterogeneity suggests that the findings may not be suitable for meta-analysis pooling. The systematic review yielded a consistent outcome supporting the potential application of NM weighted imaging in detecting pigmented SN neuronal loss in detecting the severity of PD and monitoring progression, identifying PD subtyping, distinguishing PD and atypical parkinsonism.

Subsequently, a prospective case-control study involving single centre data from a multi-centre study, Parkinson's MR Imaging Repository (PaMIR)

project with a moderately large cohort of 65 PD patients and 60 controls were scanned from two NM sequences using automated and semi-automated methods. The SNpc CBR of both sequences was normalised and a marked decrease was observed among Parkinson's than controls. In the dorsal SNpc and ventral SN, normalised CBR for healthy controls was significantly higher compared to that of PD patients using both methods ($p<0.0001$). The ventral SNpc demonstrated the best performance in distinguishing healthy controls and PD patients in both methods.

(iv) to explore whether NM- MRI may be sensitive to detect pathology in cases with uncertain parkinsonism alternative to DaTSCAN

The final work explored NM-MRI among clinically uncertain parkinsonian syndrome using the data gathered from an ongoing sub-study of nigrosomal iron imaging in PD (N3iPD). A total of 36 clinically uncertain Parkinson's cases were diagnosed by DaTSCAN and scanned with NM-MRI. Seven of them were identified as negative DaTSCAN group while the remaining 29 were classified as positive DaTSCAN group. NM-MRI using both methods showed consistent reduced CBRs in the positive DaTSCAN group compared to those with negative DaTSCAN group especially in ventral SNpc.

Conclusion

In conclusion, NM-MRI was demonstrated as a useful non-invasive diagnostic biomarker of PD and based on preliminary data that there might be clinical value in using NM to detect earlier with clinically uncertain

parkinsonism. The automated processing approach used limited human resources while the faster acquisition protocol produced higher image contrast in comparison to the counterpart, satisfying the demand of a busy clinical setting.

A major limitation of this research is lack of gold standard data from postmortem study hence data with patients diagnosed using PD Brain Bank Criteria, MDS Parkinson Criteria and DaTSCAN may have a possibility to include false positive and negative data.

Declaration

Except where acknowledged, I declare that this thesis is entirely my own work and is based on research carried out in Radiological Sciences, University of Nottingham between October 2016 and October 2019. I plan to submit and have published some experimental parts of this thesis in peer-reviewed journals as first and second author. I received support from co-authors in the way of feedback and useful suggestions for corrections and improvements while writing up the manuscripts. The materials contained in this thesis has not been presented, nor is it currently being presented, either wholly or in part for any other degree or qualification.

Abdul Halim Sapuan

Acknowledgments

First and foremost, I would like to thank the volunteers contributed to this study so this PhD journey could progress and completed.

Secondly, I would love to thank my supervisors Prof Dorothee P Auer and Dr Yue Xing for the guidance and supports throughout this colourful journey. Thanks to Dr Saadnah Naidhu and Rebecca Stevenson as a colleague in the team for being supportive in the teamwork.

Thirdly, special thanks to my family especially my wife and son as a source of motivation to keep this journey running. Not forgetting my parents, siblings, parent in-law that indirectly involve in this journey.

Last but not least thanks to everyone that directly or indirectly contributed to the study which I did not recognised.

List of Author's publications and abstracts

Peer reviewed Publications:

Xing, Y., **Sapuan, A.,** Dineen, R. A., & Auer, D. P. (2018). Life span pigmentation changes of the substantia nigra detected by neuromelanin-sensitive MRI. *Movement Disorders*, *33*(11), 1792–1799. https://doi.org/10.1002/mds.27502

Xing, Y., **Sapuan, A. H.**, Martín-Bastida, A., Naidu, S., Tench, C., Evans, J., Sare, G., Schwarz, S. T., Al-bachari, S., Parkes, L. M., Kanavou, S., Raw, J., Silverdale, M., Bajaj, N., Pavese, N., Burn, D., Piccini, P., Grosset, D. G., & Auer, D. P. (2022). Neuromelanin-MRI to Quantify and Track Nigral Depigmentation in Parkinson's Disease: A Multicenter Longitudinal Study Using Template-Based Standardized Analysis. *Movement Disorders : Official Journal of the Movement Disorder Society*. https://doi.org/10.1002/MDS.28934

Abstracts and congress contributions:

Neuromelanin-weighted MRI in revealing human development and age-related changes in locus coeruleus and substantial nigra - Division of Clinical Neuroscience Research Day 2017, School of Medicine - Best poster presenter

Neuromelanin and Free-water MR Imaging as Promising Biomarkers in Parkinson's disease: A Meta-Analysis of the Disease Effect Size and Diagnostic Accuracy, International Society for Magnetic Resonance in Medicine (ISMRM) 27th Annual Meeting & Exhibition 2019 Conference, Abstract ID: 7299-Accepted.

Authors: Abdul Sapuan, Rebecca Stevenson, Saadnah Naidu, Yue Xing, Stefan T. Schwarz and Dorothee P Auer

Peer reviewed publications in progress

Meta-Analysis of Neuromelanin and Free-water MRI as Diagnostic Biomarker in Parkinson's Disease,

Authors: Abdul Sapuan, Rebecca Stevenson, Saadnah Naidu, Yue Xing, Stefan T. Schwarz and Dorothee P Auer

Submitted to 'Movement Disorder 'in 2019, resubmission plan in 2022.

Abbreviations and Definitions

3T	3 Tesla magnetic field strength
ACE	Addenbrooke's cognitive examination test
AUC	Area under the curve
BA	Bland Altman, a method of plotting to analyse agreement between two measurements
CBD	Corticobasal degeneration
CBR	Contrast-to-background ratio
CI	Confidence Interval
CNR	Contrast-to-noise ratio
CUPS	Clinically uncertain parkinsonian syndrome
DaTScan	Dopamine transporter Single Photon Emission Computed Tomography
DLB	Dementia with Lewy bodies
DTI	Diffusion Tensor Imaging
ET	Essential tremor
FA	Fractional Anisotropy
FSE	Fast Spin Echo
GRE	Gradient recalled echo
HC	Healthy control
H&Y	Hoehn and Yahr
ICC	Intraclass correlation coefficient
iRBD	Idiopathic Rapid Eye Movement Sleep Behaviour Disorder

LC	Locus coeruleus
LOA	Limits of agreement
LSPD	Late-stage Parkinson disease
MD	Mean diffusivity
MDS	Movement Disorder Society
MRI	Magnetic resonance imaging
MSA	Multiple System Atrophy
MSA-P	MSA with pre-dominant Parkinsonism
NIH	National Institutes of Health
NM	Neuromelanin
NM-MRI	Neuromelanin sensitive magnetic resonance imaging
OCT	Optical coherence tomography
PD	Parkinson's disease
PET	Positron emission tomography
PIGD	Postural Instability and Gait Disturbance Patients
PSP	Progressive Supranuclear Palsy
RBD	Rapid Eye Movement (REM) sleep behaviour disorder
Reliability	measures of precision or the extent to which test results can be replicated
Repeatability	variability of recurring measurements on subjects in a single condition
Reproducibility	variability of repeat measurements on subjects in different conditions
ROI	Region of Interest

ROC	Receiver Operating Characteristics
SBR	Striatal specific binding ratio
SCA	Spinocerebellar ataxia
SD	Standard deviation
SE	Standard error
SMD	Standardised mean difference
SN	Substantia Nigra
SNpc	Substantia nigra pars compacta
SNr	Substantia nigra pars reticulata
SNR	Signal-to-noise ratio
SPECT	Single Photon Emission Computed Tomography
T1W	T1 weighted
TCS	Transcranial Ultrasound
TD	Tremor Dominant
TSE	Turbo Spin Echo
UK	United Kingdom
UPDRS	Unified Parkinson's disease Rating Scale

1. Introduction

1.1 Clinical aspects of Parkinson disease

Parkinson's disease (PD) was first discovered by James Parkinson in 1817 but was given its name much later in 1872 by Jean-Martin Charcot (Parkinson, 2002). Originally termed "paralysis agitans", James characterised the disease through tremors, posture/gait abnormalities, and muscle strength reduction. In 1912 Frederic Lewy identified an association between protein aggregates in the brain with PD patients, the particles were later dubbed "Lewy bodies" after himself (Holdorff, 2002).

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease, affecting approximately 1% of the world population aged 60 years old and above (de Lau & Breteler, 2006; Hirtz et al., 2007; Tysnes & Storstein, 2017). Individuals affected by PD were approximately 6.1 million in 2016 compared to 2.5 million in 1990. The number is expected to double by 2040, thus posing a global cause for concern (Dorsey, Elbaz, et al., 2018). The anticipated spike could potentially increase the societal burden, especially with the growing ageing population. Besides, the economic impact on healthcare costs attributed to PD individuals is set to rise following diagnosis. The estimated mean annual health care cost difference between healthy vs PD populations in 2013 was £2471 (US$3716) per patient in the first year upon diagnosis. Meanwhile, the cost appeared to rise significantly to £4004 (US$6021) by 10-year postdiagnosis as healthcare utilisation increased (Weir et al., 2018). The future prevalence and impact of the disease are set to increase as life

expectancy grows. To date, no treatment has been established to stop the progression of PD, but some treatments have proven to reduce associated symptoms.

PD is corroborated through both motor symptoms and non-motor symptoms. The four main associated motor dysfunctions are tremor, rigidity, postural instability, and bradykinesia. All symptoms may not manifest concurrently, and the disorder does not have to exist in its entirety for confirmatory diagnosis, especially during the early stages of the disease.

Tremor is typified through the rhythmic shaking of hands, arms, or legs when at rest by a frequency of 4 to 6 Hz. It usually starts slowly on one side, subsequently affecting both the left and right sides of the body. In some instances, resting tremors may also occur in the jaw, tongue, mouth, and chin. Rigidity describes abnormal stiffness or tightness of limbs or parts of the body. It is often mistakenly assumed as an orthopaedic problem or arthritis, particularly in the early stages of PD. Meanwhile, postural instability refers to impaired balance or difficulty in standing or walking, besides the inability to maintain upright posture or to prevent falls. Bradykinesia or slowed movement is a common movement disorder symptom. Bradykinesia typically co-occurs with reduced facial expression and rate of blinking coupled with fine motor coordination difficulties. These core symptoms are physically examined and associated with nigrostriatal dopaminergic loss (Marras & Chaudhuri, 2016).

Non-motor symptoms on the other hand are usually more obscure compared to motor symptoms and may be present during the prodromal

phase for several years prior to motor symptoms. Apart from severe mental disorders such as depression and anxiety, non-motor symptoms are often regarded as a significant source of decreased quality of life (Chaudhuri et al., 2006). Premotor symptoms in PD may include dysautonomia (such as orthostatic symptoms), urinary symptoms and constipation, REM-sleep behaviour disorder (RBD), excessive daytime sleepiness, restless legs, pain, and anosmia (Lang, 2011; Noyce et al., 2016). Irregularity on the symptom development which is tailored to individual differences makes PD inconspicuous. This event was suggested as a sign of risk, which requires screening and monitoring for early detection of the disease (Hughes et al., 1992).

The diagnosis of PD is usually performed clinically and it's a non-straightforward process. Several guidelines and diagnostic criteria have been introduced over the past three decades to facilitate clinicians in identifying specific signs or symptoms denoting the presence of parkinsonian syndromes, besides assisting the classification of patients into idiopathic Parkinson's or other types of parkinsonism. The introduction of these diagnostic criteria has improved PD diagnostic accuracy up to 82% (Gelb et al., 1999) compared to the prior 75% from the gold standard post mortem examination (Rizzo et al., 2016).

Queen Square Parkinson Brain Bank Criteria, commonly known as United Kingdom (UK) Brain Bank Criteria (Hughes et al., 1992), diagnose PD using a three-step process. The first step focuses on diagnosing parkinsonian

syndrome, the second step observes PD exclusion criteria, and the third evaluates prospective supportive criteria of PD.

The three-step approach is summarised as follows:

Step 1: Diagnosis of Parkinsonian syndrome requires the presence of the following:

- Bradykinesia
- at least one of the other symptoms such as muscular rigidity,
- 4-6 Hz rest tremor,
- postural instability not caused by primary visual, vestibular, cerebellar, or proprioceptive dysfunction.

Step 2: Exclusion criteria for Parkinson's disease

Checking for any medical history of the patient that allows exclusion of Parkinson's.

- history of repeated stroke with a stepwise progression of parkinsonian features,
- history of repeated head injury,
- history of definite encephalitis,
- oculogyric crises,
- neuroleptic treatment at the onset of symptoms,
- more than one affected relative,
- sustained remission,
- strictly unilateral features after 3 years,
- supranuclear gaze palsy,

- cerebellar signs,
- early severe autonomic involvement,
- early severe dementia with disturbances of memory, language, and praxis,
- Babinski sign,
- presence of a cerebral tumour or communication hydrocephalus following an imaging study,
- negative response to large doses of levodopa in absence of malabsorption,
- MPTP exposure.

Step 3: Supportive prospective positive criteria for PD

Three or more criteria as follows are required for PD diagnosis inclusive of step 1 above:

- unilateral onset,
- rest tremor present,
- progressive disorder,
- persistent asymmetry affecting side of onset most,
- excellent response (70-100%) to levodopa,
- severe levodopa-induced chorea,
- levodopa response for 5 years or more
- clinical course of ten years or more.

An improved version known as the Movement Disorder Society Parkinson Disease Criteria (MDS-PD criteria) was introduced in 2015 (Postuma et al., 2015). Though established mainly for research purposes, it could also be used within clinical settings for the PD diagnosis (Postuma et al., 2015). The MDS-PD criteria maintained motor syndrome as the core criterion, but also incorporated several non-motor manifestations, which is crucial during the prodromal stage or early PD (Daniela Berg et al., 2015).

Employing a similar three-step principle of the UK Brain Bank criteria, the MDS-PD proposed examining the prodromal PD and certainty levels, thus allowing the characterisation of Parkinsonian and no Parkinson conditions. Given that this revised criteria is fairly new, a cross-sectional study involving 434 PD patients recruited from eight was conducted to compare the accuracy between the two methods (Postuma et al., 2018). Resultantly, by taking expert clinical diagnosis as gold standard, the accuracy for MDS-PD criteria was 94.5% while UK Brain Bank criteria was 86.4% (Postuma et al., 2018).

Clinical assessment scales of disease severity in Parkinson's

Upon diagnosis, physician will assess extent of motor and non-motor symptoms of patient based on certain criteria to determine the severity of the disease. Hoehn and Yahr were the first to describe tremor as a progression marker of PD in 1967 (Connie Marras et al., 2002). Five severity scales (i.e., I, II, III, IV and V) were formed to facilitate patient management. The updated version of the scales were the rating scales developed to assess motor impairment - the Unified Parkinson's disease rating scale (UPDRS) and

revised version labelled Movement Disorder Society (MDS-UPDRS) (Goetz et al., 2007). The UPDRS assessment includes the Modified Hoehn and Yahr scale, as well as the Schwab and England Activities of Daily Living Scale. UPDRS assessment comprises the following six sections:

1. Mentation, behaviour, and mood
2. Activities of daily living
3. Motor examination
4. Complication of therapy
5. Modified Hoehn and Yahr scale
6. Schwab and England Activities of Daily Living Scale

Each section is quantified through the scores provided based on the assessment of an expert clinician, followed by allocating a total score ranging from 0 (no disability) to 199 (total disability). The rating process is skill and experience-dependent, which could be affected by the patient's medication status during the examination (Post et al., 2005). Withdrawal of dopaminergic medication for more than 12 hours is often regarded as 'off medication'. However, some pharmacological studies reported that withdrawal effects could take up to two weeks to completely clear (Hauser et al., 2000). These factors and variations make rating for PD severity more challenging.

1.2 Age as the primary risk factor in Parkinson disease

Ageing is the most important risk factor for PD (Rao et al., 2006). The prevalence of PD is low in populations below 50 years old (de Rijk et al., 2000). Nevertheless, the prevalence increases with age affecting 5% of the population above over 85 years (de Lau & Breteler, 2006). A recent meta-analysis pooled four North American populations and concluded that the prevalence of PD had risen from less than 1% aged 45 to 54 years to 4% and 2% of males and females aged 85 years or older (C Marras et al., 2018). The global prevalence is expected to double in the next two decades as the population of 60 years and above is predicted to triple by 2100 (Dorsey, Sherer, et al., 2018; Ferri et al., 2005). Therefore, these epidemiological studies highlighted the critical need for understanding normal ageing and its impact on the development of PD. Ageing is a complex process that is still poorly understood. Furthermore, ageing is an important risk factor that increases the susceptibility toward PD and attenuates the capacity for self-repair (Y. Wang et al., 2017).

1.3 The importance of neuromelanin in ageing and PD

Anatomy of substantia nigra and neuromelanin

Substantia nigra (SN) (Figure 1) is a tiny brain tissue located inside the midbrain that is connected to the basal ganglia and participates actively in the motor and reward system. SN is divided into pars compacta (SNpc) and pars reticulata (SNr), with both offering different roles and being connected to separate neuronal networks. SNpc is located at the dorsal SN and is

composed of dopaminergic neurons that provide input to the basal ganglia and supply the striatum with dopamine, which are both crucial areas for PD pathophysiology. On the other hand, SNr is located ventrally and primarily serves as a relay centre, transmitting signals from the basal ganglia to other brain areas. Substantia nigra is a Latin word that translates to 'black substance'. Hence, it appears darker than its neighbouring region as a result of high neuromelanin (NM) pigment content in the neurons from dopaminergic metabolism (Double et al., 2003; Zecca et al., 2003).

Neuromelanin is a dark complex and insoluble molecule measuring approximately 0.5 to 2.5 μm. It is observed in the form of pigmented granules that are present in the brain, especially among dopaminergic neurons of the SNpc and in the noradrenergic neurons of locus coeruleus (L Zecca et al., 2004; Luigi Zecca et al., 2008). The function of neuromelanin remains elusive but investigators have hypothesised that neuromelanin in SN protects neurons from oxidative stress, which is mediated by free metals or free radicals (Braak et al., 2003). It also prevents neuronal toxicity by reacting with hydrogen peroxide (Fabio A. Zucca et al., 2014). Due to its high tendency to chelate iron, it seems that the accumulation of iron determines whether this molecule acts as a protective agent blocking redox-active metal ions or promotes the formation of cytotoxic radicals in the presence of excess iron (Zucca et al., 2004).

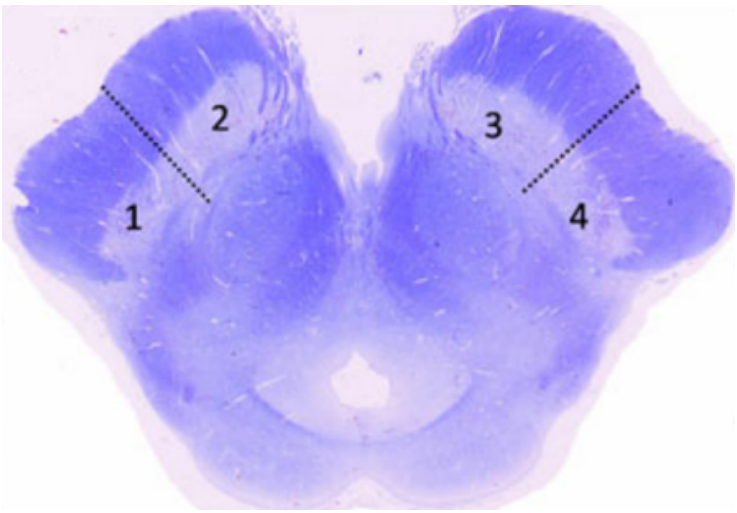


Figure 1. Substantia nigra in an axial slice of a myelin-stained section. Subdivided into four areas (1: right ventral, 2: right dorsal, 3: left dorsal, 4: left ventral). Image adapted from (Kitao et al., 2013).

Role of substantia nigra and neuromelanin in ageing and PD

The involvement of SN in the pathogenesis of PD was discovered by Paul Oscar Blocq and Georges Marinesco in 1893 (Braak et al., 2003; Jellinger, 2012; Schulz-Schaeffer, 2010; Spillantini et al., 1997). Later in 1919, a Russian neuropathologist, Konstantin Tretiakoff, in his dissertation described nigral cell loss in Parkinson's. His work was doubted by many, including his own colleagues (Parent & Parent, 2010). However, his findings were later supported by several neuropathological studies denoting a late recognition of his initial work in the field.

Motor dysfunction associated with PD manifests as a result of the damage of the critical components of the motor system, commonly attributed to dopaminergic loss in the SNpc and striatum, as well as the nigrostriatal pathway system (Schrag et al., 2015). The loss of SN neurons in the areas most affected by ageing (Costa, 2014) has been associated with the depigmentation of NM, which is a fundamental diagnostic criterion for the

disease (L Zecca et al., 2001, 2004). In fact, pathological diagnosis is the current gold standard of PD diagnosis in which the accuracy of clinical diagnosis is commonly referred to.

Previous studies discovered a reduced level (Mann, 1983) and an inverse relationship between the amount of NM in SNpc and the relative vulnerability of these cells to PD (Hirsch et al., 1988). Additionally, this loss is non-uniform and varies across regions, where some appear more vulnerable than others (Esiri, 2007; Pakkenberg et al., 1991) with the ventrolateral tier being the most severely affected (Gibb & Lees, 1991). Apparently, neuromelanin concentration in SN is depleted by approximately 60% in PD compared to healthy controls (HC), thus confirming the loss of NM-containing neurons of the SN (Luigi Zecca et al., 2002).

During normal human lifespan, previous histological studies have demonstrated the presence of neuromelanin granules as early as 2 to 3 years old (Cowen, 1986). The pigment either accumulates with age (L Zecca et al., 1996; Fabio A. Zucca et al., 2014) or stabilises in size and number after 20 years of age (Fearnley & Lees, 1991; Ma et al., 1999; McGeer et al., 1977). A few studies also reported age-related loss of NM (Fedorow et al., 2005). Based on these findings, studies relevant to NM changes during ageing or across human lifespans may be beneficial in understanding the emergence of PD.

1.4 Imaging in Parkinson disease and the need for biomarker

Parkinson's disease is highly heterogenous and its symptoms largely overlap with parkinsonism and atypical parkinsonian disorder, especially during the early course of the disease. Despite the establishment of clinical criteria, the clinical diagnosis remains difficult. Motor symptoms are subtle at early stages and physical examination does not yield any vital diagnostic information. Even if PD is diagnosed at early stages, the development of other symptoms over time may cause a change in diagnosis.

A recent meta-analysis of the clinical diagnosis of PD in the last 35 years from 1988 to 2014 concluded that the accuracy is unsatisfactory (Rizzo et al., 2016). A total of 20 studies using pathologic diagnosis as gold standards were analysed and the pooled accuracy was 80.6% pooled accuracy. The accuracy of studies that employed clinical diagnosis by nonexperts was 73.8%, 79.6% of initial assessment by movement disorder experts, and 83.9% after follow-up. Meanwhile, a pooled accuracy level of 82.7% was documented in studies that used the widely accepted standard diagnostic criteria, the United Kingdom (UK) Brain Bank criteria (Rizzo et al., 2016).

A study involving the monitoring of 100 PD patients from diagnosis until death was conducted using the UK Brain Bank criteria. During the post-mortem examination, 24% of the cases were identified as other differentials of PD, such as progressive supranuclear palsy (PSP), multiple system atrophy (MSA), and vascular parkinsonism due to lacunar infarcts, which were co-located with the SN. Resultantly, an error rate of approximately 10 to 25%

was highlighted in the clinical diagnosis, and the statistics included cases of misdiagnosis of PD and missed PD (Hughes et al., 1992, 2002). These findings depict the need for diagnostic tools or biomarkers to achieve reliable diagnoses even in non-highly specialised clinical settings.

'Biomarker' is a term often used to describe changes in a gene or protein expression. However, the National Institutes of Health (NIH) Biomarkers Definitions Working Group broadly defines a biomarker as "*a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention*" (Rizzo et al., 2016). Specific criteria and processes need to be fulfilled for a substance to be qualified as a clinical biomarker.

Presently, an established biomarker for PD is dopamine transporter single-photon emission tomography (DaTScan). The tool is however limited to differentiating essential tremor from PD and is unable to diagnose PD from normal healthy controls. Furthermore, the outcome of this biomarker can only be translated into diagnosis by a movement disorder specialist. Another limitation of this tool is its high cost due to special storage and transportation requirement.

Transcranial ultrasound (TCS) is an inexpensive and radiation-free alternative practised by European countries like Germany and Austria. This method has shown the accuracy of up to 82% by identifying increased echogenicity in the SN among PD patients (Van De Loo et al., 2010).

Nonetheless, it is machine- and operator-dependent and requires a good temporal bone window for diagnosis.

A more recent biomarker is diffusion MRI-derived free water estimates, which serve as an indirect marker of SN neurodegeneration by estimating regional free water maps using a multi tensor model (Ofori et al., 2015). A large MRI repository data, Parkinson's Progressive Marker Initiative (PPMI), have shown region-specific increased free water in the posterior SN in PD (E Ofori et al., 2015). The increased levels of free-water reflect increased water pool in the extracellular space attributed to index cell loss and atrophy (Albi et al., 2017; Pasternak et al., 2009).

A few possible beneficial biomarkers in PD need to be explored. The most important would be a diagnostic biomarker for early detection and accurate diagnosis of PD. Other than that, biomarkers to monitor disease progression or estimate PD risk and identify disease severity would also be highly beneficial. The ultimate objective of unearthing biomarkers is to identify those that are reliable with high sensitivity and specificity, non-invasive, widely available, and reasonably priced.

Magnetic resonance imaging has advantageous features as a good biomarker with high potential given that it is widely available, poses no radiation risk, and non-invasive. Despite the extensive research in PD diagnosis, there are no established MRI biomarkers for PD to date. The chapters in this study seek diagnostic markers to assist in differentiating PD from healthy control as the current standard of practice lacks acceptable

accuracy and require experience and high levels of expertise (Postuma et al., 2015).

Recent decades have seen developments in neuroimaging techniques that revolutionised PD research. Advanced MR imaging techniques with stronger magnetic field increases signal-to-noise ratio (SNR), improves contrast and spatial resolution, thus allowing visualisation of tissue microstructure. The advancement in technology has garnered interest among researchers to better understand the pathogenesis of PD in order to develop disease-modifying therapies, detect PD in its initial stages, and uncover new biomarkers to assist in diagnosis and monitoring of disease progression.

1.5 MRI and NM detectability

Application of MRI on visualisation of NM to study PD was first reported in 2006 (Sasaki et al., 2006) in which a high signal intensity of SN was associated with the presence of NM upon utilising the 3T T1 weighted turbo spin-echo (TSE) protocol. This technique was later known as "neuromelanin-sensitive MRI" (NM-MRI), and it has successfully provided the contrast between NM-rich and the surrounding areas.

Since the inception of NM-MRI, several researchers (Langley et al., 2015; Schwarz et al., 2013) have reported a reduction in signal intensity corresponding to the SN among PD patients. These aforementioned results reflect the potential application of NM for the diagnosis of PD (Cassidy et al., 2019). Wang et al. (2019) also concluded that NM might be a reliable biomarker for PD severity and subtype identification. Collectively, the findings

corroborate that NM-related signal intensity loss reflects the depigmentation of SN neurons. In addition, a post-mortem study (Kitao et al., 2013) that directly compared NM signal intensity and neuropathological findings revealed that the quantity of pigmented neurons is closely associated with the NM signal intensity. This makes NM-MRI a promising technique for visualising pathological changes in PD, which may allow the quantification of the neuromelanin signal loss. Furthermore, extensive research has focused on applying the technique to diagnose PD, and monitor the disease progression, aside from differentiating between PD and other types of parkinsonism.

The use of high-resolution spin-echo T1 weighted NM-MRI can elicit the neurons with NM in the brain to display a high signal area, thereby reflecting the signal intensity and NM distribution (Sulzer et al., 2018). The loss of high signal area can be detected by NM-MRI sequence during the early phases of PD, thus facilitating the implementation of a non-invasive approach that focuses on the degree of NM signal intensity to elucidate the disease progress. This was demonstrated in a recent study that entailed the relationships between NM concentration in *in vivo* SN and PD progression and various motor subtypes by utilising NM-MRI (Wang et al., 2019). The MRI images of PD patients demonstrated significantly decreased contrast to noise ratio (CNR) values in medial and lateral SN compared to the healthy or control subjects. Likewise, the CNR values in the lateral SN region decreased linearly with PD progression. Interestingly, the decrease in CNR values was more pronounced in the postural instability and gait difficulty group compared to the tremor dominant group.

With NM-MRI, the SNc containing the highest NM content appears as an area of higher signal intensity while other structures with lower NM content than SNc, such as the caudate nucleus, globus pallidus, and putamen, do not reflect intense NM-MRI signals. This phenomenon is believed to generate the primary source of contrast in NM-MRI, which results from an increased contrast detected with magnetisation transfer (MT) preparation of NM-MRI pulse sequences (Sulzer et al., 2018). An observable MT effect is the enhancement of contrast between SN as compared to surrounding areas, hence it is regarded as the primary source of NM-MRI contrast (Nakane et al., 2008).

In human SN, the main iron protein is ferritin with a characteristic low paramagnetic property. Meanwhile, NM-iron complexes are the main proteins in pigmented SN neurons which are highly paramagnetic passes some of its magnetic properties to the surrounding protons (Bolzoni et al., 2002; Enochs et al., 1989, 1997; L Zecca et al., 2001). This difference in paramagnetic properties is attributed to the composition of the chelating groups, which are catechols with a stable free radical in NM-iron complexes found in pigmented SN neurons while the ferritin in normal human SN contains amino-acids (Engelen et al., 2012).

The mechanistic relationship between NM and MRI signal is vital in understanding the potential role of NM in PD. Magnetisation transfer occurs when longitudinal magnetisation is transferred from the macromolecular ‘bound’ hydrogen nuclei with restricted motion to hydrogen nuclei of water or ‘free’ water pool, which is highly mobile with higher degrees of freedom.

In other words, contrast is increased by MT pulses by utilising minimal differences of the magnetic moments of macromolecular bound protons and free water protons (Sulzer et al., 2018).

Protons must possess long enough T2 relaxation times above 10 ms for signal to be detected in MRI. Such a relaxation time facilitates the spatial encoding gradients to be played out between excitation and acquisition prior to the complete decay of the signal. Nevertheless, endogenous proteins such as neuromelanin are characterised by protons with a very short T2 relaxation time, which is usually less than 1ms. Thus, they are not directly detectable by MRI. Magnetisation transfer pulse is thus combined with the T1 weighted sequences to enhance contrast of neuromelanin to the surrounding brain areas. As described by Henkelman et al. (2001), some spin properties of the protons that are physiologically associated with the bound protons or macromolecules are transferred to the surrounding 'free' water pool.

The Larmor equation: 127.728 MHz for free water protons in an external magnetic field of 3 Tesla, describes frequency of the precession of the magnetic moments of water protons in relation to an external magnetic field. The 'bound' protons are precessed at a slightly different frequency due to the influence of macromolecules, which is the main feature exploited in MT imaging. The protons associated with the macromolecules become saturated to a "free-liquid state" upon application of an 'off resonance' radiofrequency pulse at a frequency lower than the Larmor frequency as shown in Figure 2 below. Grossman et al. (1994) described MT as the alteration in longitudinal magnetisation where the restricted protons energy

exchange between rigid macromolecules and free water protons, thereby reducing the signal (partial saturation) of the free water. This event heightens the contrast of macromolecular bound protons and the 'free water pool' (Henkelman et al., 2001). In comparison to the line-shape of the liquid 'free' water pool, the precession frequency of magnetic moments of the macromolecular 'bound' proton pool is much broader.

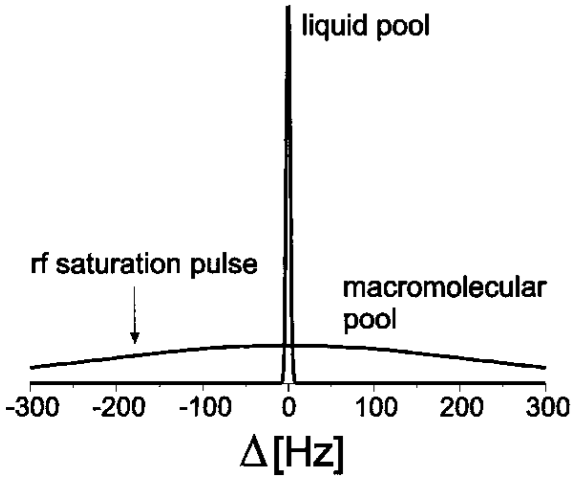


Figure 2: Diagram illustrating a macromolecular pool which is saturated selectively by an 'off' resonance of a saturation pulse without affecting the liquid pool (Adopted from Henkelman et al., 2001).

Since the discovery of contrast NM-MRI sequence, no standardised or established protocol has been accepted as the best protocol to image NM. Therefore, two dedicated MRI protocols were assessed in this thesis to validate their usability. Figure 3 depicts the NM contrast produced using one of the dedicated NM-MRI protocols at Sir Peter Mansfield Imaging Centre, University of Nottingham.

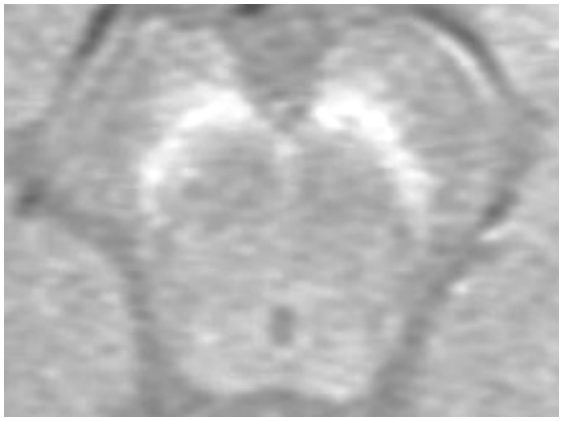


Figure 3. Hyperintense neuromelanin concentrated areas of a 35-year-old healthy person scanned by an NM-MRI protocol at our dedicated MRI equipment.

1.6 Quantification of NM-rich region

Since NM-MRI investigates a region-specific area in the brain, the region of Interest (ROI) would be the most suitable method of analysis. ROI analysis within this context refers to selecting a group of voxels from the SN or whole SN region to investigate the regional effect. The selection is performed using a commercially available image analysis software, either by creating a small space (commonly a sphere with a certain radius) or selecting the whole SN hyperintensity as an area metric or a volumetric approach. A cluster of voxels of low intensity adjacent to the SN is also selected as the background region to identify the regional effect. Image processing methods are conducted through three techniques: manual, semi-automated, or automated methods, depending on the software and technical expertise available within a research team. Automated segmentation automatically detects the SN region using a reference atlas while a semi-automated method is an assisted manual technique that requires the observer to determine the ROI selection.

Upon ROI selection, NM filled voxels are computed using two measurement methods: contrast ratio or/and 2D/3D morphometry. The most common method is the contrast ratio or contrast-to-background ratio. This method could yield a relative SN density or concentration measure, with the lower ratio corresponding to dopaminergic neuronal density loss. The mean intensity within the circled ROI is computed against the background ROI, followed by calculating the contrast ratio using any of the following two formulas:

$$\text{Contrast ratio} = \frac{(\ SN\ ROI\ intensity - \ background\ ROI\ intensity}{background\ intensity} \quad \text{(Formula 1)}$$

$$\text{Contrast ratio} = \frac{SN\ region}{background\ region} \quad \text{(Formula 2)}$$

The second method is a morphometric method that is measured by calculating the SNpc volume, which indirectly indicates the actual size of the SNpc in the human brain. It can also be performed by using 2D morphometry, which measures the two-dimensional size, width, and length of the SNpc.

Similar to the protocol, there is no standardised NM signal quantification method. Researchers have developed and employed their own dedicated methods to analyse NM signals. This condition makes it more challenging for this tool to be applied in clinical settings. Therefore, the validity of post-processing measurement methods and technical variation

assessment is equally important. This thesis started with the reliability test of quantification approaches and protocols for validation.

1.7 Measures of diagnostic test performance

Diagnostic test performance or diagnostic accuracy refers to the efficiency of a diagnostic procedure in differentiating abnormal from normal health conditions. This is represented through the measures of accuracy such as sensitivity, specificity, and the area under the curve (AUC) or receiver operating characteristic (ROC). These measures cover not only discriminative aspects but other aspects of performance such as predictive ability as presented in Table 1 below.

Table 1. Measure of diagnostic performance and brief explanation of each term.

<table>
<tr><th>Measure</th><th>Brief Explanation</th></tr>
<tr><td>Sensitivity</td><td>Defined as ratio of positive test result on diseased patient. Refers to how a test recognises the patient with targeted condition (gold standard positive).</td></tr>
<tr><td>Specificity</td><td>Complementary to sensitivity, it is the ratio of negative test result on non-diseased subjects</td></tr>
<tr><td>Receiver Operating Characteristic (ROC) curve</td><td>A plot devised based on sensitivity and specificity tested across a range of values. The shape and the area under the curve (AUC) estimate discriminative power between patient and non-patient. The larger the AUC and closer the curve to the upper-left hand corner, the better its discriminating performance. The AUC ranges between 0 and 1, where 1 denotes the best performing test.
<table>
<tr><th>Area</th><th>Diagnostic accuracy</th></tr>
<tr><td>0.9-1.0</td><td>Excellent</td></tr>
<tr><td>0.8-0.9</td><td>Very good</td></tr>
<tr><td>0.7-0.8</td><td>Good</td></tr>
<tr><td>0.6-0.7</td><td>Sufficient</td></tr>
<tr><td>0.5-0.6</td><td>Bad</td></tr>
<tr><td><0.5</td><td>Test not useful</td></tr>
</table>
</td></tr>
</table>

The diagnostic test performance of NM-MRI is evaluated in chapter 6 and 7 of this thesis in order to understand its discriminative performance between PD patients and healthy subjects.

2. Thesis aims, sub studies and hypotheses

The main aim of this thesis was to assess the clinical applicability of NM MR imaging protocol in the prediction and diagnosis of Parkinson's disease towards the study of SN. This was achieved by (i) optimisation of the acquisition and quantification protocol via assessment of its test-retest reliability, (ii) *in-vivo* examination of age-dependent changes in NM pigment, (iii) a meta-analysis of case-control studies, (iv) a moderately large case-control study, and (v) exploration of sensitivity to dopaminergic deficits in clinically uncertain parkinsonism.

The thesis work is structured in five study chapters addressing the following specific aims/research questions:

(1) Optimisation and Test-retest reproducibility of NM-MRI

   a. Compare intrarater and interrater reliability for two image processing methods: automated template-based and semiautomated method.
   b. Evaluate test-retest reliability of two NM-MRI protocols through repeated scanning of healthy participants and analysing error bias from the measurement of contrast-to-background ratio.
   c. Examine two NM-sensitive MRI sequences of similar individuals in healthy and Parkinson's patients to visualise the NM-rich region

(2) NM-MRI and physiological brain ageing from adulthood to old age
    a. Evaluate physiological ageing effects in normal human population on nigral neuromelanin contrast using NM-MRI *in-vivo* from 20 years until 82 years of age.
(3) Systematic Review and Meta-analysis of diagnostic value of NM-MRI
    a. Determine NM-MRI diagnostic performance via systematic review and meta-analysis.
    b. Compare NM-MRI diagnostic accuracy with free water MRI via systematic review and meta-analysis
    c. Discover NM-MRI application in PD and parkinsonism
(4) Diagnostic ability and accuracy of NM-MRI in PD: A single centre case-control sub-study
    a. Quantify neuromelanin-related MRI signal changes in SN and its subregions in PD patients and controls
    b. Assess interrelationships between disease progression and severity with neuromelanin signal changes
(5) Diagnostic potential of NM-MRI in CUPS: An explorative transmodality comparison with DaTSCAN
    a. Explore whether neuromelanin MRI metrics differ between DaTSCAN positive and DaTSCAN negative participants with uncertain parkinsonism
    b. Explore whether neuromelanin MRI metrics in participants with uncertain parkinsonism differ from those in established PD and healthy controls (objective 4)

The hypotheses postulated were:

1. Automated template-based is a more consistent and reliable method of processing than semiautomated method.
2. An optimised sequence will generate with higher contrast to background noise that will perform better in terms of reliability.
3. NM changes with age in a non-linear fashion with a preferential decrease from 50 years of age.
4. Meta-analysis and case-control study may consistently show significant loss of NM-concentrated region in Parkinson and parkinsonism in comparison to healthy subjects, with high diagnostic accuracy.
5. NM-MRI would be able to detect reduced SN intensity in degenerative compared to non-degenerative parkinsonism in clinically uncertain parkinsonism as defined by positive and negative DaTSCAN.

3. General Methodologies

3.1 Introduction

All experimental chapters within this thesis were developed and processed through common methodologies. This chapter describes the general methodology involved, including participant recruitment, MRI acquisition protocols, and SN measurement methods. Additional methods or modifications specific to any experiment have been incorporated into the methods section of the said experiment.

3.2 Volunteer selection process

The recruitment process involved the selection of eligible and quality volunteers based on certain criteria. Subjects were generally included if they:

1. Participated voluntarily and were able to give informed consent.
2. Had good command of written and spoken English to complete the cognitive (mini-mental state examination) and depression (Beck's depression inventory) tests.
3. Had no contraindication to MRI, such as claustrophobia or metal implant.
4. Were not pregnant or potentially pregnant.

Inclusion criteria for healthy participants were as follows:

1. Good health with no significant past medical history
2. No diagnosis of any major neurological, neurosurgical, or psychiatric condition; or
3. Not currently on medications that could have dopaminergic effects.

Parkinson's patients were recruited from a large study, Parkinson MR Imaging Repository (PaMIR) fulfilling the following criteria:

1. Diagnosed using the UK Brain Bank criteria for PD.
2. Disease severity was recorded using the UPDRS (Rutledge et al., 1987) and Hoehn and Yahr clinical rating scales for PD (Hoehn & Yahr, 1967).

3.3 Image acquisition

All volunteers were scanned using two NM-weighted protocols performed by utilising an MRI facility available at the Sir Peter Mansfield Imaging Centre at the University of Nottingham. The Scanner used was a GE (Discovery MR750; GE Healthcare, Milwaukee, WI) with 3T magnetic field strength. Scanning was performed by the scanner operator in charge, and in some cases, by the second supervisor, Dr Yue Xing. Subsequent sections elaborate the image acquisition methods.

3.3.1 Preparation before the scanning

All volunteers underwent an MRI safety check prior to scanning to ensure no MR contraindication. Informed consent was also obtained before the scan. Patients' neurological examinations were conducted by Dr Saadnah Naidhu, a research colleague and clinical research fellow in neurology with two years of experience in movement disorders. This procedure yielded clinical data to confirm the absence of underlying neurological conditions among healthy control as described in the 'Volunteer selection process' subsection above.

Participants in the scanner room were affixed to a 32-channel head coil stabilised with an inflated foam padding to prevent motion artefact. Conventional sagittal T1 images were first obtained to set up the scanning field, followed by the NM-MRI sequence of either one of two sequences, or both sequences as detailed below.

3.3.2 Imaging protocols

Two NM MR imaging sequences were involved in the imaging protocols. Sequence 1 was developed later than Sequence 2 in order to reduce the acquisition time and slice gap. These procedures were performed considering the comfort of PD patients and the hectic schedule of clinical practice, without jeopardising the image quality.

3.3.2.1 NMMR sequence 1

The following parameters were used to acquire a 3-dimensional spoiled gradient recalled with magnetisation transfer (MT):
TR, 38.4 milliseconds; TE, 3 milliseconds; flip angle, 20; slice thickness, 2 mm; 0.375 x 0.375-pixel resolution; no gap; Field of view (FOV), 19.2; matrix: 480 × 192; Scanning time 3.25 minutes. Meanwhile, 30 axial image slices were obtained for each individual, with parallel alignment to the anterior callosum-posterior callosum line by covering no less than the entire midbrain.

3.3.2.2 NMMR sequence 2

T1 weighted spin-echo sequence with additional “off-resonance” MT pulse TR, 600 milliseconds; TE, 10 milliseconds; slice thickness, 2.5 mm; 0.38x0.38-pixel resolution; 0.3 mm gap; 12 slice number; acquisition time, 9 minutes 32 seconds with the field of view (FOV) aligned parallel to the anterior callosum-posterior callosum line by covering no less than the entire midbrain like sequence 1.

3.4 Image processing

3.4.1 Image quality control

Raw data were transferred and converted into compatible NIFTI format upon each scan to allow for software processing. After conversion, visual quality control assessments of data and images were performed to identify extrinsic and intrinsic artefacts that might obscure the signal intensity of the region of interest. This step was especially important as motor symptoms may manifest during scanning, affecting the image acquired. The author of this thesis conducted the quality control during his first experience in NM-MRI and double-checked with the second supervisor, Dr Yue Xing. Subsequently, all the image quality control was performed by the PhD candidate. The examples of intrinsic (i.e., from the MR equipment) and extrinsic artefact (i.e., from the subject) are presented in Figure 3. Both artefacts prevented the inclusion of data for further processing.

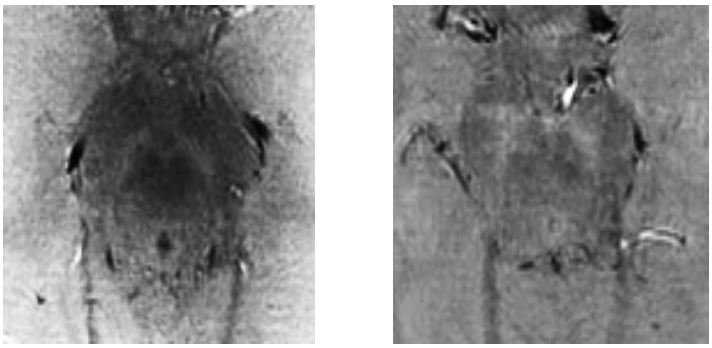


Figure 3. Sample of images that would be excluded during the quality control process. On the left is an image with low contrast on the area of the brainstem arising from the MR acquisition process and on the right is an image with motion blurring artefact.

3.4.2 Image Processing methods

Various processing methods are available to measure the signal intensity of the neuromelanin-rich region. This research aimed to analyse SN relative intensity by measuring the contrast ratio or contrast-to-background ratio (CBR) (both terms are used interchangeably throughout the thesis). CBR compares the desired signal levels to the background selected at the anatomy with a saturated signal near the region of interest, SNpc. Meanwhile, another quantification method, the contrast-to-noise ratio (CNR) quantifies image quality based on signal noise taken outside the brain.

The background of the anatomy - two sides of cerebral cri) with saturated signal near ROI was selected in this study. This is supported by a recent reliability study which found that cerebral cri was a more stable background region than red nucleus (van der Pluijm et al., 2021). The measurement is the relative quantification of the SN signal intensity to the saturated background signal (BG) resulting from the MT sequence employed in the anatomy of interest, neuromelanin. This image processing method was

chosen as described in our previous published work (Schwarz et al., 2011; 2017, Xing et al., 2018) and most of the related research in the field. For instance, earlier researchers quantified contrast ratio by employing the CBR method by analysing background from various anatomical areas, such as the cerebri cri (Isaias et al., 2016), superior cerebellar peduncle (SCP) (Ohtsuka et al., 2013), and dorsal side of the midbrain (Matsuura et al., 2016).

The two measurement methods utilised in this study were semiautomated manual drawing and automated template-based. The automated method only became available at the late stage of writing this thesis. The automated method was developed by the research team led by Prof Dorothee Auer.

3.4.3 Processing method 1: Semiautomated manual drawing method

The effects of the cluster of voxels detected from the MRI were investigated using ROI analysis. The process involved selecting the voxels via commercially available software, either by setting space with a fixed size (typically a sphere) or from an anatomical region.

The semiautomated method of ROI analysis employed in this study had been used since the beginning of the researcher's postgraduate studies as published by the research team in a recent article (Xing et al., 2018). Region of interest (ROI) placement and signal intensity distribution were determined through the Matlab software (MATLAB software version 2018a) scripted by Dr Yue Xing. A semi-automatic method was applied to measure the contrast-to-background ratio on three adjacent slices covering the SNpc.

High-intensity SNpc areas were selected using a fixed-size circular ROI while the low-intensity background was manually drawn at the low-intensity area of crus cerebri. This was first performed by using MRIcron software to display the participants' MRI images. Three slices of high signal intensity covering the SNpc, beginning from the first slice with the red nuclei as landmarks were then identified and recorded in a Microsoft Excel 2016 spreadsheet (Microsoft, Redmond, Wash).

For the SNpc signal(I_{SN}), four circular 3 mm diameter ROI placed at bilateral areas were identified as dorsal and ventral (red-filled circles in Figure 4). The background (I_{BG}) was manually drawn, covering areas of evenly low intensity at both sides of the cerebral cri (yellow coloured drawing in Figure 4).

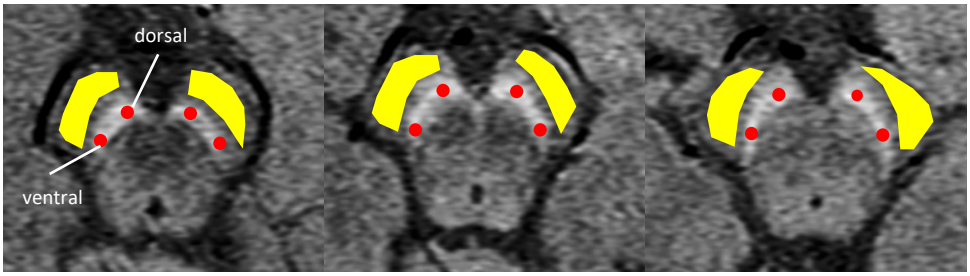


Figure 4. Fixed-size circular (red filled circle) region of interest at the SN areas (dorsal and ventral) and manually drawn backgrounds at both sides of cerebral cri (yellow coloured drawing).

Mean intensity of both sides (left and right) of the dorsal and ventral as well as the backgrounds, the red- and yellow-coloured ROIs on all slices were collected to compute CBR. The contrast ratio was calculated using the following formula 1: $CBR = \frac{I_{SN} - I_{BG}}{I_{BG}}$, where I_{SN} = intensity at SNpc areas and I_{BG} = intensity at background areas.

3.4.4 Processing method 2: Automated template-based method

The second image processing method employed was automated template-based. In summary, this technique utilised a standardised template for mapping to retrieve SN signal intensity and background intensity. It automatically retrieved SN intensity from coordinates of a template mapping to the individual image. Image processing was performed using FSL (FMRIB Software Library) and MATLAB version 2018a software. The process involved 4 main steps: affine registration, transformation, Bayesian classification and SN contrast estimation. The overall process is simplified as per the flowchart in Figure 5 adapted from Dr Yue Xing.

3.4.4.1 Affine registration

Analysis began with the registration process of each individuals involving two of the following affine registration steps:

i) Registration of NM-weighted images to T1 weighted images

ii) Registration of T1 weighted images to standard space

The first step was to linearly register the individual NM-weighted images (NMWI) to their respective T1 weighted images (T1WI) and registration of the T1WI to a standard space created by Dr Xing (second supervisor) as guided by Prof Dorothee Auer, main supervisor and advised by Dr Christopher Tench (senior statistician with 20 years' experience in MRI image analysis).

The standard space or template was created from 27 images of healthy controls selected from several sites, including London, Manchester, Glasgow, Newcastle, and Nottingham, whom the supervisors of this thesis have collaborated in a research project. Their NMWI is affinely registered to a standard brainstem space (ICBM152 space, http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009). This produced a brainstem template tailored to age-specific healthy control, containing information on respective spatial prior probability maps with three tissue classes: SN, brainstem outside of the SN, and irrelevant background.

Image registration proceeded with the application of the 'flirt' command in FSL with settings, to specify an angular range for initial optimisation search stage, 'searchx', 'searchy', 'searchz'. 'Degrees of Freedom (DOF)' were chosen from 6,7,9 or 12 for full 3D registrations. Visual checking is important following registration steps, to ensure images are perfectly aligned together.

3.4.4.2 Transformation

Output from the first registration steps was transformed into the standard space. A similar 'flirt' command was operated, with the standard space selected as a reference volume and the 'applyxfm' mode applied to save transformations to the input image. Manual visual checking was also performed in this step to ensure the output matched the standard space.

3.4.4.3 Bayesian classification and SN contrast estimation

Based on prior probability map of three tissue classes, Bayesian approach was applied in both image intensity and in space to obtain respective individual spatial posterior probability. The intensity of the three tissues for each transformed image was obtained by positioning a spherical ROI (kernel = 3) at the coordinates of the bilateral substantia nigra based on the ICBM152 template space (ICBM152 space, http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009):

i) ventral SNpc: right (107 115 60) and left (89 115 60),

ii) dorsal SNpc: right (107 112 57) and left (89 112 57).

The CBR was calculated using a similar formula as specified in the first method. The process was simplified in Figure 5 below.

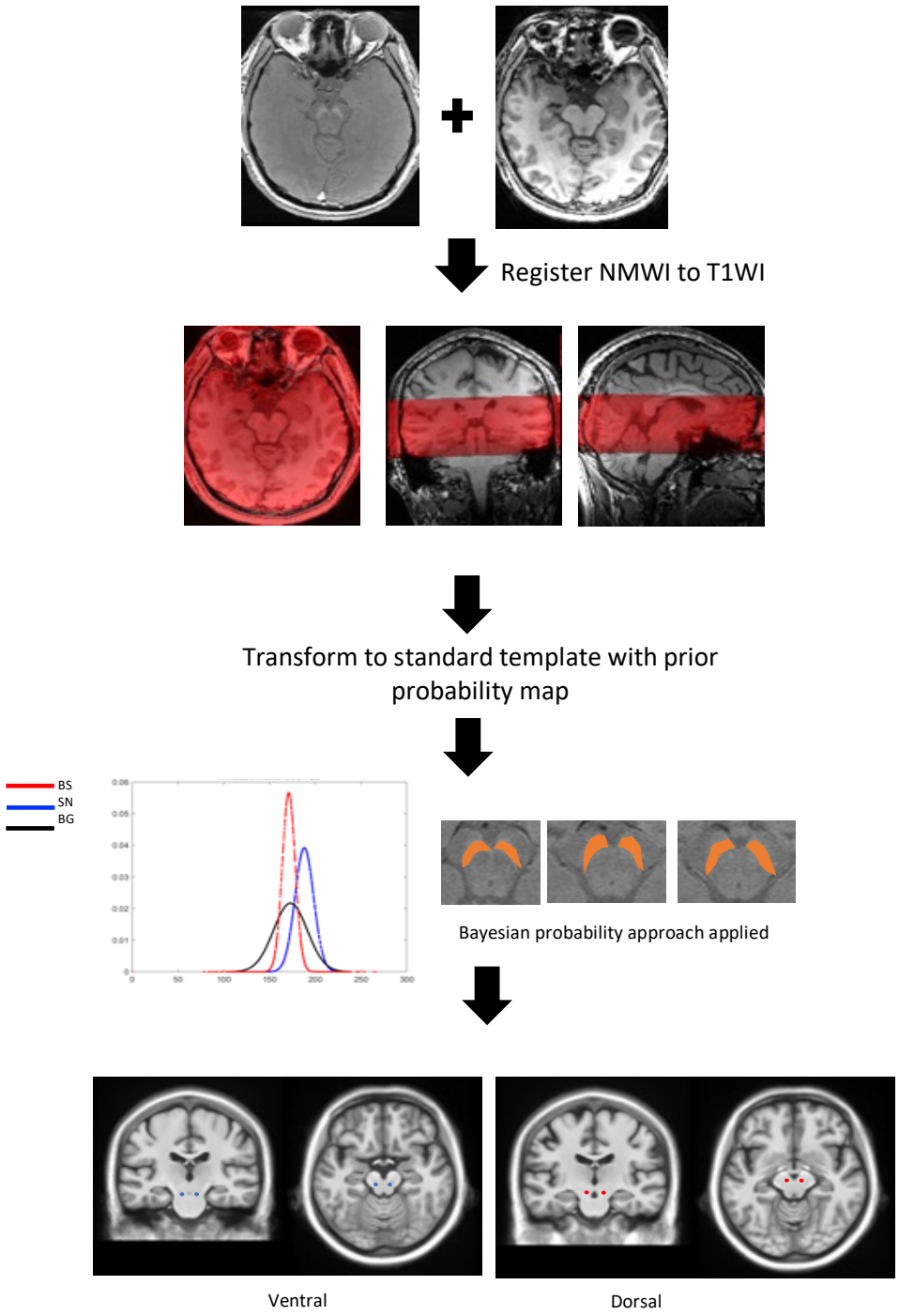


Figure 5. Flowchart of SN automated template-based image processing step. T1WI MRI images with red band overlay in 3 planes displayed linear registration of the NM-MRI to the T1WI. The orange areas are areas of SN posterior probability on three slices as a result of Bayesian approach. Bottom images shown the positions of regions of interest where the mean intensity of the SN was extracted based on the ICBM152 template space: bilateral dorsal SNpc (red dots) and ventral SNpc (blue dots).

4. Optimisation and Test-retest reproducibility of NM-MRI

4.1 Introduction

No standardised criteria are currently required for biomarkers to be accepted for use in clinical practice. Nevertheless, consistency and accuracy are among the most important criteria of a biomarker, especially when measuring certain quantities. Test-retest reliability examines the consistency, reproducibility, and agreement of one tool on two measurements towards a single individual under the same condition when no change of outcome is expected. The measurement tool is assessed based on whether it can produce an accurate representation of the subject. Hence, the outcome could be utilised for clinical purposes like diagnosis or differentiating between groups. It is also to ensure that change in observations obtained from using the tool across occasions is attributed to the individual itself instead of tool instability. Overall, test-retest reliability is crucial in determining instrument quality.

The growing interest in NM imaging as a tool for PD research and clinical application has not only seen major improvements in image processing techniques but also technical differences and sequence development. Accumulated evidence has depicted NM-MRI as a promising biomarker for PD. Despite the high diagnostic value of NM-MRI demonstrated in several studies, it is still important to assess the reliability of this tool prior to clinical implementation. High test-retest reliability is pertinent to ensure that measurements from the MRI can be replicated multiple times for the same individual. This is particularly important as consistent outcomes need to be produced each time a scan is taken. Hence,

upon clinical implementation, it can be confirmed that the outcome is attributed to disease-related changes and not from patient or MRI-modality related intrinsic or extrinsic factors.

Test-retest reliability analysis is performed by repeatedly scanning the same individual twice over a short period. Langley et al. (2017) were among the first researchers to perform test-retest reliability of NM-MRI. The first experiment in this thesis assessed the test-retest reliability of bespoke NM-MRI processing techniques and sequences since the accuracy of measurement to quantify pigmented dopaminergic neurons is still unknown.

Utilising the Bland-Altman approach for assessing reliability offers in-depth examination of data across samples, therefore facilitating evaluation of reliability. Bland-Altman Limits of Agreement (LOA) (Martin Bland & Altman, 1986) was introduced as a statistical method to assess agreement based on descriptive statistics (average and differences between repeated data) accompanied by a plot to facilitate data inspection and interpretation. Observed measurement should lie within the minimum-maximum range, thus indicating tool variability and whether it can be widely implemented. The reliability increases as the bias get closer to zero and the difference becomes smaller. The Bland-Altman plot could also be used to identify the distribution of data and outliers.

The first sub-study of this chapter aims to investigate the reliability of the semi-automated manual drawing method and the newly developed automated template-based method of image processing. This study also attempts to establish a dataset that could be utilised in optimising the

reliability by comparing various dedicated image processing techniques. This includes the one yielding the lowest variance between test-retest given no biological effect is expected in the short-term or medium-term.

The second part of this chapter aims to evaluate NM-sensitive MRI sequences, Sequence 1 and a previously developed sequence, Sequence 2 (mentioned in Chapter 3) at the Sir Peter Mansfield Imaging Centre, University of Nottingham. Presently, no agreement has been reached regarding the best and standard sequence.

4.2 Test-retest Reliability of two image processing methods

4.2.1 Materials and Methods

4.2.1.1 Participants

The study comprised six healthy participants, four males, and two females aged between 20 and 50 years old (mean age: 31.3 $\pm$ 5.5 years). None of the participants had a history of smoking or was presently a cigarette smoker or recreational drug user. Ethical approval for this study was granted by the local Ethics Committee and the Research and Development department (reference number: B12012012a-051, refer to Appendix 1 for details of the project). This study entailed a prospective design coordinated by the researcher, including the recruitment of volunteers, data acquisition, and image and statistical analysis. Both supervisors involved in the study design while the second supervisor, Dr Yue Xing also involved in data acquisition.

4.2.1.2 Acquisition

All participants underwent a short term rescan and a medium-term rescan. Short-term rescan refers to two repeated NM-MRI scans conducted in a single session with repositioning. Meanwhile, the medium-term or 2-weeks interval rescan refers to a separate MRI session conducted after the first scan, ranging approximately from 4 to 17 days later (mean±SD: 10.7 ± 5.1 days). The flow of the scan for an individual volunteer is presented in Figure 6, whereas Figure 7 displays three slices of the SN produced from the three scanning sessions conducted. One male participant was unable to perform a short-term repeat scanning due to time constraints. All participants were scanned by Dr Yue Xing. Images were acquired using the recently developed NM-MRI protocol (Sequence 1) described in Chapter 3.

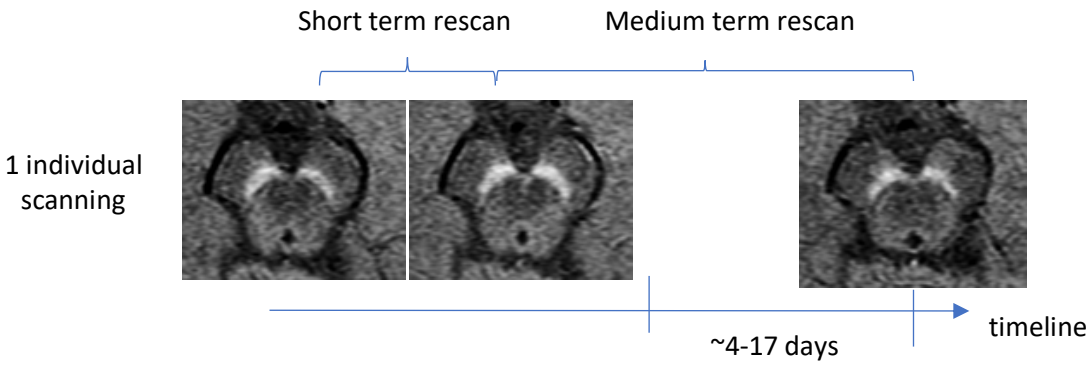


Figure 6. Flowchart timeline of a participant for scan-rescan of NMMR imaging.

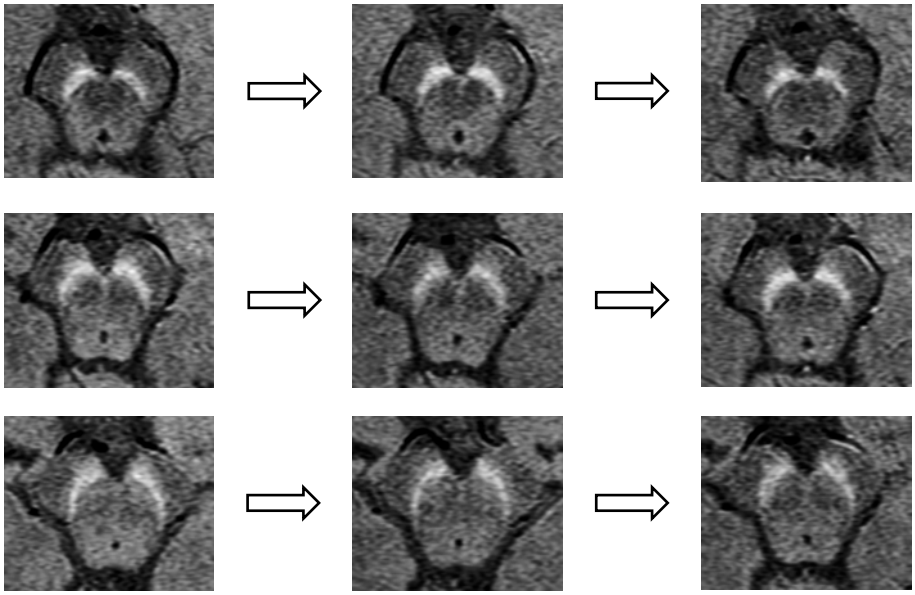


Figure 7. SNpc high signal intensity areas of a 35-year-old participant for the first scan, short-term rescan and medium-term repeat scan in 3 slices.

4.2.1.3 Image processing and analysis

Analysis conducted aimed to assess CBR of SNpc subregions via two processing methods; semiautomated manual drawing and automated template-based method as described in Chapter 3. Four different aspects of reliability were examined; intrarater reliability, interrater reliability, short-term rescan reliability, and medium-term rescan reliability. For intrarater reliability, the researcher as the first rater measured the SNpc from the first scan twice for each volunteer. Meanwhile, the interrater reliability involved a second-rater employing similar step of both methods to blindly measure the SN intensity of the same first scan upon intrarater measurement. The second rater had seven years of experience in MRI, which included four years in brain imaging research.

The reliability of the SNpc contrast-to-background ratio was assessed through Bland-Altman Limits of Agreement (LOA). Bland-Altman LOA was used to examine the agreement between two tests by plotting score differences against the mean score for each pair of measurements. The following formula was used to calculate the LOA:

$$\text{Upper LOA} = \text{bias} + 1.96SD$$

$$\text{Lower LOA} = \text{bias} - 1.96SD$$

Bias refers to the mean difference in measurements while SD refers to standard deviation of differences between each pair of measurement. The calculations were then visualised in a scatter plot, where the y—axis represents difference between a pair of tests while the x-axis represents the mean of two tests.

Bias refers to the mean difference in measurements while SD refers to the standard deviation of differences between each pair of measurements. The calculations were then visualised in a scatter plot, where the y—axis represents the difference between a pair of tests while the x-axis represents the mean of two tests.

The bias, as well as the upper and lower boundaries LOA of automated and semiautomated processing techniques in dorsal and ventral SNpc, were plotted according to intrarater, interrater, short-term rescan, and medium-term rescan. All figures were plotted using IBM SPSS Statistical software version 24 (IBM Corporation, Somers, NY, USA).

Since there was no established clinical standard for using this technique and a lack of reference study that employed BA plot in measuring test-retest

reliability, the researcher considered the threshold for LOA range as 0.1. A bias denoted a high degree of agreement when the value is close to zero in addition to a 95% distribution of data lying within the upper and lower boundaries of LOA and within the threshold range. On the other hand, there is a problem in agreement when the LOA range is either further away from the threshold of 0.1 or more than 5% outlier data (i.e., lying outside of the LOA boundaries).

4.2.2 Results

4.2.2.1 Intrarater reliability

The Bland-Altman analysis for repeat measurements of SNpc is summarised in Table 2 while the plots are shown in Figure 8. The BA plots demonstrated high reliability and agreement of both methods as the LOA range was within the threshold and no outliers were observed beyond the upper and lower limit of agreement.

Table 2. Bland-Altman limits of agreement (LOA) for intrarater reliability of the SNpc subregions contrast-to-background ratio (CBR) measurement using automated template-based method and semi-automated manual drawing method.

Measurement/ Subregion	**Automated template-based method**			**Semiautomated drawing method**		
	Bias	Upper LOA	Lower LOA	Bias	Upper LOA	Lower LOA
Dorsal SNpc	0	0	0	0.004	0.0160	-0.007
Ventral SNpc	0	0	0	-0.008	0.006	-0.022

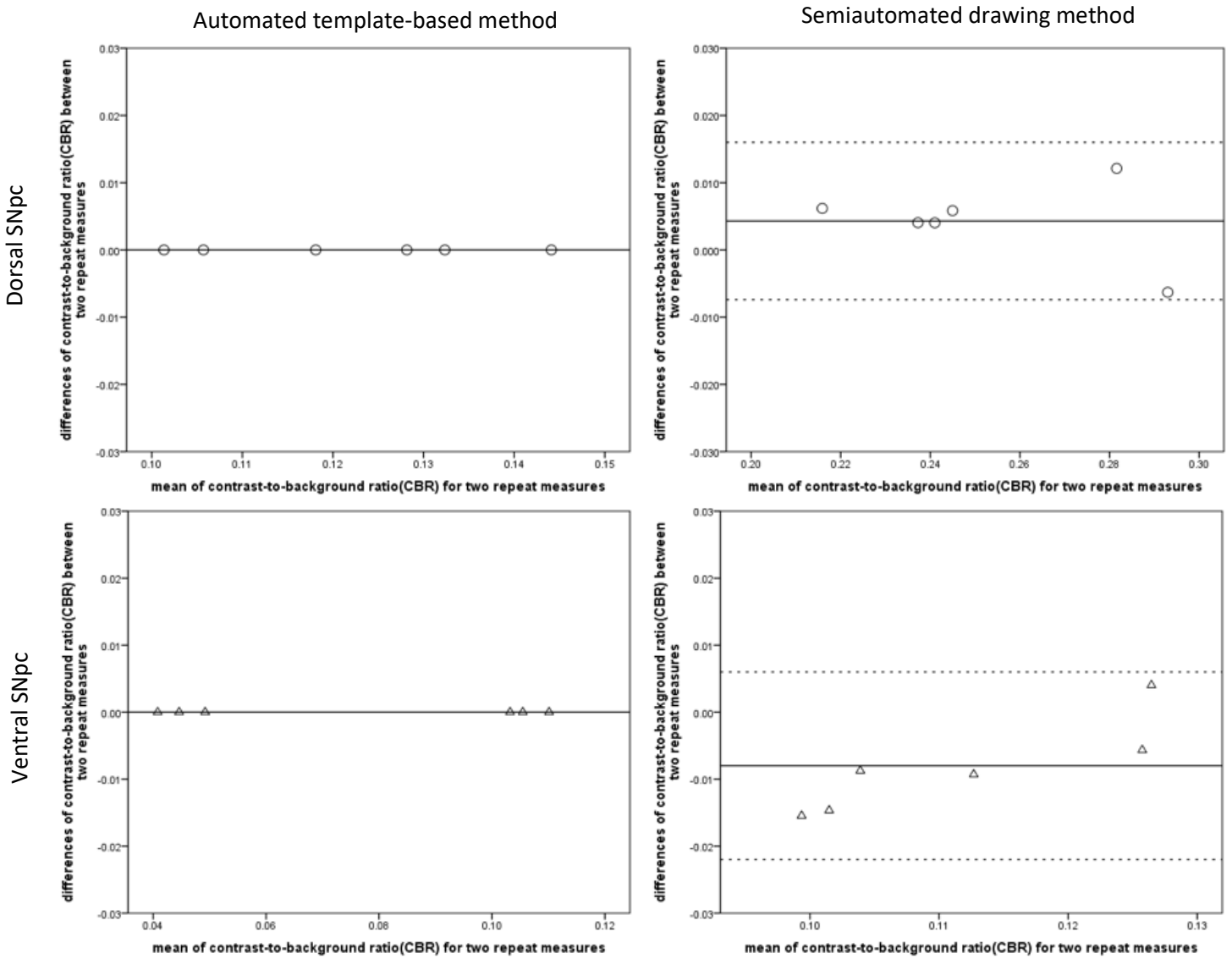


Figure 8. Bland-Altman plot for intrarater reliability of automated template-based and semiautomated drawing methods to show that both sequences reliably measure SNpc with almost all the data lying within the limit of agreement with minimal outliers (data from Table 2).

4.2.2.2 Interrater reliability

The Bland-Altman from the first and second-rater are presented in Table 3. The BA plot (Figure 9) illustrated no outliers outside upper and lower LOA for intrarater reliability and the LOA boundaries did not exaggerate the threshold, indicating high reliability and agreement for both semi-automatic and automatic methods. As the automated method was not fully automated (i.e., need for visual checking and parameter change), a non-zero bias was expected.

Table 3. Bland-Altman limits of agreement (LOA) for interrater reliability of the SNpc subregions contrast-to-background ratio (CBR) measurement using an automated template-based method and semi-automated manual drawing method.

Measurement/ Subregion	**Automatic**			**Semiautomatic**		
	Bias	Upper LOA	Lower LOA	Bias	Upper LOA	Lower LOA
Dorsal SNpc	-0.001	0.002	-0.004	-0.017	-0.005	-0.030
Ventral SNpc	0.000	0.004	-0.003	-0.032	0.016	-0.080

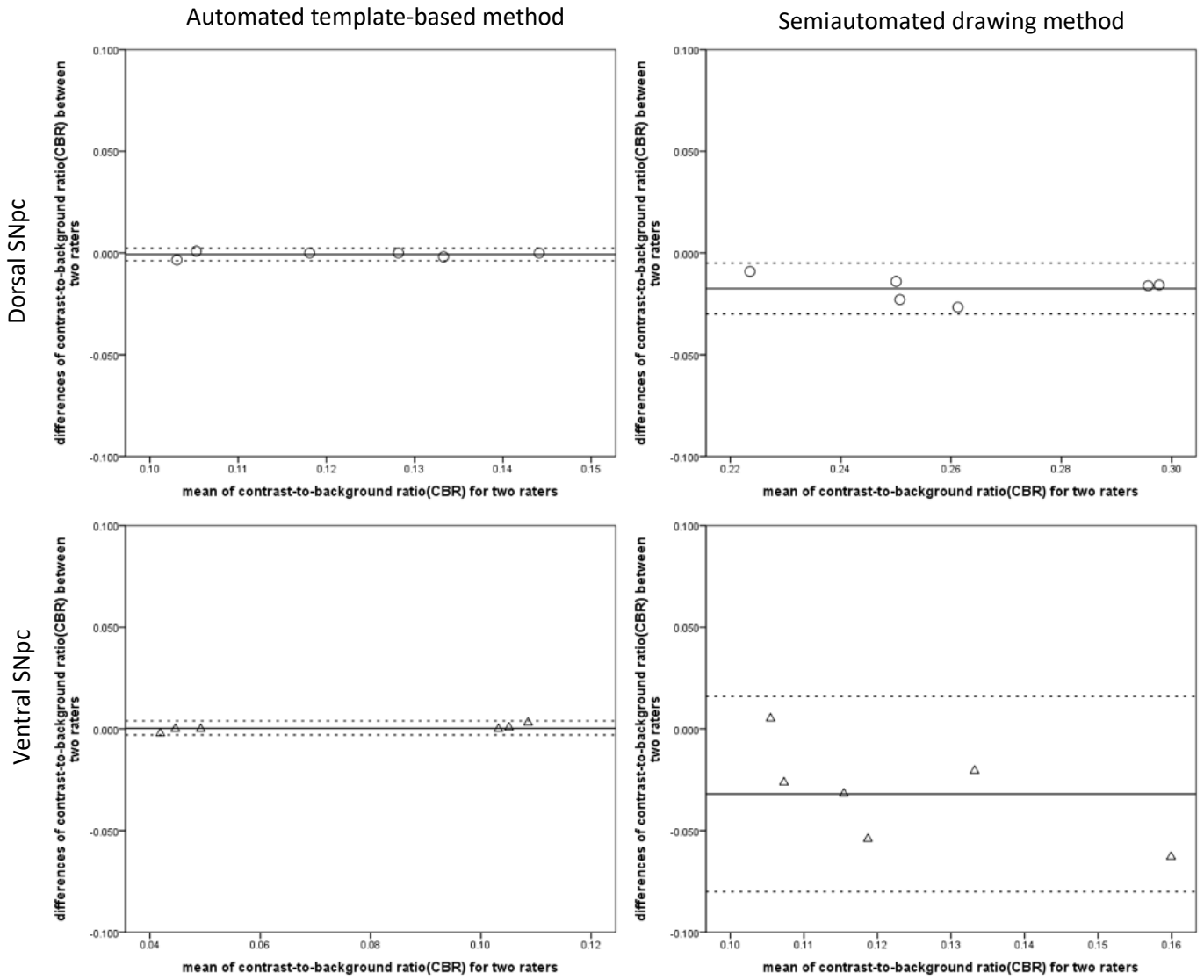


Figure 9 Bland-Altman plot for interrater reliability of automated template-based and semiautomated drawing methods to show that both sequences reliably measure SNpc with almost all the data lying within the limit of agreement with minimal outliers (data from Table 3).

4.2.2.3 Short-term rescan reliability

For short-term rescans, data points fell within the lower and upper LOA (see Figure 10) plus the upper and lower LOA range were below the limit threshold, enhancing reliability and agreement of the template-based and manually drawn methods. Bias measurement on the manually drawn method showed a higher mean difference than the template-based method, therefore pointing to lower consistency and agreement between scanning. The Bland-Altman of short term rescan reliability is as presented in Table 4.

Table 4. Bland-Altman limits of agreement (LOA) for short term reliability of the SNpc subregions contrast-to-background ratio (CBR) measurement using an automated template-based method and semi-automated manual drawing method

Measurement/ Subregion	**Automatic**			**Semiautomatic**		
	Bias	Upper LOA	Lower LOA	Bias	Upper LOA	Lower LOA
Dorsal SNpc	0.001	0.008	-0.008	-0.001	0.019	-0.022
Ventral SNpc	0.001	0.007	-0.006	-0.006	0.018	-0.030

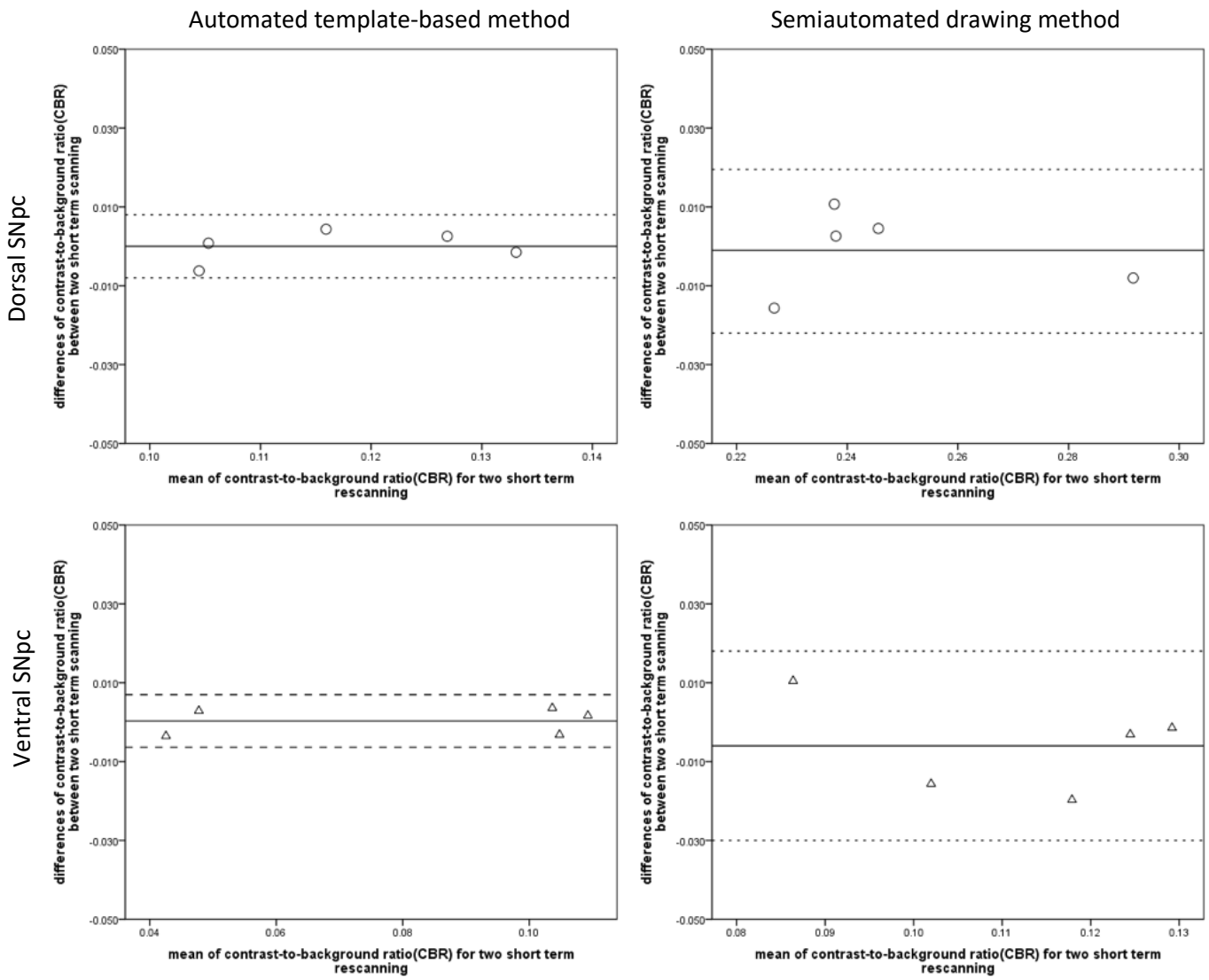


Figure 10. Bland-Altman plot for short term rescan reliability of automated template-based and semiautomated drawing methods to show that both sequences reliably measure SNpc with almost all the data lying within the limit of agreement with minimal outliers (data from Table 4).

4.2.2.4 Medium-term rescan reliability

The final aspect of reliability elucidated in this study was medium-term rescan reliability. Similar to the findings above, no data evaded the upper or lower boundaries of LOA (Figure 11) and the upper and lower LOA boundaries were lower than the threshold limit. Hence, both measurements demonstrated high reliability and agreement. The trend of bias for the automatic method was lesser compared to that for the semi-automatic method, which implies higher consistency and agreement between both methods. The Bland-Altman of medium-term rescanning is summarised in Table 5.

Table 5. Bland-Altman limits of agreement (LOA) for medium term rescan reliability of the SNpc subregions contrast-to-background ratio (CBR) measurement using automated template-based method and semi-automated manual drawing method.

Measurement/ Subregion	**Automatic**			**Semiautomatic**		
	Bias	Upper LOA	Lower LOA	Bias	Upper LOA	Lower LOA
Dorsal SNpc	-0.002	0.013	-0.017	0.009	0.031	-0.013
Ventral SNpc	-0.003	0.011	-0.017	-0.035	-0.012	-0.059

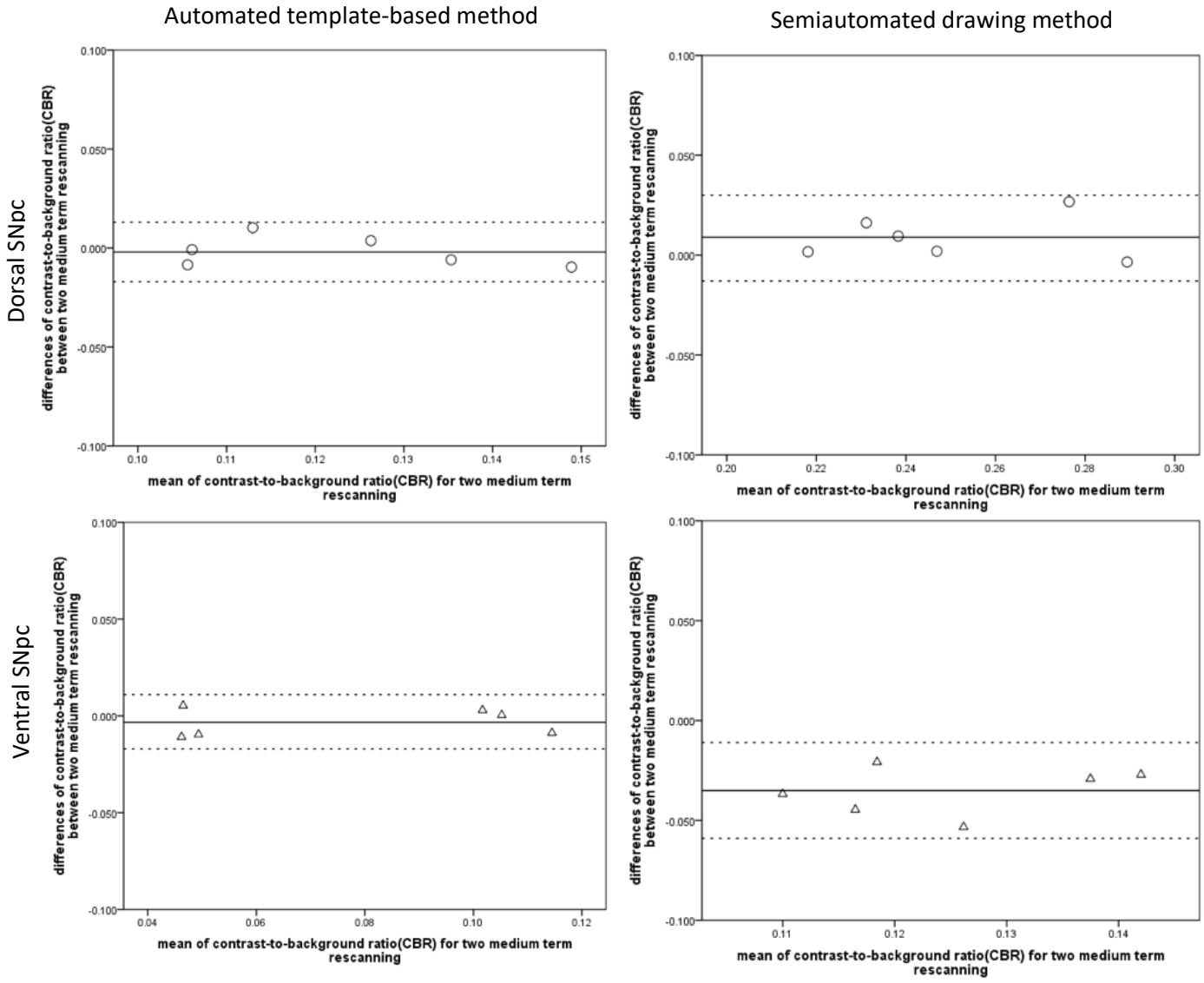


Figure 11. Bland-Altman plot for medium term rescan reliability of automated template-based and semiautomated drawing methods to show that both sequences reliably measure SNpc with almost all the data lying within the limit of agreement with minimal outliers (data from Table 5).

4.2.3 Discussion

This study involved six healthy subjects with images from three scanning sessions; the first scan, short-term rescan on the same day, and medium-term rescans within a range of 4 to 17 days later. This study attempts to assess the reliability of two image processing methods of NM-sensitive imaging; semi-automated drawing and automated template-based method. Both methods were compared through a series of tests to observe their reliability and agreement in quantifying SN intensity.

Results obtained revealed all four tests; intrarater reliability, interrater reliability, short term rescan, and medium-term rescan recorded high reliability and agreement of CBR in dorsal and ventral SNpc. All data distributions were within the upper and lower boundaries of LOA in the BA plot. None of the data distribution was beyond the threshold LOA of 0.1.

In the intrarater reliability, perfect agreement with zero bias of the automated method vs non-zero bias of the semiautomated method suggests an ideal condition. This was evident in the consistent image visual checking taken by the same rater during the registration and transformation steps hence similar choice of angular range and DOF were selected (refer to Chapter 3.4.4). Apart from the variation in the manual drawing of SN and BG in the semiautomated method, zero bias was also explained by the CBR measured in three decimal points. Repetitive CBR measurements revealed that the bias between them was very near to zero, hence the zero bias was taken. Notably, visual checking might explain the non-zero bias in interrater reliability of the automated approach as visual acuity variation from one rater

to another leading to varying angular range and DOF selection. However, near-zero bias with a narrow range of LOA could be achieved by utilising the manual drawing-based method, suggesting a minor to an insignificant effect of visual acuity subjectivity. The researcher's three-year experience in NM imaging analysis compared to the second-rater who was a first-timer in NM imaging might also contribute to the outcome, especially when identifying the SN subregions and the background space.

Short-term and medium-term rescans were conducted to investigate whether slight differences in patient positioning and selection of the field of view (the slice orientation) affected the measurement of NM signal intensity in response to normal body physiological change over a period. An increased variation in bias and wider LOA range of both short-term and medium-term rescans in comparison to intrarater reliability was observed. Meanwhile, bias and range of LOA were considerably similar to that of interrater reliability, suggesting that the differences did not significantly affect SN intensity measurement although with different physiological status and scanner performance. All measurements were highly reliable since data distribution did not exceed the upper and lower boundaries of LOA. Comparing the range of LOA between dorsal and ventral SNpc displayed a wider range in ventral SNpc than the dorsal SNpc. This might be due to the significance of the ventral SNpc in nigrostriatal neuronal vulnerability for PD (Fearnley & Lees, 1991; Langley et al., 2020).

The findings align with a previous short term rescans reliability study that scanned 11 young adults (mean: 28 years old) twice with repositioning

(Langley et al., 2017). Their SNpc volume and SNpc MT contrast measurements reflected substantial to excellent reproducibility. Although Langley et al. (2017) used a different testing method, their findings were similar to the present study as both reports revealed that NM-MRI is a highly reliable tool. Similarly for medium term rescan, an assessment on test-retest reliability study with a period of 3 weeks interval revealed that three NM-MRI sequences and three measurement methods are highly reproducible measure on test-retest as well as interrater and intrarater reliability (van der Pluijm et al., 2021).

An overall trend of lesser or near-zero bias and LOA of a narrower range was observed for the automatic method, suggesting its higher consistency compared to the semiautomatic method. Therefore, the automatic approach has an edge in comparison to the semiautomatic approach. Nevertheless, van der Pluijm et al., 2021 discovered that their manual method which is near to the semiautomated method of this study had substantial to almost perfect test-retest intraclass correlation coefficient (ICC) ranged from 0.79-0.84 (on three rater) while automated method scored only moderate test-retest ICC of 0.72 on CBR measurement. The differences could be attributed by many factors such as interindividual rater skill as they employed three rater as first timer and variation in the automated analysis method applied such as the template creation and pre-processing registration steps which this study used FSL software and Matlab while they used Statistical Parametric Mapping (SPM 12; Wellcome Trust UK) software.

The present dataset could also be used to assess the reliability of alternative current and/or future processing methods. Although the ventral SNpc appeared less reliable among the young healthy group in this study, such anomaly is predicted to be accentuated in older subjects, especially in PD patients. A multi-centre study of the same participants may be worthwhile for standardisation and harmonisation of the processing methods. This is a crucial consideration in the application of NM imaging in PD, which will be uncovered in chapter 6.1 of this thesis.

This study has important limitations. For instance, the participants were mostly young adults and healthy and no older adults or PD patients were involved. In older adults and PD patients, reliability may be reduced since neuromelanin loss will decrease CBR and symptom manifestation may cause artefacts. The next subchapter will explore the reliability of both measurement methods on normal volunteers and PD patients.

Secondly, disparity of experience in NM-MRI research between rater 1 (myself) with 3 years' experience and first time for rater 2 might affect reliability which supported by many MR reliability studies (Fu et al., 2014; van der Pluijm et al., 2021; Visser et al., 2019). Reliability is expected to increase with experience. This study also lacked true NM concentration, hence unable to further characterise the reliability of the information regarding the processing methods in terms of their true value. True value can only be gathered via post-mortem morphological study or synthetic NM phantom. During the course of this thesis, a synthetic NM phantom was attempted to

build with help of a chemist collaborator, Meike Katharina Leu, who has 3 years of experience. However, the work was not completed.

The current image alignment procedure is rater-dependent and involves visual checking. Nonetheless, artificial intelligence or machine learning approaches could make the procedure become hands-off, thereby ensuring a perfect alignment of images. Hence, image processing time could run faster with lesser involvement of human resources.

4.2.4 Conclusion

This experiment aimed to assess the reliability of automatic template-based and semi-automated manual drawing methods for measuring SN intensity. The methods were tested for reliability via Bland-Altman plot. High reliability was observed in both methods though the automatic template-based presented better consistency and more reliable compared to the semi-automated method. The high reliability observed in this study provides evidence for NM-MRI as a reliable non-invasive *in-vivo* tool in investigating the role of the dopaminergic system in PD.

This work is a vital step to assess the technical validity of establishing promising biomarkers for clinical application. The reliability of both measurement approaches among healthy and PD patients is further evaluated in the subsequent section. Further optimisation on higher resolution images with higher magnetic field strength MRIs such as 7T could be attempted in future studies. However, such optimisation is beyond the scope of this thesis and may require advanced scanner technologies.

4.3 Assessment of two NM-weighted MRI sequences

4.3.1 Materials and methods

This study was a single-centre sub-study of an ongoing collaborative multicenter research, PaMIR. Ethical approval was obtained from the local Research Ethics Committee (reference number: 14/EM/0061) (refer to Appendix 2 for study description). Data were collected from five MRI facilities across the UK, including our research facility, Sir Peter Mansfield Imaging Centre at the University of Nottingham. This work utilised and analysed datasets only from our MR facility, which began recruitment in February 2015. Participant recruitment was led by Dr Yue Xing and facilitated by Dr Saadnah Naidhu. The author of this thesis was involved in data organisation and safekeeping, fulfilling participants' logistical needs, preparation for data acquisition, and data analysis (i.e., image and statistical analysis). This study utilised available data from participants scanned with two locally developed NM-MRI sequences (Sequence 1 and Sequence 2) in one session without repositioning. Data from Sequence 1 were also used in the case-control study in Chapter 7 and the next chapter (Chapter 5) involving healthy subjects to elucidate age-related changes of neuromelanin pigment from adulthood to old age.

4.3.1.1 Participants

This study obtained data from 73 participants scanned via two NM imaging protocols. Upon image quality control (as described in Chapter 3.4.1), data exposed to severe motion artefact in either one or both

sequences were removed. The final analysis included a total of 69 participants, 35 healthy participants (17 males) and 34 PD patients (19 males) with age ranging from 49 to 83 years old and from 51 to 83 years old, respectively. The overall mean age ± SD for healthy and PD participants was 66.58 ± 9.21 and 71.96 ± 7.34, respectively. Demographic data are presented in Table 6 below.

Table 6 Demographics for healthy controls and Parkinson patients.

	Controls	PD
No. (female/male)	35 (17/17)	34 (15/19)
Age, years (mean±SD)	66.58 ± 9.21	71.96 ± 7.34
Age range (min-max)	48.52 - 83.08	50.98 - 83.38

4.3.1.2 Acquisition

Both groups were scanned by two NM-MRI sequences (refer to Chapter 3) in one session. Figure 12 displays the examples of NM-weighted images of the same individual from two sequences, Sequence 1 and Sequence 2. Notably, Sequence 1 has a slice thickness of 2mm and acquisition time of approximately 3 mins, whereas Sequence 2 has a slice thickness of 2.5 mm and acquisition time of approximately 9 mins.

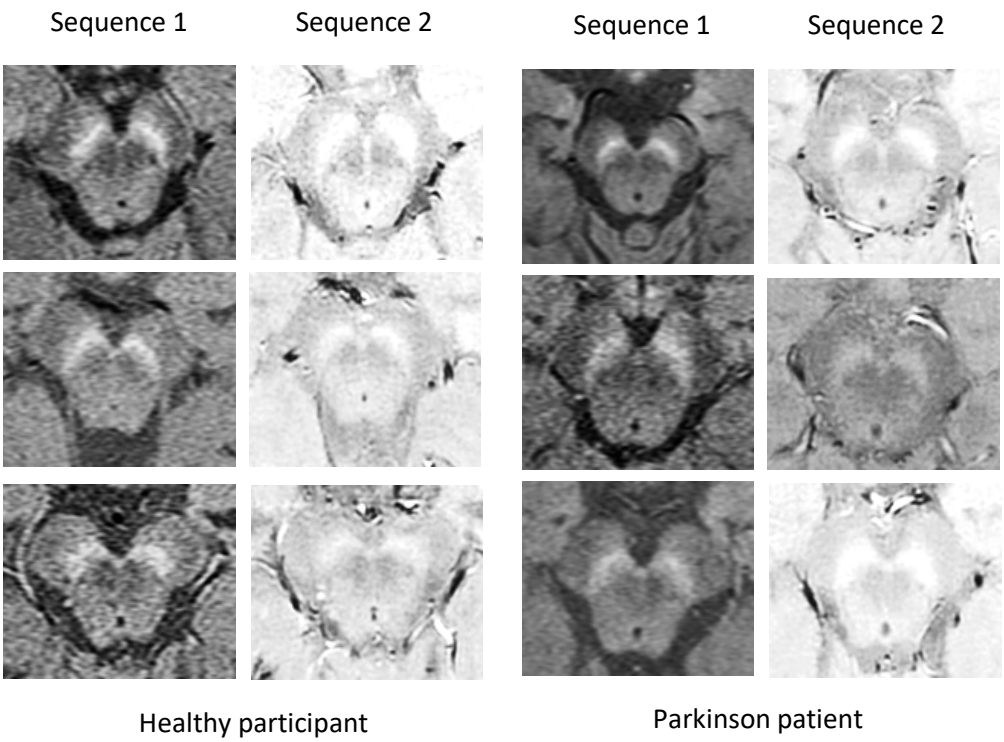


Figure 12. Example 3 slice-images of healthy and Parkinson patients being imaged with two neuromelanin-sensitive MR Imaging sequences.

4.3.1.3 Image processing and analysis

Image processing was performed similarly to the principles described in Chapter 3, the semiautomated manual drawing and automated template-based methods. A slight modification was made to the semiautomated technique for this study, where only one slice was selected based on the visually brightest SNpc intensity, instead of the usual three slices.

Scatter plot and Bland-Altman analyses were conducted to assess reliability and differences between the sequences. The LOA range with a high threshold was chosen as 0.2 since this study involved technical differences between two sequences in which SN contrast variation is expected. Besides, this study recruited both healthy aged and clinical populations (i.e., PD patients) in which degradation of SNpc may exist (further details in Chapter 5 and Chapter 7). Correlation analysis was also conducted to evaluate the strength of the relationship between both sequences, thereby providing information on SN measurements.

4.3.2 Results

4.3.2.1 Relationship between sequence 1 and 2

The relationship between the pair of sequences was observed through scatterplot and correlation analysis. Scatterplots for Sequences 1 and 2 were plotted together on the y-axis being CBR of Sequence 1, whereas the x-axis was CBR of Sequence 2. The plots were separated to enable the comparisons based on different processing techniques and participant groups (i.e., volunteers' group, healthy control, and PD patient).

The scatter plot of CBR (Figure 13) revealed that all data were distributed above the unity line in the middle. This finding suggests that Sequence 1 had higher SNpc CBR measurement compared to Sequence 2. Correlation analysis revealed a strong positive linear association between CBR of sequence 1 and CBR sequence 2 as shown in Table 7 below.

Table 7. Correlation analysis performed on image processing methods on healthy control and PD patients demonstrated that both sequences were highly in agreement with a strong relationship denoted by high a R^2 value

Groups	Automated method		Semiautomated method	
	R^2	P-value	R^2	P-value
Healthy control	0.865	0.0001	0.767	0.0001
Parkinson patient	0.908	0.0001	0.688	0.0001

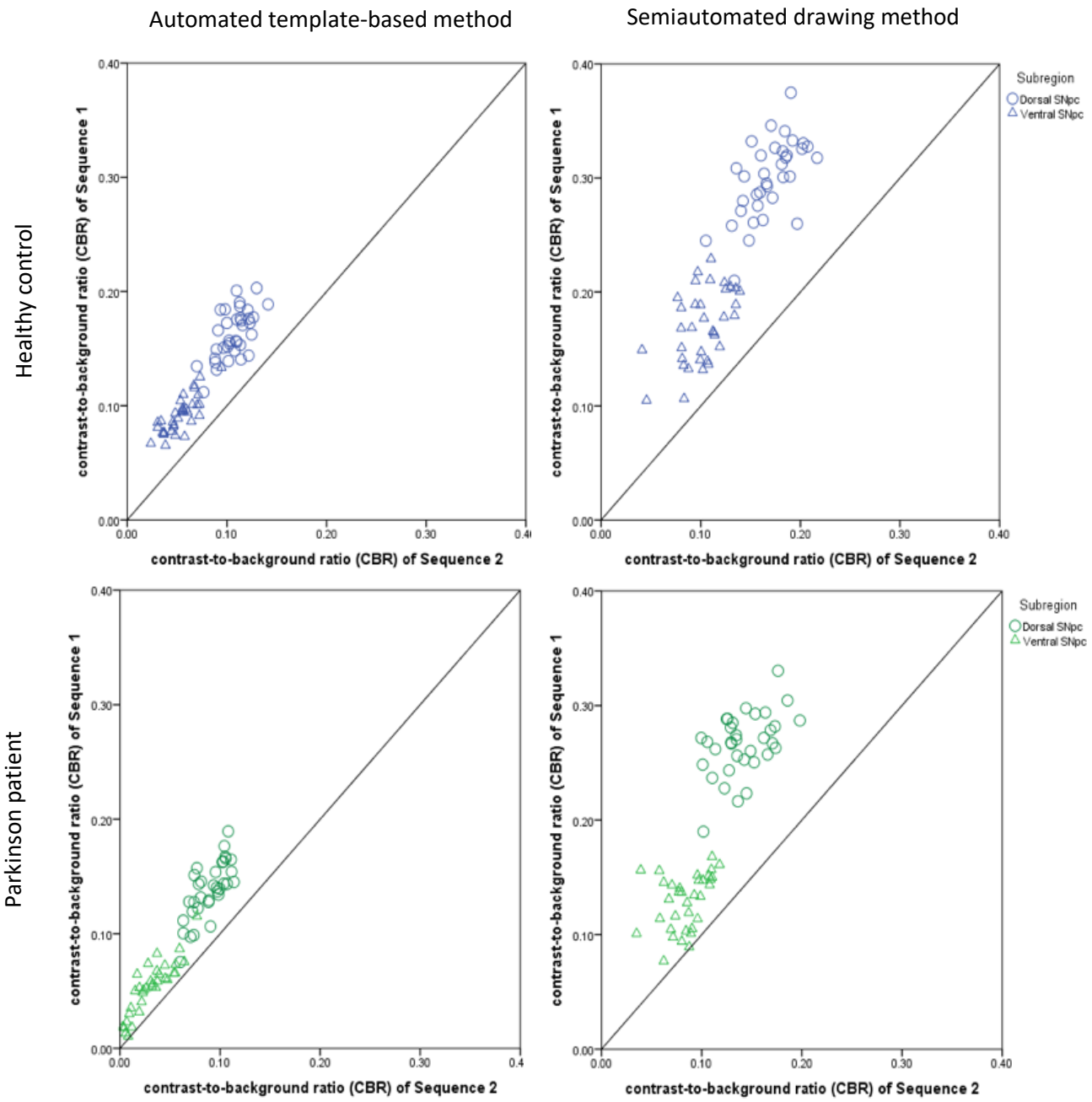


Figure 13. Scatter plots image processing methods on healthy control and Parkinson patients presented almost all data points reside above the unity line indicating SNpc CBR measured by sequence 1 were higher than sequence 2. The circle is for dorsal SNpc and the triangle is for ventral SNpc.

4.3.2.2 Agreements and differences between Sequence 1 and Sequence 2

Further analysis was performed by constructing a Bland Altman plot to investigate the agreement and differences between both sequences. The CBRs were separated by subgroup and processing techniques to investigate whether each method corresponded to different bias measures. Figure 14 (A and B) provides evidence that both sequences regardless of population group were highly reliable tools to measure SNpc, with the LOA range within the threshold while the majority of the data were found within the LOA and only a few outliers were observed (less than 5%). Table 8 depicts the bias, as well as upper and lower.

Table 8. Summary of bias and limit of agreements of automated template-based method and semiautomated manual drawing method.

	Automated method			**Semiautomated method**		
Healthy control	Bias	Upper LOA	Lower LOA	Bias	Upper LOA	Lower LOA
Dorsal SNpc	0.056	0.089	0.023	0.131	0.184	0.078
Ventral SNpc	0.039	0.058	0.020	0.069	0.125	0.013
Parkinson patient						
Dorsal SNpc	0.049	0.081	0.016	0.125	0.178	0.072
Ventral SNpc	0.023	0.044	0.001	0.045	0.094	-0.003

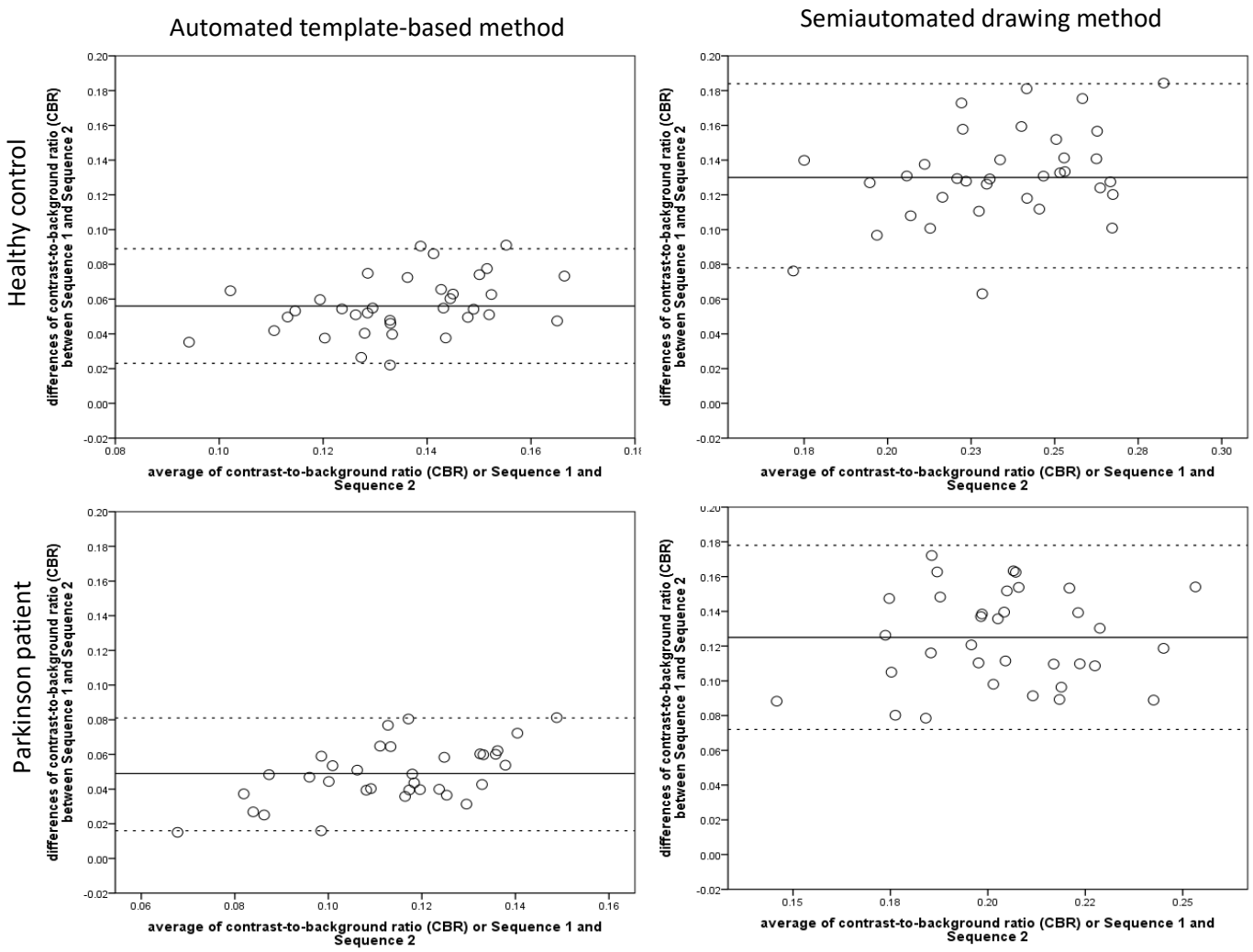


Figure 14A. Bland-Altman plot of automated template-based and semiautomated drawing methods for dorsal SNpc to show that both sequences reliably measure SNpc with almost all the data lying within the limit of agreement with minimal outliers.

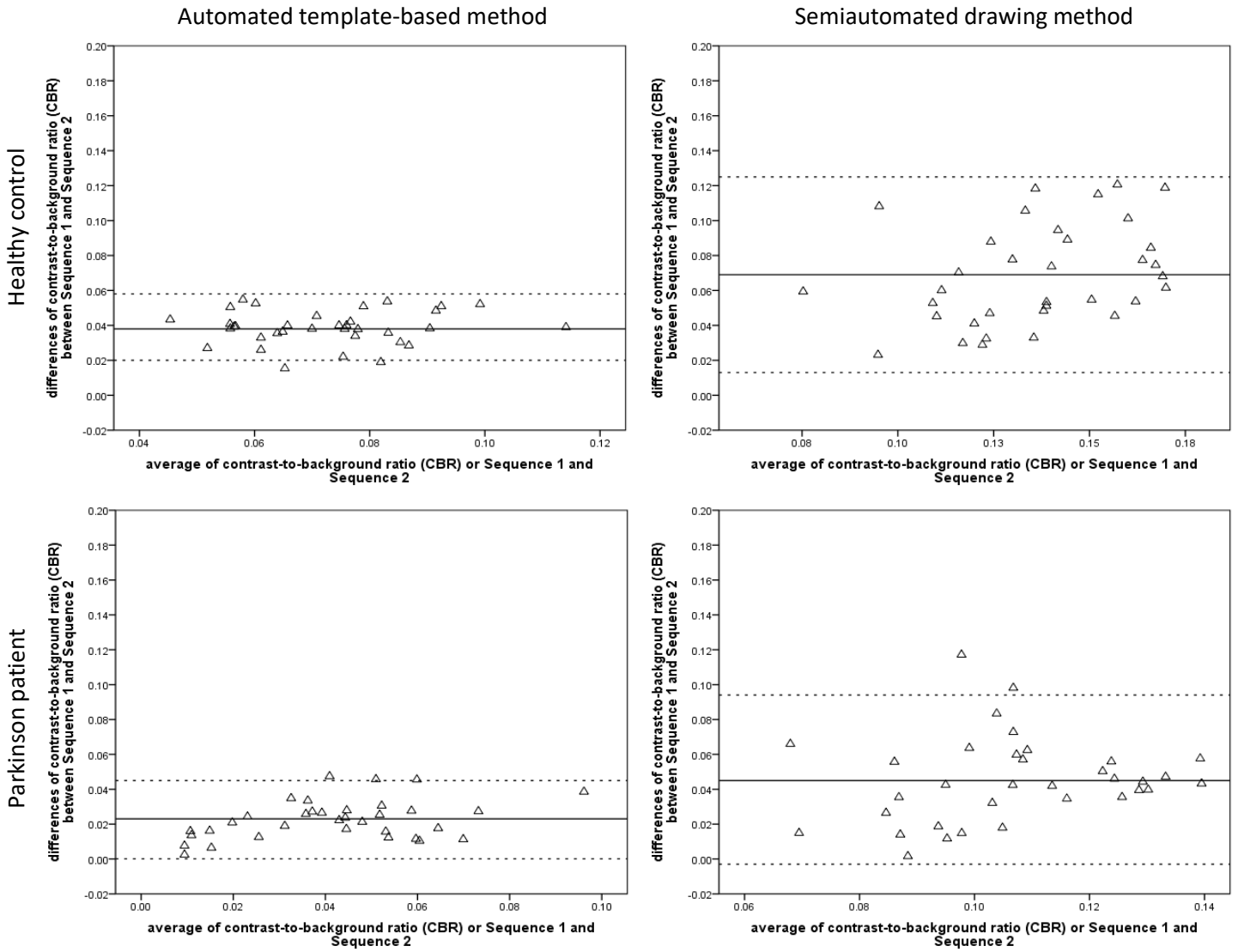


Figure 14B. Bland-Altman plot of automated template-based and semiautomated drawing methods for ventral SNpc to show that both sequences reliably measure SNpc with almost all the data lying within the limit of agreement with minimal outliers.

4.3.3 Discussion

This work aimed to evaluate the relationship between two NM-MRI sequences. Specifically, SN intensity was measured via automated template-based and single-sliced semiautomated methods to investigate reliability, agreement and differences that might affect the measurement of SNpc hyperintensity.

The relationship between both sequences was visualised using scatterplots with sequences 1 and 2 as the y-axis and x-axis, respectively. Scatterplots revealed all data points to be above the unity line, suggesting a higher SNpc CBR measure in Sequence 1 compared to Sequence 2 when scanned on the same individual. Intrinsic factors from the MRI machine and patient positioning affecting sequence signal and noise could be controlled since scanning was conducted within the same session. The NM signal contrast increased rather than the background noise, thus reflecting that Sequence 1 produced a higher contrast compared to Sequence 2.

Correlation analysis in the present study demonstrated a strong positive linear correlation between Sequence 1 and Sequence 2. The R^2 for both sequences was high and consistent with each other. Additionally, reliability analysis of the BA plot displayed high reliability and agreement of Sequence 1 and Sequence 2 in measuring the SNpc intensity regardless of population group. This finding is important as a study reported various acquisition and preprocessing parameters affect the strength and test-retest reliability of the signal intensity obtained from NM-MRI in determining

optimised protocols (Wengler et al., 2020). Further, a recent prospective study that assessed the test-retest reproducibility of three widely used NM-MRI sequences and three methods of analysis found that the NM-MRI gradient recalled echo (GRE) with MT pulse (close to Sequence 1) produced a higher contrast and higher reliability for contrast ratio measurements compared to NM-MRI with spin echo sequence with MT (close to Sequence 2) or without MT (van der Pluijm et al., 2021).

These finding also suggests that differences in diagnostic performance could be obtained from Sequence 1 compared to Sequence 2, aligning with a recent meta-analysis assessing the significance of various MRI parameters such as slice thickness in diagnosing PD (Cho, Bae, Kim, Kim, Baik, Sunwoo, Choi, Jung, et al., 2021). This is pertinent given the significance of the SNpc in PD etiopathogenesis.

Overall, these findings highlight that sequence 1 and 2 had high agreement despite technical differences. This chapter did not analyse NM-MRI capability in distinguishing healthy control from PD patients, nonetheless, Sequence 1 had a shorter acquisition protocol (about 3 minutes). Therefore, the protocol was more comfortable for participants and less prone to motion artefact. This is particularly useful when dealing with movement disorder patients. A case-control study was conducted and discussed later in Chapter 7.

A limitation of this study is that the standardisation method to normalise the measurement of different sequences was not investigated. A normalisation technique is applied in Chapter 7 in which different individuals'

data were scanned from either sequence 1 or sequence 2 were analysed together.

4.3.4 Conclusion

Healthy controls and PD patients were scanned via two neuromelanin-sensitive MRI sequences in a single session to investigate the correlation with SNpc hyperintensity. Both sequences demonstrated high consistency and agreement with strong positive correlations. Sequence 1 revealed superior CBRs compared to Sequence 2, hence suggesting higher contrast. Besides, sequence 1 had a shorter acquisition protocol that offers a better choice for PD patients since clinical symptoms may introduce artefacts.

5. Neuromelanin pigment changes from adulthood to old age

5.1 Introduction

Neuromelanin (NM) is considered neuroprotective given its capacity to act as a scavenger by removing potential toxins either via binding redox-active metal ions or the autooxidation of catecholamines (Sulzer & Zecca, 2000). However, NM may elicit neurodegeneration by inducing immune-based pathogenic mechanisms following their release from degenerating neurons (Zucca et al., 2014).

Accelerated ageing has been linked to the development of PD due to the heightened oxidation stress in the substantia nigra (SN) and environmental toxins, where NM plays a vital and catalytic role (Enochs et al., 1993). The SN and locus coeruleus (LC), being the major sites of NM in the brain, have been recognised to reflect reduced T1 hyperintensity in previous NM imaging studies (Kashihara et al., 2011b) . Meanwhile, positive associations were reported between the quantity of NM and symptom severity (Kashihara et al., 2011; Ohtsuka et al., 2013; Schwarz et al., 2017). An in-depth understanding of age-related physiological variations of SN pigmentation in vivo remains unclear.

Pigmented granules were found in human SN as early as 2 to 3 years with a linear effect as the accumulation increased with age (Cowen, 1986). Other histological studies demonstrated that the size and number of pigmented granules either remained stable from 20 years of age (Halliday et al., 2006), age-dependent reduction (Ma et al., 1999) or continued to increase (Zecca et al., 2001).

Post-mortem exploration of age-dependent NM change is still unclear and insufficient in associating the changes to the active human brain. *In-vivo* studies offer great potential for better understanding by quantifying NM in the active human brain. The advancement of imaging technology allowed for this thesis to explore age-related changes in NM-rich regions *in-vivo* using magnetic resonance imaging.

The first *in-vivo* study on NM-rich noradrenergic neurons of LC suggested a nonlinear age effect (Clewett et al., 2016; Shibata et al., 2006). This study, however, was interested in investigating NM signal changes in the SN from young adulthood using an optimised NM-MRI protocol. A similar work analysed using semiautomated approach has already been published in the journal Movement Disorder in 2018, where the author of this thesis served as the second author (Xing et al., 2018). This chapter aimed to revisit the relationship via an automated template-based method to assess the population of young adults till old age. Since the standard template created for automated analysis was made specifically for adults, data from childhood subjects on the semiautomated method were removed.

5.2 Materials and Methods

This work was a collaborative work that garnered data from healthy subjects from a few studies at the institution that was approved by the local Research Ethics Committee. Data for younger adults were obtained from the participants recruited using the similar study of the previous chapter 4.2, (REC B12012012a-051) while data for middle age and older adults was based on two studies, PaMIR (REC 14/EM/0061) and 'Brain Network Dysfunction in

Multiple Sclerosis (BRaNDy-MS) (REC 14/EM/0064). Data from PaMIR were also a subset of the data in Chapter 4.3 and Chapter 7. Data from BRaNDy-MS were collected and contributed by Dr Robert A Dineen, the chief investigator. Description of the projects is as detailed in Appendix 1, Appendix 2 and Appendix 3.

For younger adult data, the author of this thesis played identical roles as elaborated in the previous chapter (Chapter 4.2) to coordinate the study, including recruiting participants, data acquisition, organise and gathering data of all three studies, and finally performing both automated and semiautomated image processing techniques and their statistical analyses. For middle-aged and older adults' data, the author of this thesis was also involved in the PaMIR project to assist the volunteers with the scanning and data management that was explained in Chapter 4.3.

The outcome of semiautomated manual drawing approach was published recently (Xing et al., 2018). As the second co-author, I conducted the analysis while facilitated by Dr Yue Xing, who also provided hands-on guidance for image processing and analysis since that was my first experience in NM analysis. All authors of the published work were involved in the study design, including acquisition, analysis, and interpretation. This chapter explains the application of an automated template-based method and compared the findings with the published work on semiautomated manual drawing method.

5.2.1 Participants

Since the analysis using semiautomated methods was published, more healthy control data was collected from the PaMIR project. Available updated data from PaMIR with seven male volunteers aged 73.3 to 83.4 was added for this automated analysis. Together, about 92.5% of the data aged from 63 were from PaMIR, and the outcome of this analysis chapter is linked to that in Chapter 7. Data from the BRaNDy-MS project was included in order to achieve representative age distribution. However, no cognitive or depression tests were performed.

Approximately 70.2% aged above 47 underwent cognitive and depression assessments, to which cores yielded were within the normal range (Ma et al., 1999). All tests and assessments of medical and medication history were administered by Dr Saadnah Naidu. Finally, there were 94 healthy volunteers free from major neurologic, psychiatric, or medical disorders were involved in this study. Notably, this analysis removed 41 children's and teenagers' data since the created template is for adults only.

5.2.2 Image analysis

Unlike the published semiautomated method, the automated method did not analyse the suprathreshold volume and the whole SN CBR as it was not a true representative of the whole SN but the product of averaged-subregional SN CBR. The term dorsal SNpc and ventral SNpc in this chapter may also refer to anterior SNpc and posterior SNpc respectively in the semiautomated technique.

Both linear and non-linear regression analyses were performed to assess the associations between age and NM-sensitive CBR. Non-linear regression analyses were also performed on curve estimation models such as quadratic, cubic, logarithmic, S-curve, and exponential regression models. To find the best fit, adjusted R^2 values were compared to assess how the strength of the relationship between age and NM-rich area in SNpc was represented. The population were then grouped by decade to investigate age-related changes by every 10 years.

A comparison of sexes and hemispheres (left vs right) was also performed via univariate or multivariate analyses of variance, with a least-squares difference, or t-tests. Non-parametric analysis was conducted for non-normally distributed data. The significance level was set at $p<0.05$ was considered. Values were presented as mean ± standard deviation (mean ± SD).

5.3 Results

The final analysis of 94 healthy participants comprised 42 males and 52 females aged between 21.66 and 83.38 years old (mean age ± SD, 55.79±17.47). Scatter plots as shown in Figure 15 demonstrated a similar projection with both methods. A linear increase in CBR of subregional SNpc could be seen, which slowed down until it reached a plateau roughly for middle-aged adults before it decreased as age progressed. Visual inspection also recognised that for the automated template-based method, the CBR declined in old age at a lesser rate than the semiautomated method.

Nonlinear fitting curves showed significant correlations between age and SN-CBR in both subregions ($p < 0.0000001$). Quadratic regression curve fit revealed the best fit corresponding to a high adjusted R^2 of 0.67 for ventral and 0.56 for dorsal. The equation was as follows: Y = -0.068X2-0.00000391X+0.005 (ventral SNpc) and Y = -0.006X2-0.00000391X + 0.005 (dorsal SNpc) (Figure 15). In contrast to the semiautomated method, adjusted R^2 were higher with 0.82 in dorsal (anterior) and 0.74 in ventral (posterior).

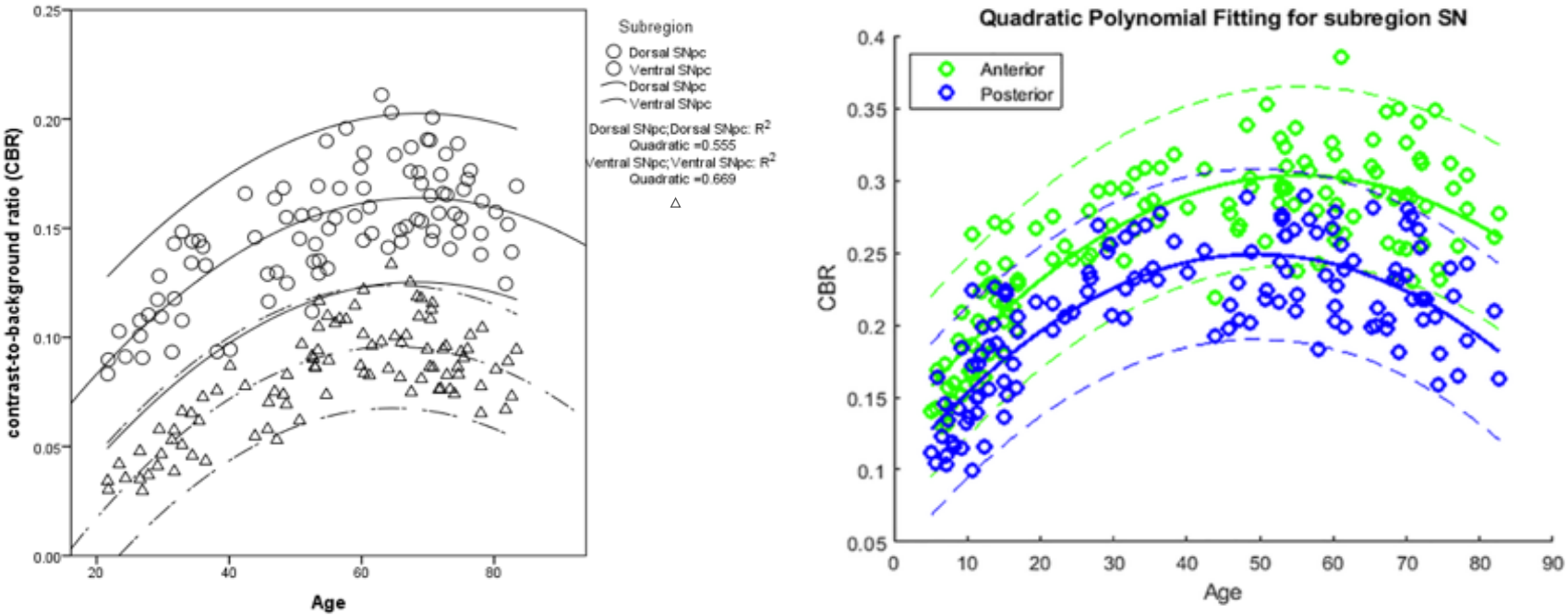


Figure 15. Scatter plot of age-related CBR changes with quadratic regression curve fit of SNpc subregions, dorsal and ventral SNpc (or anterior and posterior SNpc) for the automated template-based method (left) and the semiautomated manual drawing method (adopted from Xing et al., 2018) (right). Note that the automated method had no childhood data below 20 years old because the template was meant for adults only.

Figure 16 depicts age groups per decade visualised in a scatter plot with error bars of 95% confidence interval. In both subregions, long error bars are observed indicating high variability of CBR in the age group: 80-90 years old. Long error bars are also observed in dorsal SNpc for two other groups, 30-40 and 40-50 years old. The remaining groups showed short error bars with more precise or concentrated data.

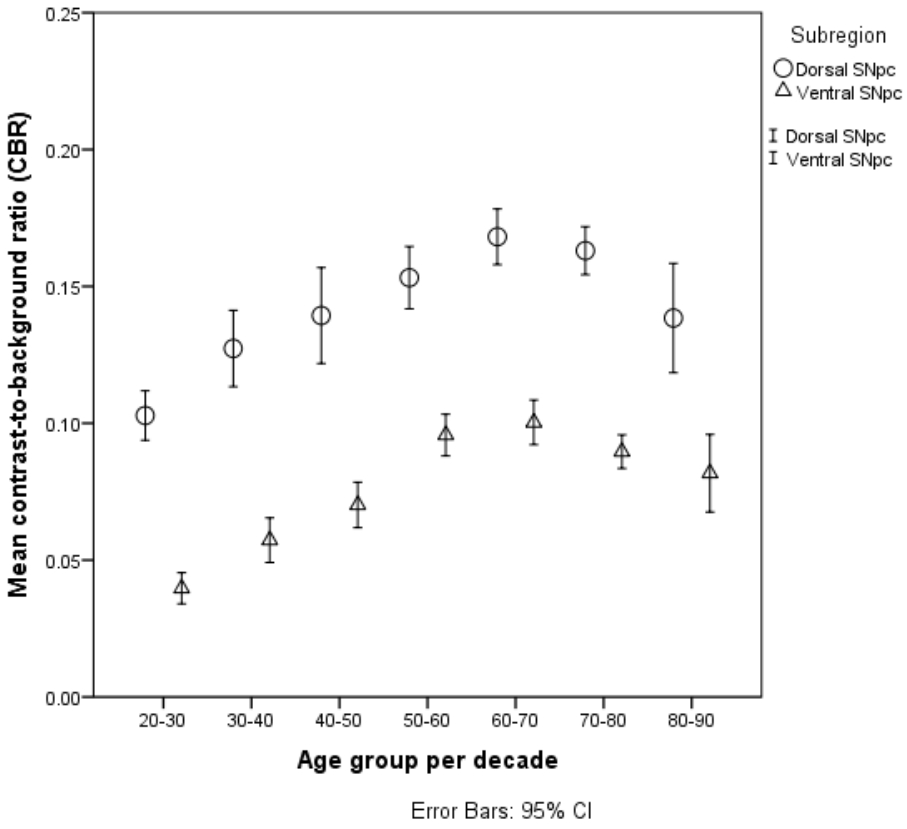
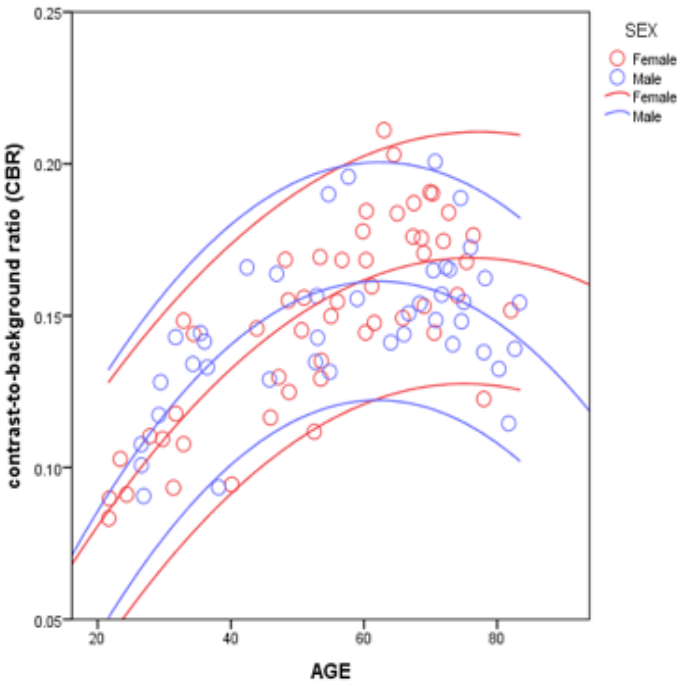
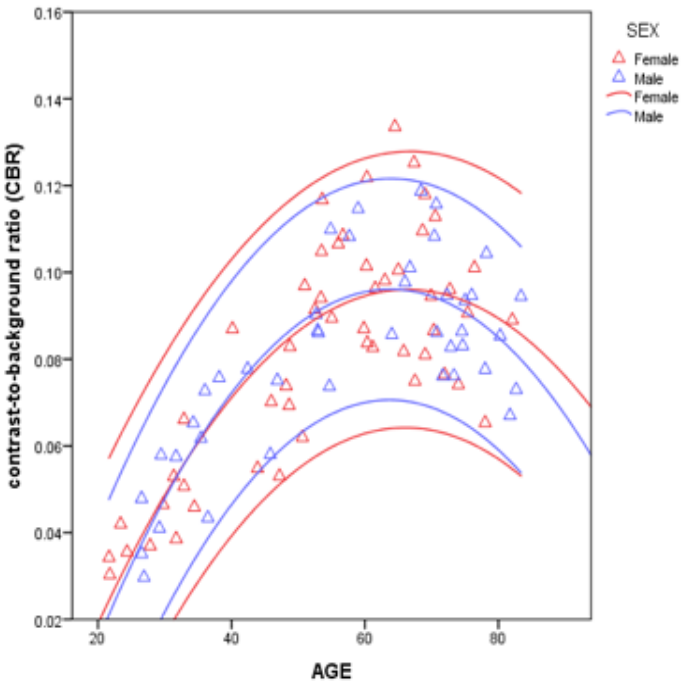


Figure 16. Scatter plot illustrating the mean CBR of each subregion, dorsal and ventral SNpc classified into age group by decades and error bars representing 95% confidence interval.

Apart from effects of age, gender, and laterality effects on SNpc were also respectively examined. No gender difference in CBR was found (female, 0.115 ±0.044; male, 0.113 ± 0.041; P = 0.715, Figure 17) and no laterality effect was observed in NM-rich CBR. Left and right pigmented SN were similar in CBR (left SNpc, 0.114 ± 0.025; right SNpc, 0.114 ± 0.024; P = 0.857, Figure 18). Therefore, gender and laterality effects confirmed the outcome of the semiautomated method. Scatter plots in Figure 17 and Figure 18 visualised age-related sex and lateral effect in dorsal and ventral SNpc measured by the automated approach. Quadratic curve fitting illustrated female curve lies above the male curve in both subregions after about age 60 suggesting less apparent loss of CBR in females than males. This might be explained by more men diagnosed with PD than women (Cerri et al., 2019).



Dorsal SNpc



Ventral SNpc

Figure 17. Scatter plots of SNpc CBR with quadratic curve fitting illustrated female curve lies above the male curve in both subregions after about age 60 for automated template-based method.

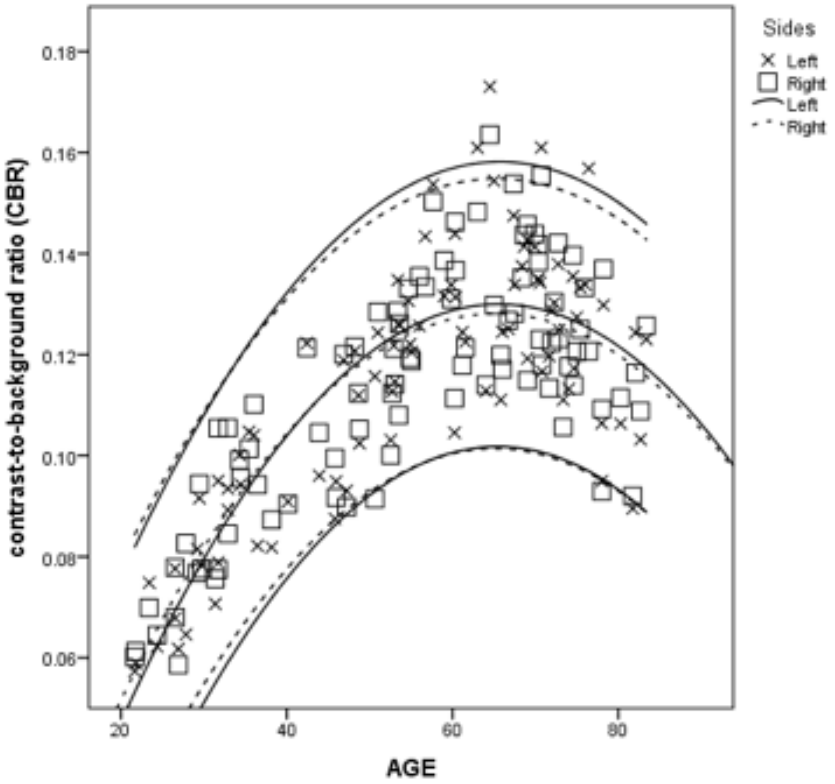


Figure 18. Scatter plot of SNpc CBR lateral effect with quadratic curve for automated template-based method.

5.4 Discussion

NM pigment detected in-vivo established a new paradigm for investigating its effect throughout the human lifespan, thereby elucidating its role in the development of neurodegenerative disorders. Replicating the published semiautomated method (Xing et al., 2018), this relatively large sample of healthy individuals demonstrated a similar trajectory. NM pigment increases rapidly from young adulthood until reaching a peak at middle age and then declines as age increases forming an inverted U-shape trajectory. CBR best fitted the quadratic curve and regional differences were found between CBR dorsal and ventral SN across all age groups.

The trajectory reveals a linear increase till adulthood, at which a plateau is reached. Thereafter, a decay at about mid-5th decade occurs until old age. As discussed in detail in the study published previously (Xing et al., 2018), the trajectory is consistent with several post-mortem lifespan studies explaining MRI CBR with neuromelanin concentration (Fearnley & Lees, 1991; Ma et al., 1999; D. M. A. Mann & Yates, 1974). Nevertheless, a similar signal reduction observed in this study suggested loss of NM-pigment in catecholaminergic nuclei during old age. Concurring with other histological studies, SN neurons decline significantly from 28.3% to 66.7% from age ~20 to 90 years old (Bartzokis et al., 2007; Luigi Zecca et al., 2001).

Post-mortem studies presented discrepancies among researchers evaluating age-related changes in pigmented SN. There were generally four views relevant to the phenomenon. Firstly, it was generally agreed that pigmented neurons are stable and no apparent loss is detected among the population of adults aged 44 to 110 (Alladi et al., 2009; Daniela Berg et al., 2015; Di Lorenzo Alho et al., 2016; Halliday et al., 1996; Kubis et al., 2000; Muthane et al., 1998; Pakkenberg et al., 1991). Meanwhile, a study on changes in lifespan among the younger population found that NM accumulation began at 3 to 4 years old (Alladi et al., 2009; Halliday et al., 2006). While pigmented neurons are steadily accumulated until 20 years old, after which they stabilised (Halliday et al., 2006). A separate study reported no significant change in melanised neurons with advancing age (Alladi et al., 2009).

The second view stemmed from the recruitment of subjects aged 1 to 97 years old, where it was discovered that NM pigment accumulates in the SN beginning at 3 years old and increases linearly till death (Luigi Zecca et al., 2002). In contrast, a separate group of researchers offered the third view based on the observation of an age-associated decrease in SN neuronal numbers from young adults (18 years) until old age (96 years) (Cabello et al., 2002; Fearnley & Lees, 1991; Ma et al., 1999; Rudow et al., 2008). The rates of nigral neuronal loss reported ranged from approximately 4.3% to 9.8% per decade, with a total loss of between 28.3% to 33% by age 90 to 96 (Cabello et al., 2002; Fearnley & Lees, 1991; Ma et al., 1999; Rudow et al., 2008).

Findings from the present study support the reports from the last group, which partially agree with the second and third views. Specifically, this study revealed that NM concentration in SN increased linearly as per the second view, up till middle age, before it began to decline somewhat around the sixth decade, as per the third view (D. M. A. Mann & Yates, 1974). Total loss of pigment by age 90 was 35% closer to the upper range of the third group (Fearnley & Lees, 1991).

Many within the research community concur that lifespan changes of NM are unclear due to methodological differences (Kubis et al., 2000; L Zecca et al., 2002), NM detection methods (Conde & Streit, 2006; Farrall & Wardlaw, 2009), cohort difference or possible conformational pigment changes during ageing (L Zecca et al., 1994, 1996, 2008; L Zecca & Swartz, 1993; Luigi Zecca et al., 2001).

Both this study and the previously published study (Xing et al., 2018) confirmed that dorsal or anterior subregions of the SNpc had higher NM signals, similar to previous histological and imaging findings (Luigi Zecca et al., 2002). While comparing outcomes from different methods, scatterplots revealed enhanced delineation of SNpc subregions when using the template-based method as there was no overlapping of CBR between the dorsal and ventral SNpc (Figure 15). The diverse results from the two methods might be caused by the ROI selection in which the automated template selected a consistent area at the mid dorsal area with high intensity and towards the end of ventral area, which is more sensitive to NM change or towards the susceptibility area of NM loss.

Consistent with the previous semiautomated study, gender and laterality showed no effect on NM concentration lifespan changes. Although the earlier study did report gender differences in volume measurement, which was argued to result from the male-female ratio in the sampled population.

The role of NM and Parkinson's pathology was also discussed in the previous study (Xing et al., 2018). Both studies upheld the neuroprotective role of NM as results revealed the presence of NM gradient from an early age which persisted over the lifespan, suggesting resilience to PD pathology (Luigi Zecca et al., 2002). Our recent study, Xing et al., 2022 and the research conducted by Gaurav et al. (2021) revealed the longitudinal changes of NM in the SN of healthy and PD patients. Nonetheless, more studies are needed to elucidate the neuroprotective role and lifespan alterations of NM, especially

in normal healthy subjects. Longitudinal study design also presents the best condition to corroborate human lifespan changes which this study was unable to conduct.

Another limitation in this experiment is that not all participants undergo cognitive or depression tests which they might have hidden underlying neurological disorder especially with the older age.

5.5 Conclusions

Alterations in NM throughout the human lifespan were detected using *in vivo* NM-MRI protocol. Using the automated method, a strong non-linear effect with an inverted U-shaped association postulated a comprehensive trajectory of NM pigmentation and depigmentation from adulthood to old age.

6. Meta-analysis and systematic review of neuromelanin-sensitive MRI in Parkinson disease

6.1 Meta-analysis of Application of Neuromelanin Weighted Imaging and Free-water Imaging in Parkinson's Disease

6.1.1 Introduction

Most PD cases are diagnosed only upon the occurrence of motor symptoms, at which point corresponds to approximately 50% to 60% of nigrostriatal dopaminergic neurons loss (Fearnley & Lees, 1991). A recent meta-analysis conducted on the UK Brain Bank Criteria, the most common diagnostic method for PD revealed an accuracy of 82.7% using pathologic examination as gold standard and high uncertainty during the early stages of the disease (Rizzo et al., 2016). This indicates unsatisfactory validity of diagnosis, especially during the early stages.

Extensive research on the novel MRI neuroimaging marker of PD has proven its high diagnostic accuracy. A recent review by Sulzer and colleagues (Sulzer et al., 2018) denotes NM-MRI as a promising biomarker in detecting depigmentation of the SNpc (Sasaki et al., 2006). Meanwhile, diffusion MRI-derived free water estimates served as an indirect marker of SN neurodegeneration (E Ofori et al., 2015). These markers are relatively new, and no systematic studies to assess their diagnostic accuracy, consistency, dependency on protocol, variations and generalisability.

This study aimed to conduct a systematic review and meta-analysis of NM MRI and free-water diffusion MRI of the SNpc as diagnostic biomarkers of PD. Readily available published studies were retrospectively collected to assess the effect size of the disease and diagnostic value and to identify requirements for future research. Literature was pooled to determine disease

effect size and compare diagnostic accuracy. The findings were also investigated for modifying factors, such as choice of protocol or regional specificity. This study has been prepared for publication in my role as the first author with input from my supervisors and all co-authors.

6.1.2 Methods

Meta-analysis and systematic review were conducted based on the 'preferred reporting items for systematic reviews and meta-analysis' PRISMA statement (Moher et al., 2015). In order to maintain the terms used in the pooled published studies, this chapter utilised the term "anterior SNpc" and "posterior SNpc" or "lateral" and "medial" to describe the dorsal SNpc and ventral SNpc, respectively.

6.1.2.1 Literature Search Strategy

A literature search of studies relevant to PD, NM, and MRI was conducted using four online databases: PubMed, MEDLINE, Web of Science, and EMBASE. Studies published until 1st October 2020 were searched for using keywords, which included "neuromelanin" OR "melanin" and "parkinsonism" OR "Parkinson disease" OR "Parkinson's disease" and "magnetic resonance imaging" OR "MRI" OR "MR imaging", respectively. Keywords used for free water studies included "free water" AND "diffusion imaging", alongside common keywords like "parkinsonism", "Parkinson disease", "Parkinson's disease" and "magnetic resonance imaging" or "MRI" or "MR imaging".

Results generated a list of articles which were then identified based on titles and abstracts. Next, identified studies were screened and their full

texts skimmed to determine if they fulfilled the predefined inclusion and exclusion criteria as shown below:

Inclusion Criteria

1. Case-control study comparing PD patients diagnosed using UK Parkinson's Brain Bank Criteria or MDS Parkinson Disease Criteria with healthy human controls without movement disorders.
2. Minimum of six participants for each group.
3. For the NM-MRI:

i) Use of dedicated T1-weighted neuromelanin protocol scanned using 3T field strength MRI.

ii) Extractable outcome of SN-related signal intensity.

iii) Studies testing diagnostic performance tests were included for meta-analysis of diagnostic test accuracy.

4. For the free water diffusion imaging:

i) Use of 3T diffusion tensor imaging (DTI) protocol with at least 32 directions.

ii) Free water index extractable from the publication.

Exclusion Criteria

1. Review articles, conference abstracts, and qualitative studies were excluded.

The literature search was conducted by the author of this thesis and Dr Saadnah Naidhu. Prior to the literature search, both individuals were trained in bibliographic database searching by an information specialist with

six years of experience, Rebecca Stevenson. Where dispute was raised regarding articles to be included in the review, a consensus was reached with supervisors.

6.1.2.2 Study quality and risk of bias assessment

Included studies were assessed for quality and risk bias via Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2). I and the information specialist independently administered the assessments. Disagreements were discussed and resolved together with approval from supervisors. All studies were finally included but due consideration was provided towards the risk of bias while drawing conclusions.

6.1.2.3 Data Extraction

6.1.2.3.1 NM-MRI

Analysis for disease effect size involved extraction of mean and standard deviation (SD) of PD patients and healthy control's SN measurement. In the case of missing SD value, SD was either calculated using the SPSS software based on mean and *p*-value (Ogisu et al., 2013) or using the formula, SD = SE * √n from the SE value extracted (Huddleston et al., 2017). Mean and SD was also estimated by converting median and range reported from a few studies (Castellanos et al., 2015; Fabbri et al., 2017; Chigumi Ohtsuka et al., 2013) as per previously published conversion techniques (Hozo et al., 2005; Matsuura et al., 2013). For studies published under the same research group with overlapping populations, the study with the highest number of subjects was selected (Fabbri et al., 2017; Chigumi Ohtsuka et al., 2013; Schwarz et al., 2017). If no whole SN values were

reported with only sub-regional data available (lateral, medial, anterior or posterior SN), combined mean and SD were estimated to obtain whole SN values (Isaias et al., 2016; Prasad et al., 2018; Xiang et al., 2017). Specificity and the cut-off value were also extracted for analysis of diagnostic performance sensitivity.

I and Rebecca Stevenson extracted the data, and disagreements were resolved by consensus. Demographic data: including age, sex and sample size, clinical data of the PD patients, details of the MR acquisition, substantia nigra segmentation method and brief outcomes were collected and summarised in the result section.

6.1.2.3.2 Free water

Mean and SD of the SN free-water index from all included studies was extracted. Since not all the included studies reported free-water index in the SN sub-regions (anterior and posterior), a whole SN meta-analysis was not performed. For studies with overlapping cohorts, the studies with a larger number of participants were selected (Planetta et al., 2016). Demographic data such as age, gender, sample size, and PD patients' clinical data were also extracted.

6.1.2.4 Statistical analysis

A meta-analysis of disease effect size was conducted using Review Manager software (RevMan, Version 5.3. Copenhagen: The Nordic Cochrane Centre, the Cochrane Collaboration, 2014). Disease effect size was represented by standardised mean difference (SMD), where SMD < 0.2 was

considered small, > = 0.2 < 0.6 was moderate, and > 0.8 was large or important. Two-tailed *p*-value < 0.05 was considered statistically significant. I^2 values (Q-statistics) represented heterogeneity assessment and a random effect model was employed when I^2 > 50%. Forest plots generated utilised the random-effects model using Der Simonian and Laird method.

For meta-analysis of diagnostic performance, the 'metandi' command via STATA 15 statistic software was utilised. Sensitivity, specificity, and the cut-off value of available data were extracted.

6.1.2.5 Investigation of Sources of Heterogeneity

Potential sources of heterogeneity were investigated through subgroup analyses. Forest plots were assessed for factors that may affect heterogeneity. Subgroup analyses were able to identify factors contributing to heterogeneity.

6.1.2.5.1 NM-MRI

Subgroup analyses were conducted based on two categories: method of segmentation and type of NM measurement metric. Methods of segmentation were automatic, semi-automatic, and manual segmentation. NM-SN intensity metrics included contrast ratios, and 2D or 3D volumetric measurement.

6.1.2.5.2 Free water

For free-water imaging, all studies were from the same research team which utilised the same image processing method for segmentation and free water fitting. Hence, subgroup analysis could not be performed.

6.1.2.6 Publication bias

Funnel plot was used to identify publication bias.

6.1.3 Results

6.1.3.1 Literature Search Results

6.1.3.1.1 NM-MRI

The literature search yielded a total of 467 articles. Upon removing duplicates and screening the titles and abstracts, 54 articles were selected for further evaluation. Twelve studies were excluded due to nine being conference abstracts, two being review articles, and one post-mortem study. Following full content assessment, 21 full-text articles were excluded due to non-eligibility based on inclusion and exclusion criteria. The final pool comprising 19 articles was eventually included, with 17 for meta-analysis of disease effect size (Castellanos et al., 2015; Fabbri et al., 2017; Huddleston et al., 2017; Isaias et al., 2016; Kashihara et al., 2011b; Le Berre et al., 2019; Matsuura et al., 2013; Moon et al., 2016; Ogisu et al., 2013; Chigumi Ohtsuka et al., 2013; Okuzumi et al., 2019; Prasad et al., 2018; N. Pyatigorskaya et al., 2018; S. T. Schwarz et al., 2017; H Takahashi et al., 2018; Trujillo et al., 2017; J. Wang et al., 2018; Xiang et al., 2017; Zupan et al., 2019) and eight articles (six already included for analysis of disease effect size) for meta-analysis of

diagnostic test accuracy (Ariz et al., 2019; Castellanos et al., 2015; Fabbri et al., 2017; Ogisu et al., 2013; N. Pyatigorskaya et al., 2018; S. T. Schwarz et al., 2017; Shinde et al., 2019; Zupan et al., 2019). No articles were found outside the online databases. The processes are simplified as per PRISMA flowchart below (Figure 19).

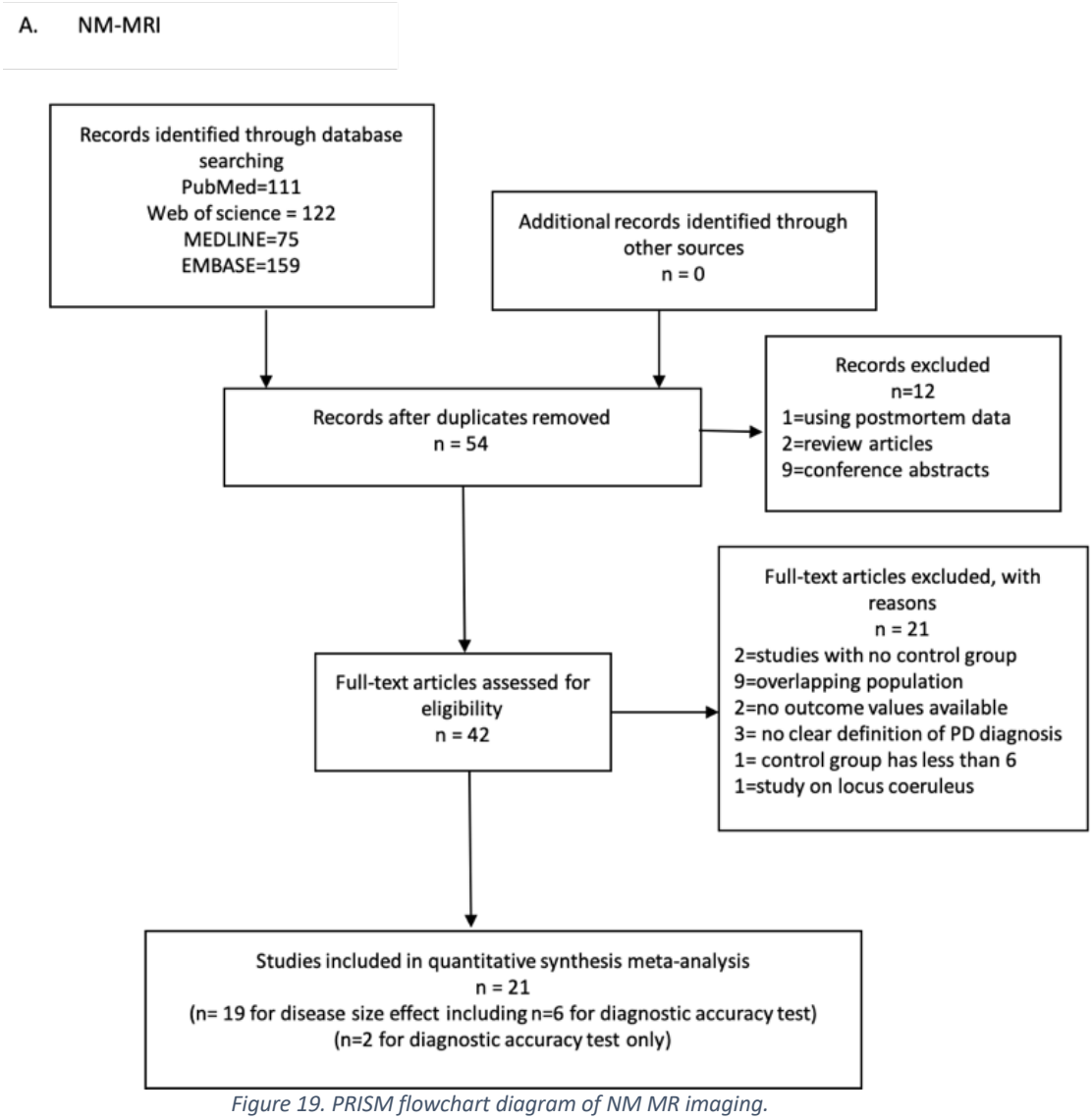


Figure 19. PRISM flowchart diagram of NM MR imaging.

6.1.3.1.2 Free water

The literature search for free water studies 61 articles, but the number was reduced to 14 after screening of titles and abstracts. Seven were excluded (two review articles, three conference abstracts, one study due to overlapping population and one study with no control group). All full-text articles were examined for inclusion criteria. The final selection included seven articles (Arribarat et al., 2019; Burciu et al., 2017; Guttuso et al., 2018; Edward Ofori et al., 2015, 2017; Planetta et al., 2016; Yang et al., 2019). No articles were found outside the online database. The processes are simplified as per PRISMA flowchart below.

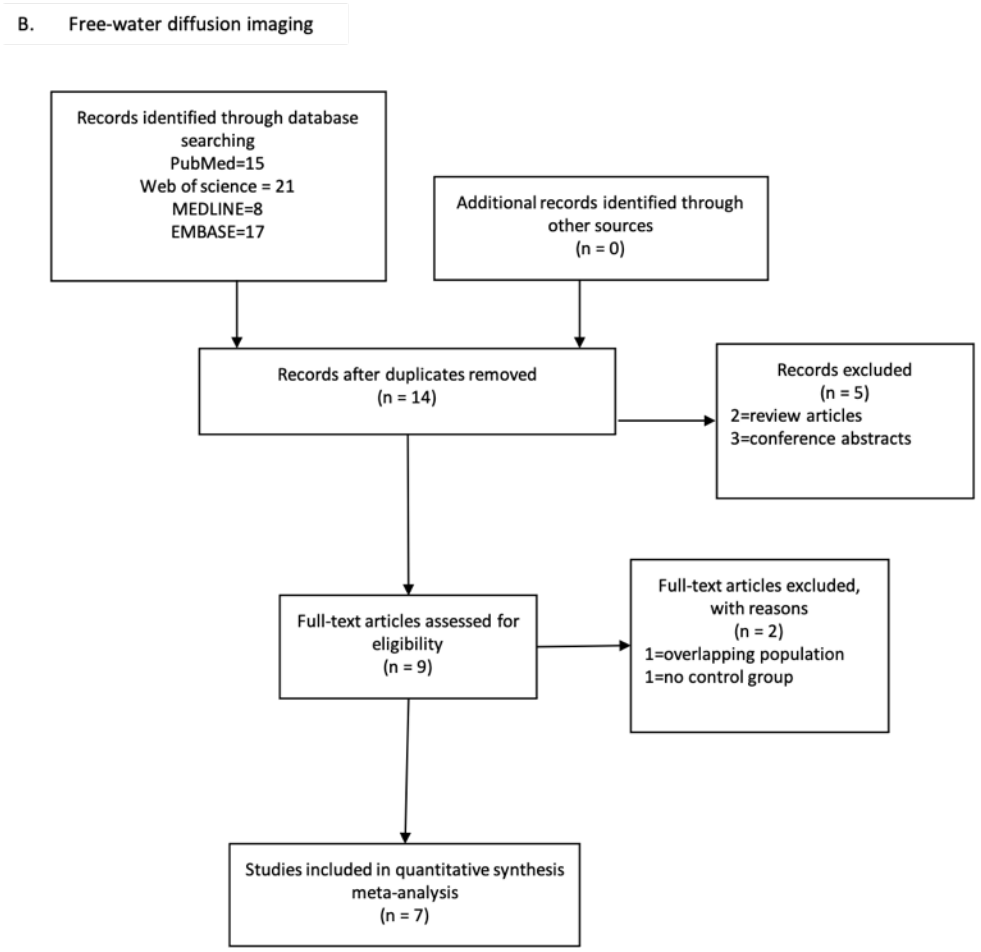


Figure 20. PRISM flowchart diagram of free water imaging.

6.1.3.2 Characteristics of studies included into the diagnostic meta-analysis

6.1.3.2.1 NM-MRI

The 19 studies involved a total of 1120 participants (636 PD patients and 484 healthy controls) from different parts of the world (i.e., United Kingdom, France, India, China, Japan, Korea, Portugal, Italy, the United States and Spain) with age ranging between 56.3 to 76.5 years old. Participants fell within a range of disease severity stage from Hoehn and Yahr I to V. Mean disease duration of pooled studies ranged from 1.6 to 14 years. Table 9 and Table 10 below briefly summarises the included studies and ensuing participants' details. Figure 21 and Figure 22 depicts the QUADAS-2 assessment of quality and risk among the included studies.

*Table 9. Demographic details of the included studies in NM-MRI. Abbreviations: ON=on medication, OFF=off medication, *=median±SD, **=MDS-UPDRS*

Studies	Healthy Control		Parkinson Disease					
	N	Age (mean±SD)	N	PD group	Age (mean ± SD)	UPDRS-III	H-Y score	Disease Duration (year)
Castellanos et al 2015	37	61.5 ± 5.3	23		63.8 ± 4.3	12.5 ± 3.5		5.8 ± 1.4
Fabbri et al 2017	10	61.1 ± 4.1	12	De novo	62.8 ± 6.13	35.1 ± 5.3 **	2.0	< 0.5
			10	2-5 years	66.7 ±2.2	26.4 ± 8.6 **	2.0	2.0-5.0
			13	LSPD	76.5 ± 3.7	49.1 ± 3.6 **		14.0
Huddleston et al 2017	19	71.3 ± 1.2	22		60.4 ± 1.8	18.7 ± 2.3		6.8 ± 0.7
Isaias et al 2016	18	60.0 ± 8.7	18		62.8 ± 8.9	14.5 ± 5.8	2.0	10.0 ± 0.842
Kashihara et al 2011	54	71.6 ± 10.8	80		70.9 ± 8.20			5.1 ± 4.7
Le Berre et al 2019	60	70.8 ± 3.7	62		70.2 ± 6.45	22.5 ± 15.7	2.7 ± 0.9	9.9 ± 6.2
Matsuura et al 2013	23	71.2 ± 8.9	32		71.4 ± 8.6		2.6 ± 1.0	6.0 ± 6.7
Moon et al 2016	13	72.2 ± 10.2	10		71.7 ± 8.0		2.3 ± 0.7	
Ogisu et al 2013	27	65.9 ± 6.7	18		68.8 ± 6.4			10.3 ± 6.1
Ohtsuka et al 2013	22	69.5 ± 5.774	30	EARLY	65.3 ± 8.1	30.6 ± 7.3	1.8 ± 0.1	1.6 ± 0.8
Okuzumi et al 2019	23	59.0 ± 19.0	35		60.0 ± 13.0	17.0 ± 11.3	2.09 ±1.2	7.6 ± 4.1
Prasad et al 2018	15	56.3 ± 8.5	44		58.0 ± 8.7	39.0 ± 15.75	2.28 ± 0.6	6.4 ± 4.1
Pyatigorskaya et al 2018	20	60.8 ± 8.3	36		60.1 ± 10.1	28.3 ± 9.6 (off)/14.8 ± 8.8 (on)	1.9 ± 0.6	9.1 ± 3.6
Schwarz et al 2017	30	65.4 ± 7.0	39	EARLY	66.2 ± 9.1		1.7 ± 0.7	4.8 ± 4.3
			31	ADVANCED	64.8 ± 6.1	53.0 ±16.7	4.0 ± 0.6	13.8 ± 4.9
Takahashi et al 2018	25	67.1 ± 6.8	39		69.2 ± 6.0	13.1 ± 7.2	2.9 ± 0.8	5.2 ± 3.6 *
Trujillo et al 2018	8	59.0 ± 7.3	16		63.0 ± 6.7			7.0 ± 3.0
Wang et al 2018	28	63.4 ±7.4	51	De Novo	63.9 ± 9.9	17.4 ± 9.3		1.8 ± 1.5
Xiang et al 2017	20	66.0 ± 9.9	14	PIGD	59.8 ± 8.4	30.2 ± 7.7	2.2 ± 0.3	3.3 ± 1.7
			9	TD	58.8 ± 7.7	32.1 ± 9.9	2.1 ± 0.4	3.8 ± 3.4
Zupan et al 2019	12	64.0 ± 4.8	20		67.1 ± 4.8		2.1 ± 0.2	3.7 ± 2.8

Table 10. Overview of the included studies in the NM imaging.

Studies	Scanner	Sequence	Segmentation method	Software /Toolbox	Outcome Measure	SNpc vs Control
Castellanos et al 2015	Siemens	T1 FSE	Automatic	N/A	Volume	Reduced in PD
Fabbri et al 2017	Philips	T1 FSE	Semi-Automatic	Osirix	Area, Contrast ratio	Reduced in PD
Huddleston et al 2017	Siemens	2D GE Pulse Sequence with MTC preparation	Semi-Automatic	FMRIB MATLAB	Magnetization Transfer Contrast	Reduced in PD
Isaias et al 2016	Philips	T1 FSE	Semi-Automatic	3D SLICER+ MATLAB	Volume, Contrast ratio	Reduced in PD
Kashihara et al 2011	GE	T1 FSE	Manual	Pentablet pointing device	Volume	Reduced in PD
Le Berre et al 2019	Siemens	T1 FSE	Automatic	Python 3.6, DL framework of Chainer 3.2.0	Area	Reduced in PD
Matsuura et al 2013	Siemens	T1 FSE	Manual	N/A	Contrast ratio	Reduced in PD
Moon et al 2016	GE	T1 FSE	Semi-Automatic	MIPAV	Volume	Reduced in PD
Ogisu et al 2013	Philips	3D T1 TFE	Semi-Automatic	Perfusion Mismatch analyzer	Volume	Reduced in PD
Ohtsuka et al 2013	GE	T1 FSE	Manual	N/A	Contrast ratio	Reduced in PD
Okuzumi et al 2019	Philips	T1 FSE	Manual	MRIcron	Area	Reduced in PD
Prasad et al 2018	Philips	High Resolution 3D T1 SPIR	Semi-Automatic	Osirix	Contrast ratio	Reduced in PD
Pyatigorskaya et al 2018	Siemens	T1 TSE	Manual	FreeSurfer	Contrast ratio	Reduced in PD
Schwarz et al 2017	Philips	T1 Spectral presaturation inversion - recovery	Manual	JIM	Volume	Reduced in PD
	Philips	T1 with MT				
	GE	T1 SE with MT				
Takahashi et al 2018	GE	T1 FSE	Automatic	N/A	Area	Reduced in PD
Trujillo et al 2018	Philips	T1 TSE	Semi-Automatic	3D slicer	Pool size ratio	Increased in PD
Wang et al 2018	GE	T1 FSE	Manual	AWD4.6 workstation	Width, contrast ratio	Reduced in PD
Xiang et al 2017	Philips	T1 TFE	Manual	N/A	Contrast ratio	Reduced in PIGD and TD
Zupan et al 2019	Philips	T1 GRE	Semi-Automatic	AFNI	Area, Volume	Reduced in PD

QUADAS-2 assessment tool demonstrated low risk of bias for all the included studies.

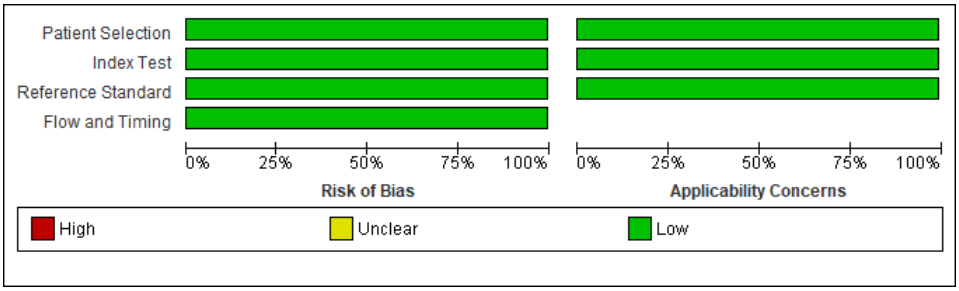


Figure 21. Summary of bias risk assessment results for QUADAS-2 for NM-MRI

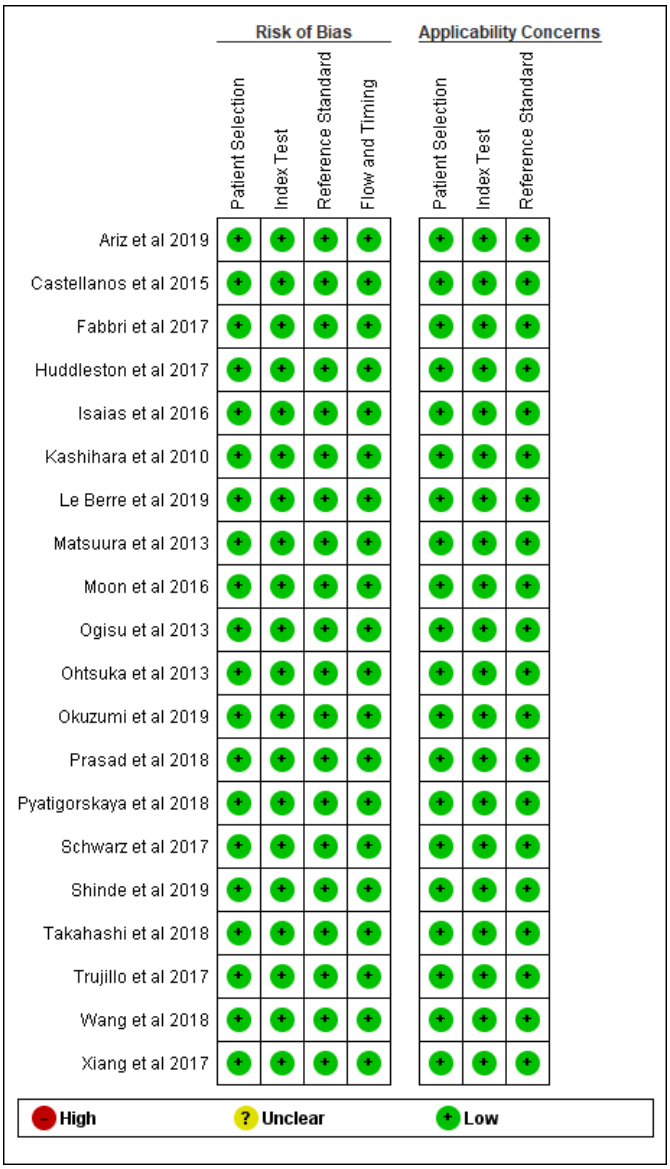


Figure 22. Quality of included studies according to QUADAS-2 guidelines for NM-MRI

6.1.3.2.2 Free water

All seven studies comprised a total population of 942 subjects (577 PD patients and 365 healthy controls) from different countries (i.e., United States, Austria, Germany, Italy, United Kingdom, France, Spain, Australia, Norway, and Israel) with age ranging from 56.5 to 68.4 years old. The mean disease duration of pooled studies ranged from 0.58 to 40.9 years. Participants fell within disease severity stage ranging from Hoehn and Yahr I to V. Data on the study overview and participant details are presented in Table 11 and Table 12. Study quality analysis and risk assessment, QUADAS-2 is depicted in Figure 23 and Figure 24.

*Table 11. Demographic details of the included studies in free water imaging. *Median **Interquartile range*

Studies	Healthy Control			Parkinson Disease				
	Group	N	Age (mean±SD)	N	Age (mean±SD)	MDS UPDRS-III	H-Y score	Disease Duration (year)
Arribarat et al 2019		21	66.0 ± 4.9	18	65.2 ± 6.6	11.4 ± 4.9		82 ± 56.3 (months)
Burciu et al 2017		49	60.3 ± 10.4	103	60.4 ± 9.5	20.1 ± 8.8	44, H&Y I 59, H&Y II	0.9 ± 0.6
Guttuso et al 2018		19	56.5 ± 10.1	19	59.8 ± 8.4	18.8 ± 7.3	1.9*, 2.1-2.5**	7.1 ± 5.1
Ofori et al 2015	University of Florida	20	62.7 ± 8.9	28	64.7 ± 8.2	29.8 ± 9.2		40.9 ± 21.5 (months)
	PPMI	56	59.5 ± 11.5	78	61.6 ± 9.2	22.9 ± 8.7		7.9 ± 7.8
Ofori et al 2017	PSHMC	30	64.2 ± 8.7	45	68.4 ± 9.6	36.7 ± 27.1	2.4±1.2	3.7 ± 5.0
	University of Florida	36	63.8 ± 9.6	54	64.1 ± 9.2	26.3 ± 10.9	1.8 ± 0.5	3.3 ± 2.4
	Medical University of Innsbruck	41	61.5 ± 8.9	85	65.1 ± 9.1	35.8 ± 12.6	2.3 ± 0.6	6.0 ± 4.3
Planetta et al 2016		18	66.9 ± 6.6	18	67.9 ± 6.1	37.3 ± 6.6		
Yang et al 2019		75	65.97 ± 11.5	129	65.38 ± 7.8	32.6 ± 14.3	2.44 ± 0.6	6.0 ±4.2

Table 12. Overview of the included studies in the free water imaging.

Studies	Scanner	Segmentation method	Software /Toolbox	SNpc vs Control	B values (s/mm²)	direction	TR/TE	FOV (mm)	Slice thickness (mm)
Arribarat et al 2019	Philips Achieva	Manual	FSL, MATLAB	Increase free water	0,1000	32	11696/76	-	-
Burciu et al 2017	Siemens Trio	Manual	FSL, MATLAB	Increase free water	0,1000	64	-	-	2
Guttuso et al 2018	GE Signa	Manual	FSL, MATLAB	Increase free water	0,900	25	-	-	4
Ofori et al 2015	Philips Achieva	Manual	FSL, MATLAB	Increase free water	0,1000	64	7748/86	224 x 224	2
	Siemens Trio	Manual	FSL, MATLAB		0,1000	64	900/88	230 x 230	2
Ofori et al 2017	Siemens Trio	Manual	FSL, MATLAB	Increase free water	0,1000	42	8300/82	256 x 256	2
	Philips Achieva	Manual	FSL, MATLAB		0,1000	64	7748/86	224 x 224	2
	Siemens Vario	Manual	FSL, MATLAB		0,1000	20	8200/83	230 x 230	3
Planetta et al 2016	Philips Achieva	Manual	FSL, MATLAB	Increase free water	0,1000	64	7748/86	224 x 224	2
Yang et al 2019	Philips Achieva	Manual	FSL, MATLAB	Increase free water	0, 800	15	8044/67	-	-

Using the QUADAS-2 assessment tool, all the included studies were generally at low risk of bias.

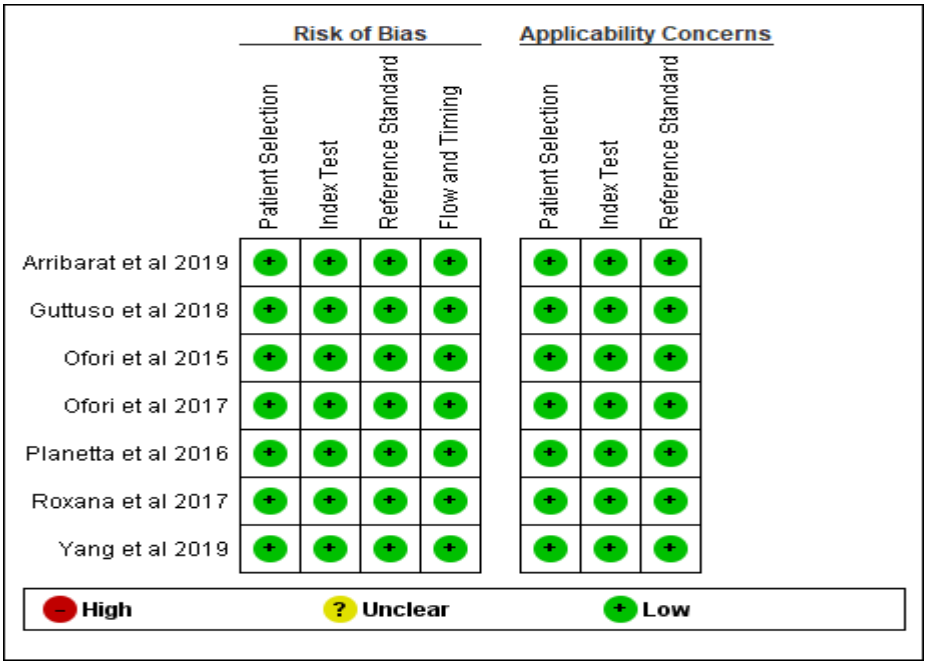


Figure 23. Summary of bias risk assessment results for QUADAS-2 for free water MRI.

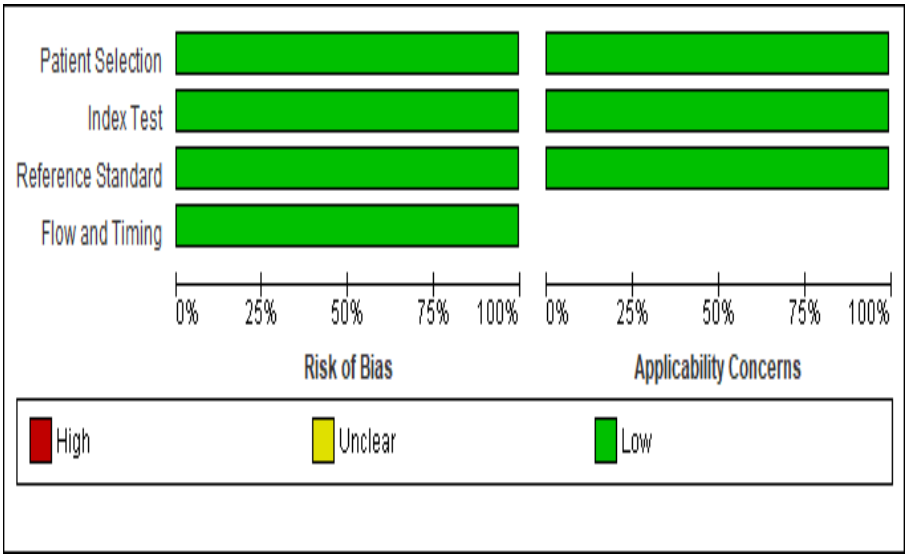


Figure 24. Quality of included studies according to QUADAS-2 guidelines for free water MRI

6.1.3.3 Main Results

6.1.3.3.1 NM-MRI

6.1.3.3.1.1 Whole SN

All studies were consistent in demonstrating reduced pigmented neuron signal intensity in SN of PD patients against healthy control, which was reflected based on an overall large effect size [SMD = 1.41, 95% CI, 1.08-1.75, $p < 0.00001$, Figure 25]. Nevertheless, moderate to high heterogeneity was observed in the population ($I^2 = 82\%$; $p<0.00001$) mainly due to one study reporting the largest disease effect size (Castellanos et al., 2015). The same event was observed in one of three studies that implemented the automated segmentation method of SN.

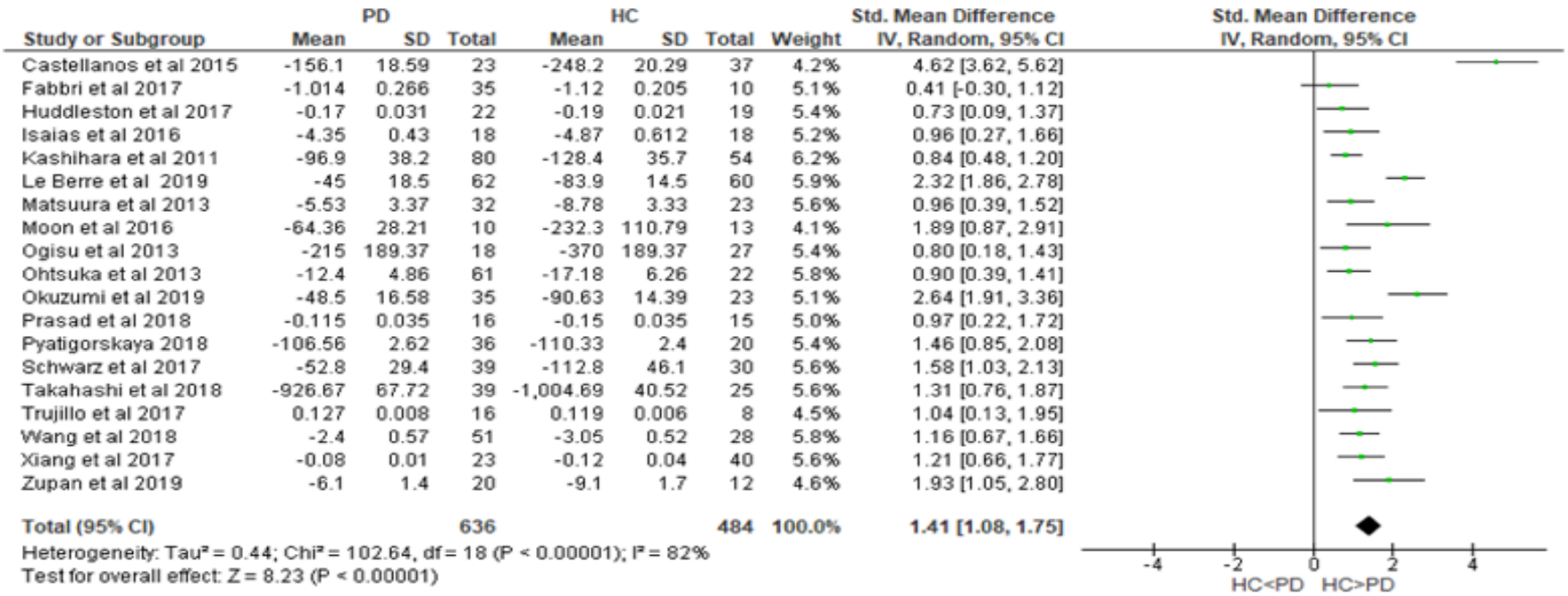


Figure 25. Forest plot of different studies of NM-MRI of the substantia nigra when comparing Parkinson patients and healthy controls. Whole SN studies demonstrate a NM decrease within the substantia nigra in Parkinson's as indicated by a highly significant positive disease effect sizes, 1.41, p<0.00001, with moderate to high heterogeneity (I^2=82%).

Besides, there were eight articles (235 = PD, 184 = HC) reporting diagnostic accuracy using respective methods as represented by the forest plot of sensitivity and specificity in Figure 26 below. The pooled sensitivity was approximately 91% (95% CI: 84% - 94%) while the pooled specificity was 83% (95% CI: 79% - 87%). There was a considerable variation in study-specific cut-off values with a range of 73 mm^3 to 285mm^3 for volumetric studies, 33mm^2 for area metric approach in one study, and two studies reported no cut-off value.

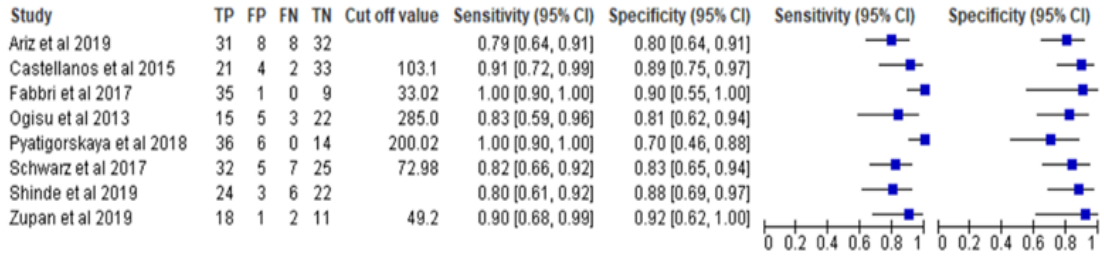


Figure 26. Forest plots of sensitivity and specificity with the pooled sensitivity was estimated to be 91% (95% CI: 79% - 96%) and the pooled specificity was 83% (76% - 88%). The cut-off values (range 73 - 285mm^3) in volumetric studies, 33mm2 in an area metric approach in one study and 2 studies reported no cut-off value.

6.1.3.3.1.2 Anterior SN

Of the 19 studies, seven reported pigmented neurons signal intensity in the anterior SN sub-region among 425 participants (PD = 257 and HC = 168). A significant moderate disease size effect [SMD=0.74, 95% CI, 0.47-1.01, $p < 0.00001$, Figure 27] and moderate heterogeneity were detected among these cohorts ($I^2 = 38\%$, $p < 0.14$). Pooled sensitivity and specificity of the anterior SN could not be obtained since no studies reported diagnostic accuracy.

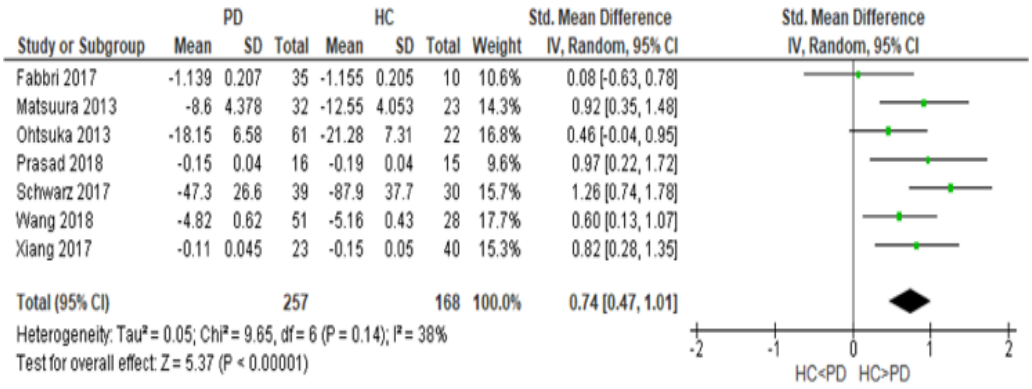


Figure 27. Forest plot of anterior SN demonstrate a NM decrease within the substantia nigra in Parkinson's as indicated by a highly significant positive disease effect sizes, 0.74, p<0.00001, with moderate heterogeneity (I^2=38%).

6.1.3.3.1.3 Posterior SN

Eight studies examining posterior SN subregion reported a significantly larger disease effect size compared to the anterior SN [SMD = 1.10, 95% CI, 0.71-1.48, p < 0.00001, Figure 28], with moderate to high heterogeneity (I^2 = 70%, p < 0.001). A total of 466 subjects (PD = 279 and HC = 187) were involved. Only two studies (Prasad et al., 2018; Schwarz et al., 2017) calculated diagnostic accuracy, thus sensitivity and specificity could not be pooled. Besides, these two studies used different approaches to measure different SN metrics in anterior and posterior SN (Prasad et al., 2018; Schwarz et al., 2017).

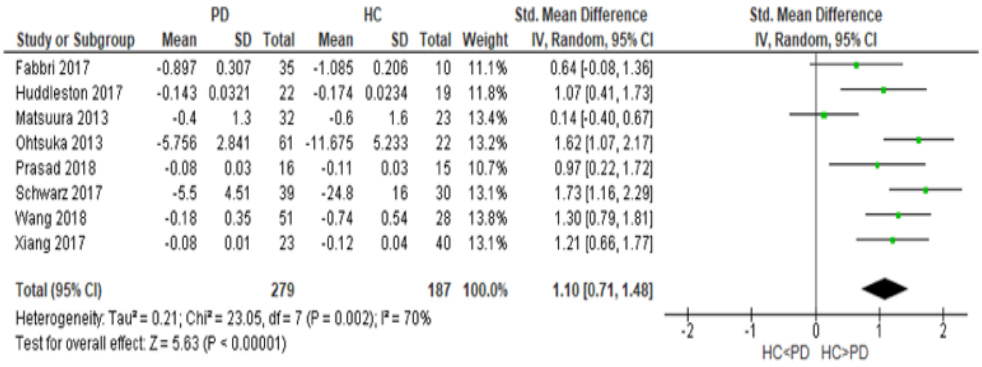


Figure 28. Forest plot of posterior SN demonstrate a NM decrease within the substantia nigra in Parkinson's as indicated by a highly significant positive disease effect sizes, 1.10, p<0.00001, with moderate to high heterogeneity (I^2=70%).

6.1.3.3.2 Free water

6.1.3.3.2.1 Anterior SN

An increase in free water was consistently reported in the anterior SN sub-region of PD patients against HC, with an overall moderate disease effect size (SMD = 0.39, 95% CI, 0.13- 0.65, p < 0.003, Figure 29). Heterogeneity was moderate to high (I^2 = 68%, p < 0.005). The SMD via this method was smaller compared to NM-MRI, but there was an overlap of confidence intervals.

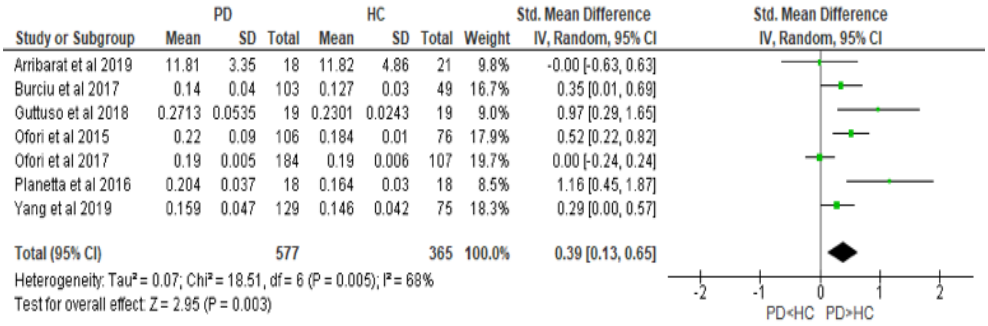


Figure 29. Forest plot of different studies of free water of the substantia nigra when comparing PD patients and healthy controls. All studies demonstrate an increase of free water in PD patients as indicated by a positive disease effect size, 0.39. There was significant (p<0.003), however moderate to high heterogeneity was observed (I^2 = 68%).

6.1.3.3.2.2 Posterior SN

A larger disease effect size was observed in the posterior SN of PD patients (SMD = 1.64, 95% CI, 0.41-2.86, p = 0.009, Figure 30) for similar free

water direction. However, the results could not be pooled for meta-analysis due to high heterogeneity (I^2 = 98%; p < 0.00001).

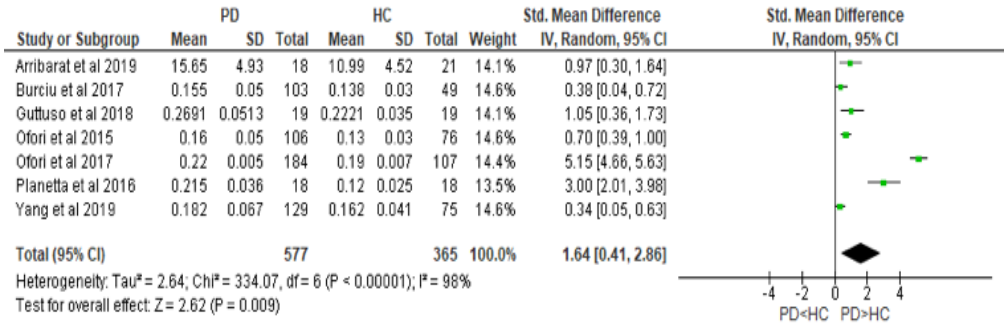


Figure 30. Forest plot of different studies of free water of the posterior substantia nigra when comparing PD patients and healthy controls. All studies demonstrate an increase of free water in PD patients as indicated by a positive disease effect size, 1.64. There was significant (p < 0.009), however high heterogeneity was observed (I^2 =98%).

6.1.3.4 Subgroup analysis

6.1.3.4.1 NM-MRI

Subgroup analysis according to 'segmentation method' and 'measurement metrices' were used to assessed heterogeneity.

6.1.3.4.1.1 Segmentation method

In this category, three studies were based on automatic segmentation, eight studies were based on semi-automatic segmentation, and the remaining eight were based on manual segmentation (Figure 31). No significant subgroup differences were found between these three methods (subgroup difference: p = 0.06, I^2 = 63.6%).

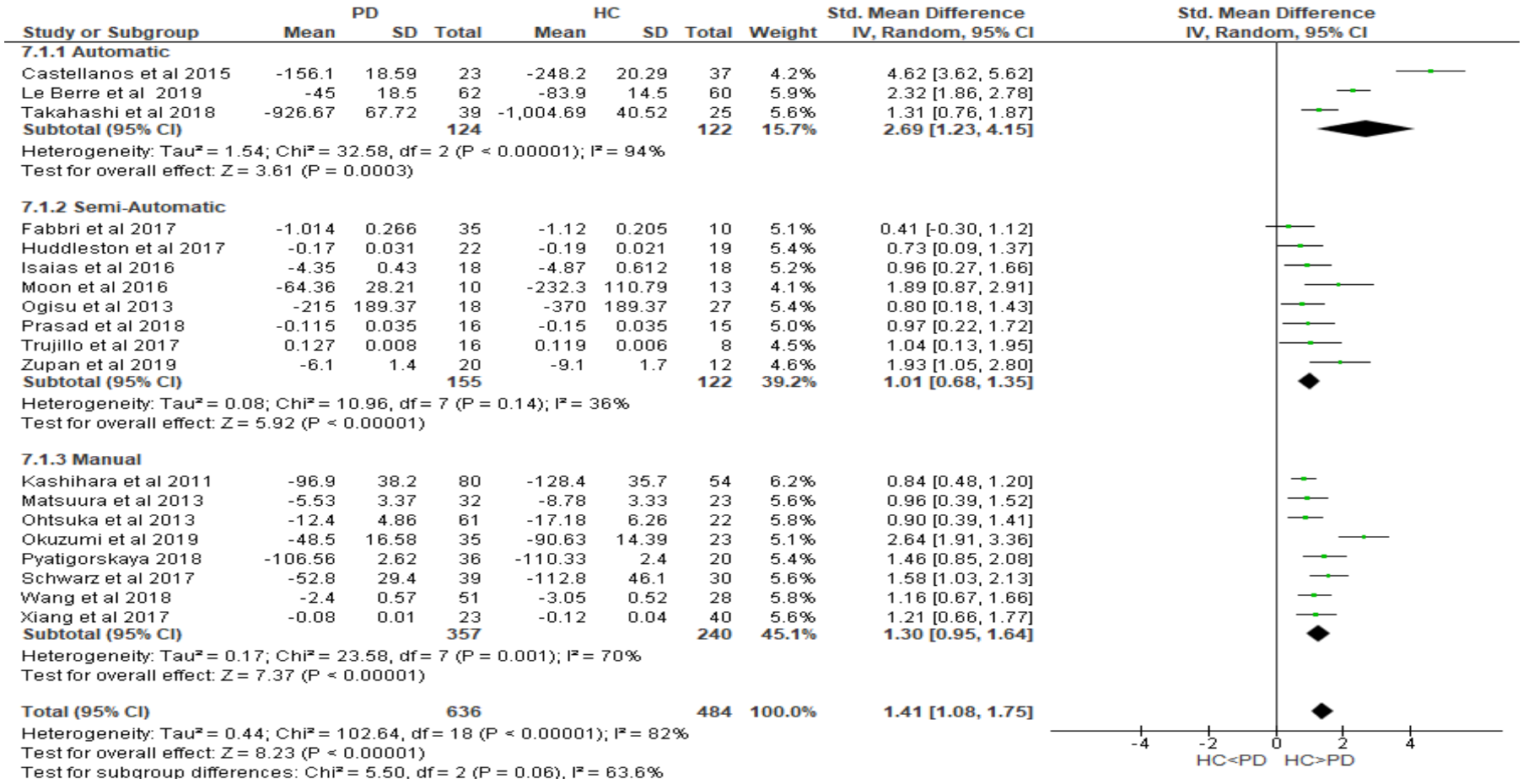


Figure 31. Forest plot of subgroups analysis for segmentation method, there is no significant subgroup differences with p=0.06.

6.1.3.4.1.2 SN intensity measurement metrices

For SN intensity measurement metrics, 10 studies estimated SN pigmentation using contrast ratio, whereas six studies used 3D and three used 2D morphometry (Figure 32). There were significant differences between subgroups, with a larger effect size for all methods ($p = 0.00001$ and $I^2 = 95.3\%$) mainly attributed to 3D volumetric studies reporting the largest I^2 of 91%. The other subgroups reported low values of I^2, 0%.

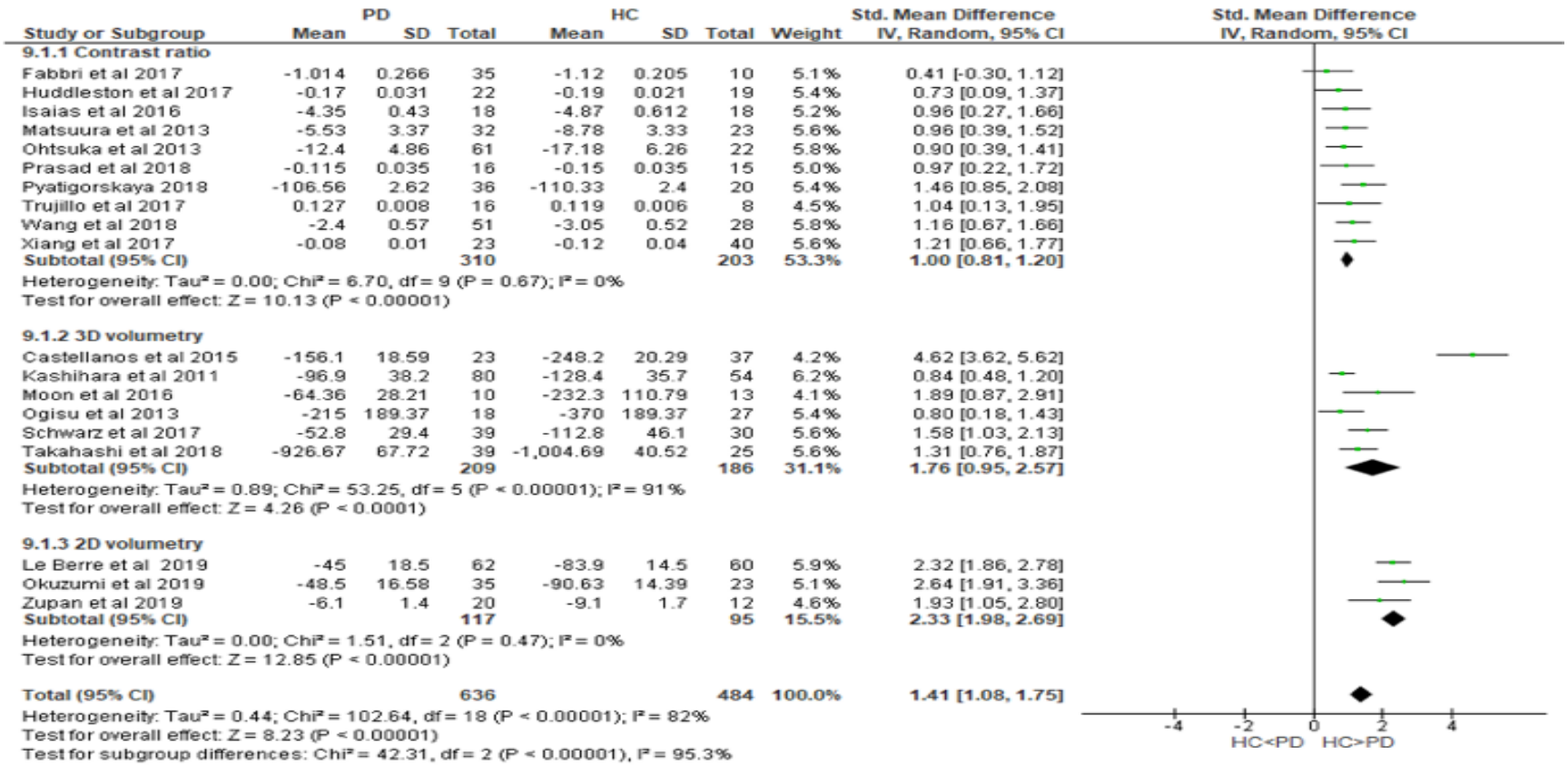


Figure 32. Forest plot of subgroups analysis for SN intensity measurement matrices, there is significant subgroup differences with p = 0.00001.

6.1.3.4.2 Free water

The source of heterogeneity could not be investigated given the high heterogeneity within the entire and posterior SN. Although moderate heterogeneity was found for the anterior SN, the patient population could not be grouped due to insufficient information. Image processing methods could not be differentiated into groups as standardised processing methods were utilised.

6.1.3.5 Publication bias

6.1.3.5.1 NM-MRI

Funnel plots analysing publication bias appeared to be asymmetric when visually inspected. Nevertheless, the asymmetry was mainly caused by an outlier study (Castellanos et al., 2015) [red symbol in Figure 33]. Therefore, it was not considered a major publication bias among included studies.

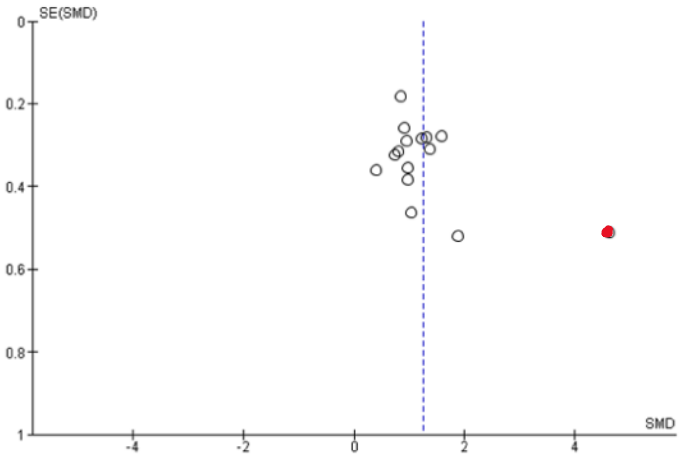


Figure 33. Funnel plot that examines possible publication bias of the NM-MRI studies.

6.1.3.5.2 Free water

The funnel plots appeared to be asymmetric when visually inspected as per Figure 34. However, it was not considered a major publication bias as only a few studies were included, especially since the free-water method was only introduced recently for PD diagnosis.

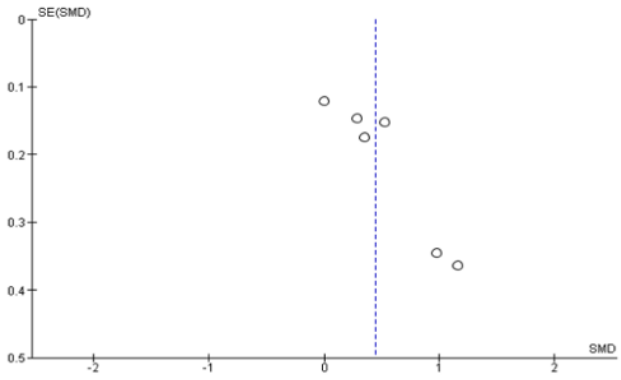


Figure 34. Funnel plot that examines possible publication bias of the Free-water studies.

6.1.4 Discussion

This meta-analysis quantitatively reviewed articles involving two promising MRI biomarkers of substantia nigra neurodegeneration in PD patients; NM-MRI and free water at 3T. Overall, NM nigral signal loss was consistently detected with a large sample population of over 2000, evidenced by a large disease effect size in the whole SN and posterior SN and high diagnostic accuracy. Nevertheless, moderate to high heterogeneity was also observed. Free water on the other hand was consistently elevated in the SN but data pooling could not be done due to very high heterogeneity for posterior and whole SN.

6.1.4.1 NM-MRI

Results obtained demonstrated consistent NM signal loss in PD patients across cohorts, acquisition protocols, and processing methods with a large pooled disease effect size. The SMD conformed to previous evidence in post-mortem studies, implying the role of NM loss in the pathogenesis of PD (Fearnley & Lees, 1991; Kitao et al., 2013).

This study also found diagnostic ability of Parkinson's with excellent sensitivity of 91% (95% CI: 84% - 94%) and very good specificity, 83% (95% CI: 79% - 87%) (eight studies). This is in line with two recent meta-analyses; a ten pooled studies showed sensitivity and specificity of 82% (95% CI: 0.74–0.87) and 82% (95% CI: 0.73–0.89) respectively (X. Wang et al., 2019) and a latest twelve studies meta-analysis found a pooled sensitivity of 89% (95% CI: 86–92%) and a pooled specificity of 83% (95% CI: 76–88%) when discriminating PD from healthy controls (Cho, Bae, Kim, Kim, Baik, Sunwoo, Choi, & Kim, 2021). The latest study by Cho, Bae, Kim, Kim, Baik, Sunwoo, Choi, & Kim, 2021 also included studies published on NM-MRI in LC.

Moderate to high heterogeneity (I^2 = 82%) was observed but subgroup analysis demonstrated significant subgroup difference between three metrics; contrast ratio, 3D volumetric, and 2D volumetric. A significantly strong disease effect size remained for all three subgroups, in addition to very low heterogeneity of 0% observed in contrast ratio and 2D volumetric subgroups. Hence, reflecting the 3D volumetric approach as a major contributor to the high heterogeneity obtained. Cho, Bae, Kim, Kim, Baik, Sunwoo, Choi, & Kim, 2021 in their meta-regression added that apart

from SN intensity measurement measure, heterogeneity also associated to slice thickness and segmentation methods.

Another interesting finding was that neuronal vulnerability in posterior SN had larger SMD (1.10, 95% CI: 0.71-1.48) vs. anterior SN (0.74, 95% CI, 0.47-1.01), consistent with prior pathologic studies (Damier et al., 1999; German et al., 1989; Gibb & Lees, 1991) of selective vulnerability to loss of melanized dopamine neurons in Parkinson's (Fearnley & Lees, 1991; Hodaie et al., 2007; Kordower et al., 2013) in the lateral-ventral area of the SN.

6.1.4.2 Free water

This is the first meta-analysis that systematically examined an advanced diffusion model that detects free-water increase in PD compared to healthy control (E Ofori et al., 2015). Free-water levels increased consistently in PD patients, which was reflected in the significant positive disease effect size. Nevertheless, high heterogeneity detected in whole SN and posterior SN suggested that findings may not be suitable to be pooled for meta-analysis.

Another limitation of the meta-analysis is that only a few studies relevant to free water diffusion MRI were available. Furthermore, the findings could be exposed to a bias of hidden overlapping populations since all the free water studies were from two research groups (Burciu et al., 2017; E Ofori et al., 2015). Pooled diagnostic performance could not be determined as only one study reported diagnostic accuracy (Planetta et al., 2016).

Despite the unexplainable heterogeneity in the reviewed studies, the pattern of free-water increase in the SN of PD patients (E Ofori et al., 2015) are thought to reflect neurodegenerative processes regarding local atrophy and inflammation in PD, not specific pathologic processes. Its clinical potential is further supported by two studies assessing free water as a PD progression marker. One study found free water increases in the posterior SN in de novo Parkinson's but remained constant in healthy controls (Burciu et al., 2017). The other large study involving more advanced disease observed elevated free water in the patient's anterior SN (Guttuso et al., 2018).

6.1.5 Limitation and future recommendation

There were several limitations regarding the scope and findings of this study. One of the main limitations was the lack of validation against the pathological diagnosis of PD, essentially the gold standard. Currently, DaTSCAN is the only diagnostic imaging test which has been clinically accepted but is yet to be validated against the post-mortem diagnosis. It has however been validated in clinical diagnosis with an accuracy of 84% in early PD and 98% at later stages (De La Fuente-Fernández, 2012). Meanwhile, clinical diagnosis using UK Brain Bank Criteria showed a pooled accuracy of only 87.2% compared to the post-mortem diagnosis in the latest meta-analysis (Rizzo et al., 2016).

This study focused on case-control which only included patients with a confirmed diagnosis according to UK Brain Bank Criteria or MDS Parkinson

Criteria. Other important populations such as patients with uncertain parkinsonism or those in preclinical manifestation were not included.

Few data with significant heterogeneity may reduce the statistical power of this meta-analysis, thus requiring justification from more studies. Subgroup analysis investigating intensity measurement metrics showed significant subgroup differences. A few other factors which were not considered in this study might have contributed to heterogeneity. Examples include grouping the populations based on severity, between-scanner, MRI protocol, post-processing differences, and SN region-of-interest (ROI) selection. Hence, suggesting standardisation and harmonisation through multi-centre studies. To date, our two NM-MRI studies (S. T. Schwarz et al., 2017; Xing et al., 2022) aimed to normalise NM-MRI signals across different imaging platforms. Meanwhile, four of the eight free-water diffusion studies were recruited from several platforms (Burciu et al., 2017; E Ofori et al., 2015; Edward Ofori et al., 2017; Yang et al., 2019). The high heterogeneity observed does require standardisation and harmonisation methods.

Finally, this meta-analysis gathered and compared data from different cohorts. So far, no study comparing NM-MRI and free-water imaging within the same cohort has been published.

6.1.6 Conclusions

Findings from this study confirmed the potential of NM- and free water diffusion MRI at 3T in diagnosing PD based on the pooling of previously published studies. Heterogeneity analyses highlighted the need for standardisation and harmonisation across many NM-MRI protocols and

cohorts. The following chapter presents a prospective case-control study using multiple protocols to investigate the harmonisation and standardisation techniques highlighted in this study. The next sub-chapter also discusses queries surrounding the potential capability of NM-MRI, such as disease progression, disease severity, and other types of parkinsonism.

6.2 Systematic Review of NM-weighted MRI in Parkinson disease and parkinsonian disorders.

6.2.1 Introduction

Functional imaging using DaTSCAN has been proven relevant in clinically uncertain parkinsonism. Meta-analysis in the previous subchapter suggested that NM-MRI can portray and quantify the depigmentation of NM loss in PD patients. Therefore, this systematic review aimed to further extend the study by synthesising published evidence of NM-MRI. Specifically, this study was designed to (1) detect and quantify pigmented NM loss in SN and (2) explore the potential application of NM-MRI in monitoring disease progression, classifying PD motor subtypes, and characterising PD from atypical parkinsonian disorder. Hence, this study systematically retrieved available data to assess diagnostic value and identify needs for future NM-MRI research.

6.2.2 Methods

The systematic review was conducted in accordance with the 'preferred reporting items for systematic reviews and meta-analysis' PRISMA statement (Moher et al., 2015). Similar to the former chapter, the term dorsal SNpc and ventral SNpc were exchanged for anterior SNpc and posterior SNpc or lateral SNpc and medial SNpc.

6.2.2.1 Literature Search Strategy

Literature related to NM, Parkinson's disease and MRI dating up to 30th March 2020 was searched within four databases: MEDLINE, PubMed,

EMBASE, Web of science. Keywords used were “neuromelanin” or “melanin” and “parkinsonism”, “Parkinson disease”, “Parkinson’s disease” and “magnetic resonance imaging” or “MRI” or “MR imaging”, respectively.

Based on the keywords, titles and abstracts of identified publications were screened. Following an exhaustive search of literature contents, articles were selected according to pre-defined inclusion criteria:

1. use of dedicated neuromelanin scan at 3T
2. Studies differentiating PD subtypes OR
3. Studies comparing PD with atypical parkinsonian disorders OR
4. Studies investigating PD staging and severity OR
5. Studies examining the PD progression OR
6. Review articles, conference abstracts, and qualitative studies were excluded.

The literature search was conducted by two researchers, me and Dr Saadnah Naidu. In the case of a dispute, both researchers discussed and reached an agreement with the first supervisor, Prof DP Auer.

6.2.2.2 Data Extraction

The following data were extracted: demographic data, including age, sex, and sample size; clinical data of the PD patients, details of the MR acquisition, segmentation method, and brief outcome. Data extraction was also performed by the author of this thesis and Rebecca Stevenson. Any disagreement during the review process was resolved by consensus.

6.2.3 Results

6.2.3.1 Literature Search Results

Initial search using the keywords as described in the method section returned a total of 459 articles. Subsequent screening of titles and abstracts reduced the number to 26. Thereafter, 8 articles were excluded due to two being conference abstracts, and six being review articles. Following an exhaustive examination of the contents, 5 full-text articles were excluded based on the aforementioned selection criteria. Fourteen articles were finally selected. No articles were found outside the online databases (i.e., in the reference list of the included studies).

Fourteen articles were selected for qualitative analysis. The articles were categorised into three groups to elucidate the utilisation of NM-MRI in discriminating motor subtypes, detecting early PD, monitoring disease progression, and differentiating PD from atypical parkinsonian disorder. The processes were simplified as illustrated in the PRISM flowchart below (Figure 35).

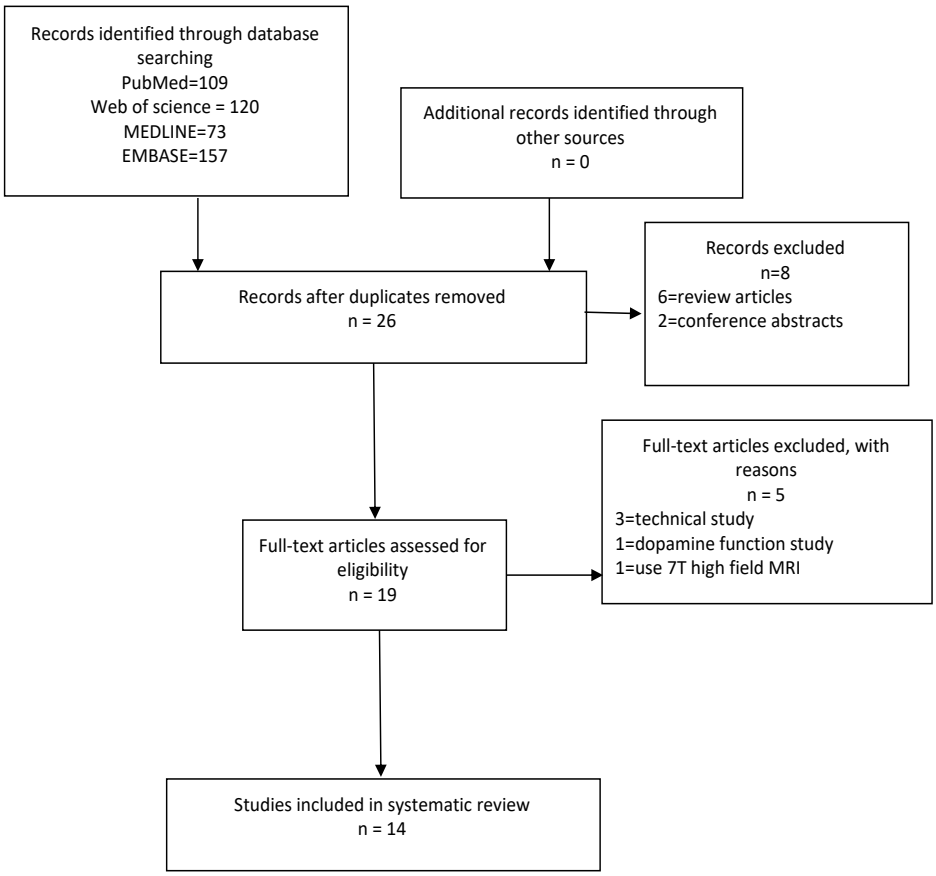


Figure 35. Flowchart of the processes involved in systematic review of the NM-MRI in parkinsonism following PRISM flowchart guidelines.

6.2.3.2 Main Results

Research for biomarkers in PD also included other domains like prognostic markers, staging markers, and theragnostic markers (Cova & Priori, 2018). This study aimed to examine whether NM-MRI would be a promising method to detect early disease, monitor disease progression, classify PD motor subtypes, or distinguish PD with atypical parkinsonian disorders. After an extensive screening of the full-text articles, fourteen studies were included.

6.2.3.2.1 Classify PD Motor Subtyping

Early clinical features of PD are highly heterogeneous. Patients are initially classified into two subtypes based on their motor and non-motor features. These subtypes are then characterised by severity of motor complications and non-dopaminergic features, reflecting complex interactions of the disease mechanism (van Rooden et al., 2011). The complex interaction of both subtypes is associated with different disease progression and neuropathological processes (van der Heeden et al., 2016). The first motor subtype is further classified into the tremor dominant (TD) or postural gait difficulty (PIGD) subtype based on the most pronounced motor symptoms. Patients with TD exhibit tremors as the dominant symptom while those with PIGD manifest balance and gait disturbances. This subtype categorisation is of utmost importance, especially during the early detection of PD to ensure correct interventions and tracking progression of the disease. Differentiating TD from PIGD is currently based on sub-scores of the Unified Parkinson's Disease Rating Scale (UPDRS) (Stebbins et al., 2013). The UPDRS is scored by clinicians, thus it is a skill and knowledge-dependent method besides being prone to human error. Alternatively, NM-MRI may offer an objective subtype classification.

For this domain, only one published article investigated the utility of NM-MRI to classify both subtypes (Xiang et al., 2017). The study involved 14 PIGD patients, nine TD patients, and 20 healthy controls. Upon image processing and analysis, it was found that the contrast ratio of medial SNpc was significantly reduced in PIGD compared to TD, suggesting medial SNpc in

PIGD degenerates more rapidly than in TD. The diagnostic accuracy however was moderately high, 81% for ipsilateral (sensitivity of 71.4% and specificity of 77.8%) and 77% for contralateral (sensitivity of 78.6% and specificity of 66.7%) depending on the clinically affected side. Results were consistent with previous neuropathological findings that revealed more severe SN neuronal losses and faster progression in the PIGD patients compared to TD patients (Burn et al., 2006; Rossi et al., 2010). Since this study involved a small sample size, further studies with a larger sample size are required to elucidate the two subtypes.

6.2.3.2.2 Detecting premotor Parkinson Disease

PD is detected during clinical diagnosis after substantial motor features had manifested. It is highly beneficial to recognise the loss of melanised cells during the development phase for early detection and intervention. It is also equally important to detect asymptomatic populations at risk of PD, in whom neurodegeneration is expected to begin (preclinical phase). The difficulty in identifying early diagnostic criteria for PD stems from the fact that there is currently no objective biomarker to predict the onset of the disease.

Non-motor symptoms may appear during the prodromal stage of the disease and may reflect degeneration in other areas before triggering the loss of dopaminergic neurons. Some of the non-motor symptoms may include olfactory dysfunction, rapid eye movement (REM) sleep behaviour disorder (RBD), autonomic dysfunction (such as constipation) and depression (D Berg et al., 2013).

One of the most important clinical signs of early PD is REM sleep behaviour disorder. RBD has the highest risk among the numerous premotor symptoms in developing PD. A longitudinal study revealed that patients with RBD will subsequently develop PD or dementia with Lewy bodies (DLB) with a median duration of 7 years (Iranzo et al., 2014).

The first study NM-sensitive imaging study investigated idiopathic RBD against healthy participants scanned with several MR sequences including NM-weighted and DTI (Nadya Pyatigorskaya et al., 2017). A significant decrease in both metrics and fractional anisotropy (FA) in the SN ROI of RBD patients was observed during manual measurement of pigmented SN neurons' signal intensity and volume compared to their counterparts. Discrimination of 21 RBD with an equal number of normal subjects yielded an accuracy of 0.86 for volume, 0.79 for signal intensity, and 0.77 for FA, suggesting volume had the highest differentiating abilities among all. Nevertheless, an excellent diagnostic accuracy of 0.92 was achieved when all three biomarkers were combined.

In summary, NM-MRI may be able to detect one of the main signs of the prodromal stage of PD. However, no link was found between the nigral degeneration with the other premotor signs, such as olfactory dysfunction, autonomic dysfunction (i.e., constipation) and depression. Therefore, NM-MRI may not apply to those signs.

6.2.3.2.3 Distinguish PD with Atypical Parkinsonian Disorders

Many atypical parkinsonian disorders such as multiple system atrophy (MSA), progressive supranuclear palsy (PSP), and corticobasal degeneration

(CBD) share similar neurological symptoms with PD, but they are pathologically distinct neurodegenerative disorders. Besides, progressive nigral cell loss exists as a pathology of these disorders. Consequently, misdiagnosis often occurs, especially in the early stages of the disease when the symptoms are still underdeveloped. NM-sensitive MRI newly emerged as a promising tool for differentiating PD and atypical parkinsonian disorders. Four studies explored dopaminergic neurodegeneration in PD and atypical parkinsonian disorders.

Kashihara et al. (2011) recruited a number of patients suffering from MSA (n = 28), PSP (n = 11), CBD (n = 10), PD (n = 80), and spinocerebellar ataxia SCA (n= 9). The aforementioned patients were compared with healthy control (n = 54) who suffered from a minor stroke, with no difference in disease duration among them (Kashihara et al., 2011a). Manual tracing of SNpc volume demonstrated a significant reduction in patients with MSA, PSP, and CBD but no reduction in SCA when compared to healthy control. The highest SNpc volume loss was observed in PD and PSP, followed by MSA and CBD, but no statistical difference was detected among groups.

Ohtsuka et al. (2014) later analysed regional selectivity of the central, medial, and lateral SNpc in the early stage of parkinsonism (i.e., PD, MSA with pre-dominant Parkinsonism (MSA-P), and PSPS). It was found that within lateral SNpc, the contrast ratio of PD and MSA-P were significantly lower than PSP. While in central SNpc, the contrast ratio of PD was significantly lower than PSPS. In terms of diagnostic performance, they compared the NM-MRI outcome with the iodine-123 meta-iodobenzylguanidine scan (MIBG).

Resultantly, comparable performances were found for PD versus PSPS of 0.84 NM-MRI, 0.88 MIBG (early) versus 0.84 MIBG (late) in the lateral area with better sensitivity (86.7% vs 76.7% vs 76.7%, respectively) and similar specificity (76.9% vs 76.7% vs 76.9%). Nonetheless, findings distinguishing PD and MSA-P were lower (0.5) compared to MIBG (0.78 for early and late phases) (C Ohtsuka et al., 2014).

Another study was conducted using NMMR imaging and midbrain volumetry to assess the relationship between pigmented neuronal degeneration and midbrain atrophy differentiation in PSP and PD (Taniguchi et al., 2018). A total of 45 volunteers were divided into 3 groups consisting of 11 PSP patients, 24 PD patients, and 10 healthy controls. Results showed that melanised SNpc areas and midbrain volume in PSP were smaller than in the PD and controls. Besides, the subregional analysis revealed that lateral and medial areas of SNpc were affected in PSP in contrast to lateral areas in PD. Best diagnostic accuracy discrimination of PD and PSP was achieved when both pigmented SNpc areas and midbrain volume were combined, with an area under the curve of 0.99, with a sensitivity and specificity of 100% and 96%, respectively. NM-weighted imaging solely produced an AUC of 0.85 with moderate sensitivity (73) but high specificity (100). While midbrain volumetry solely produced a higher AUC of 0.96 and superior sensitivity and specificity (91% and 96%). Combining NM-MRI and midbrain volumetry was proposed to best distinguish PSP from PD.

A recent study using multimodal MRI involved 11 PSP patients, 51 PD, and 26 healthy subjects (Nadya Pyatigorskaya et al., 2020). Comparisons of

three groups found a significant decrease in pigmented SN among PSP patients. This evidence supports NM-MRI as a promising biomarker to discriminate PD with atypical parkinsonian disorders, especially in PSP. Nevertheless, larger cohort studies are needed to explore the usability of NM-MRI before applying it to the population. Further research on MSA and CBD might also assist in understanding and differentiating SN neurodegeneration from PD.

6.2.3.2.4 The effect of severity

The UPDRS rating and Hoehn and Yahr Rating Scale (H&Y scale) are currently the standards for disease severity rating. Assessment is usually performed by neurological experts. A Higher UPDRS rating score and H&Y scale indicate higher severity of PD. Given the ability of NM-MRI to provide an objective estimation of nigral loss, six NM-MRI studies investigated nigral loss in relation to disease severity.

Studies investigating disease severity via NM-MRI revealed consistent results. Ohtsuka et al. (2013) compared three groups: early PD (i.e., Hoehn and Yahr stage I to II) with disease duration ranging from six months to three years, advanced PD (i.e., Hoehn and Yahr stage III to V) with disease duration between seven and 24 years, and healthy control. The contrast ratio for three circular ROIs was measured at the lateral, central, and medial parts of SNpc. No group differences were found between early and advanced PD but the advanced PD group depicted a lower SN contrast ratio compared to early PD for all areas (Chigumi Ohtsuka et al., 2013).

Schwarz et al. (2011) previously investigated neuronal degeneration in six early-stage PD patients with H&Y stage 1 or 1.5 and UPDRS 8 to 25, and four patients with late PD: HY stages 2 and 3, and UPDRS of 26 to 60. The ROI SNc area, including the contrast ratio of the SNc, was drawn and manually calculated. Similar to Ohtsuka and colleagues (2013), early and late-stage PD SNpc contrast ratio was also found to overlap. Nevertheless, area metric measurement conveyed a significant reduction of SNpc area between the late-stage and early-stage patients. The calculated SN area reduction of 47% in mild PD and 78% in moderate PD were consistent with progressive loss during the course of the disease: H&Y stage I by 40%, H&Y stage II by 75% (Rudow et al., 2008). A similar study was conducted by the same research team in 2017 with a higher number of participants. The measurement of SNpc based on the three MRI protocols indicated an association between SNpc volume loss and increased UPDRS score. Patients presented mild to moderate disease severity with UPDRS scores ranging from 2 to 60 and H&Y stages ranging from 1 to 3 (S. T. Schwarz et al., 2011).

A larger case-control NM-MRI study by Kashihara et al. (2011) was conducted on 80 PD patients and 54 healthy volunteers. SNpc were manually traced with a pen tablet pointing device to produce SNpc volume loss as per the H&Y stage. The mean volume loss obtained was 5.8% of the control value at Hoehn and Yahr stage I, 15.0% at stage II, 28.6% at stage III, 36.8% at stage IV, and 50.2% at stage V (Kashihara et al., 2011b).

On the other hand, Miyoshi et al. (2013) compared NM-MRI and MIBG to evaluate disease progression between early PD (n = 13), late PD (n =3 1), and Alzheimer's Disease. The early PD group comprised patients with Hoehn and Yahr stages I and II while the late PD group included stages III, IV, and V patients.

A reduced contrast ratio of lateral SNpc was observed in early and late PD. Comparisons of early and late PD also revealed a stage-dependent reduction in CR of medial SNpc (Miyoshi et al., 2013).

NM-MRI has been established as a potential marker for disease severity. Nevertheless, another study assessed NM-MRI alongside another Parkinson's biomarker, 123I-β-Carbomethoxy-3β-(4- iodophenyltropane) (CIT) SPECT (Hiroto Takahashi et al., 2019). The study involved 40 PD patients, where 22 were early-stage (H&Y I-II) while 18 were of advanced stage (H&Y III-V). Analysis disclosed a significantly higher SN SNR and striatal specific binding ratio (SBR) among the early PD group compared to the advanced group. Correlation analysis indicated SNR and SBR are negatively correlated with H&Y and UPDRS motor rating scores. Both portrayed high diagnostic performances, though DAT SPECT (0.89) was better than NM MRI (0.73) in distinguishing early-stage from advanced-stage PD.

From the above, it is apparent that NM imaging sequences hold great potential as a PD severity marker. The volumetric approach consistently revealed gradual SN signal intensity loss with disease severity, implying greater sensitivity in the volumetric approach, more than in the contrast ratio approach. Some studies evaluating the contrast ratio approach did report overlapping outcomes. However, those studies consistently highlighted a reduction in SN contrast ratio at an early stage of PD and later stages of PD.

6.2.3.2.5 Monitor disease progression

PD symptoms may change as the disease slowly progresses. There is currently no method to stop or delay disease progression in PD. Likewise, no biomarker is capable of detecting brain changes and monitoring the progression of the PD (Athauda & Foltynie, 2015). Researchers investigated

how NM-MRI could assist in monitoring progressive loss of dopaminergic neuronal loss *in-vivo*. Two studies explored NM-MRI's potential for monitoring neuropathological alteration over time in PD patients.

A preliminary longitudinal study by Matsuura et al. (2016) involving 14 PD patients found that SNpc total area and CR of the patients were significantly smaller during the follow-up scan compared to the initial scan, and were negatively correlated with disease duration (Matsuura et al., 2016). This unique outcome led to the implementation of a cross-sectional study by Fabbri (2017), who scanned 13 late-stage PD patients (LSPD) with a median disease duration of 14 years, 12 de novo PD patients, 10 PD patients with a 2 to 5-year disease duration, and 10 healthy subjects. Group comparison displayed a significantly smaller NM area in the LSPD group than in the de novo group. Though the comparison with the 2 to 5-year PD group was not statistically significant, tendency for the pigmented neuron area to decrease was observed (Fabbri et al., 2017).

6.2.4 Limitation and future recommendation

Some limitations were relevant to the scope of this systematic review. Most of the studies discussed comprised a small sample size. Therefore, studies with larger cohorts are required to properly determine the potential of NM-MRI for larger populations.

6.2.5 Conclusions

This systematic review corroborated pigmented SN neuronal loss in PD and parkinsonism. All studies yielded consistent outcome for the domains discussed, which includes identifying PD subtyping, distinguishing PD with

atypical parkinsonism, detecting severity and monitoring progression. These events form the basis of a potential biomarker for PD and neurodegenerative parkinsonism. Several aspects in this systematic review were not covered within this thesis, but it is crucial for future related research or ongoing studies. More research is needed with a larger sample size to increase statistical power.

7. A case control single centre sub study of PaMIR of neuromelanin-weighted imaging in detecting Parkinson disease

7.1 Introduction

Meta-analysis of NM-MRI in the previous chapter (Chapter 6.1) described case-control studies in NM MR imaging following the recent discovery of the protocol. Collectively, all studies denoted strong positive disease size effect although moderate to high heterogeneity was found. As standardisation and harmonisation of method was proposed, this study recruited a moderately large number of patients with Parkinson and healthy controls to investigate the normalisation of two NM-sensitive imaging sequences to discriminate healthy control and PD patients. Diagnostic performance and sensitivity towards SNpc sub regional difference was also investigated.

7.2 Materials and methods

This entire project was part of the ongoing multi-centre Parkinson imaging research of PaMIR (REC 14/EM/0061; refer Appendix 2 for details of the project). This work utilised and analysed datasets from our MR facility only and majority of the data were also used in Chapter 4.3. As described in detail in Chapter 4.3, the author of this thesis was involved in data organisation, data acquisition, image analysis, and statistical analysis. A manuscript on this multi-centre PaMIR project which covered all data and analysis from other centres has been published recently (Xing et al., 2022). The author of this thesis also contributed as the second author. While the published manuscript only reported the outcome of the automated

template-based method, this single-centre study analysed the data using both automated and semiautomated image processing methods.

7.2.1 Participants and data acquisition

A total of 131 volunteers scanned using two NM MR sequences, Sequence 1 (n = 85) and Sequence 2 (n = 46), were recruited in this single centre data. In contrast to Chapter 4.3, this study first selected all participants with available data of Sequence 1 to prioritise the optimised sequence with faster acquisition and then recruited those with only Sequence 2 without Sequence 1 data to maximise available data and to attempt normalisation. All the data were first inspected for quality control before processing, leading to the removal of six participants' data.

The semiautomated method was previously analysed in 2018, so participants that were scanned afterwards were not included in this study. Meanwhile, the analysis for the automated method was performed recently during the writing stage of this thesis. The data comprised 18 additional participants. Demographic data of the subjects are presented in Table 13 below.

Table 13. Demographic data for healthy controls and Parkinson patients.

	Automatic		**Semiautomatic**	
	Healthy	PD	Healthy	PD
No. (female/male)	60 (33/27)	65 (26/39)	45 (31/14)	62 (24/38)
Age, years (mean±SD)	69.46 ± 8.54	67.0 ± 8.8	68.11 ± 8.9	67.1 ± 9.0
Age range (min-max)	46.43-83.72	47.1-85.29	46.43-83.72	47.1-85.29
Hoehn & Yahr scores		1.75 ± 0.84		1.77 ± 0.82
Total UPDRS scores		40.59 ± 22.30		41.9 ± 23.3
UPDRS III		22.11 ± 13.05		23.0 ± 14.0
Disease duration		2.42 ± 1.72		2.4 ± 1.7

For the automatic method of analysis, a final total of 125 volunteers comprising 65 PD patients (39 males, 26 females, mean ± STD age: 67.0 ± 8.83 years, range = 47.1-85.29) and 60 controls (27 males, 33 females, mean ± STD age: 69.46 ± 8.54 years, range = 46.43-83.72) were included in this study for further processing and analysis.

There was no significant difference in the age distribution (P = 0.116) between Parkinson's patients and healthy controls. PD patients had a mild to moderate disease severity according to the Hoehn & Yahr scores, ranging from 1 to 4 (n = 64, mean ± STD: 1.75± 0.836). Total UPDRS scores ranged from 9.5 to 124 with mean UPDRS of 40.593 ± 22.30 (n = 59). The UPDRS III had the lowest and maximum scores of 3 and 77, respectively (mean ± STD; 22.11 ± 13.052, n = 63). The Mean disease duration was 2.418 ± 1.719 years (range: 0.173 – 6.847 years).

For the semiautomated method, the analysis included 107 participants with 62 PD patients (38 males, 24 females, mean ± STD age: 67.1 ± 9.0 years, range = 47.1-85.29) and 45 controls (14 males, 31 females, mean ± STD age: 68.11 ± 8.9 years, range = 46.43-83.72) were included. No significant difference in age (P = 0.547) was found between PD patients and healthy controls. PD patients had mild to moderate disease severity according to the Hoehn & Yahr scores, ranging from 1 to 4 (n = 62, mean ± STD: 1.77± 0.82). Total UPDRS scores ranged from 9.5 to 124 with a mean UPDRS of 41.9 ± 23.3 (n = 52). The UPDRS III recorded the lowest and maximum scores of 4 and 77, respectively (mean ± STD: 23.0 ± 14.0, n = 52). The mean disease duration was 2.4±1.7 years (range: 0.2 – 6.8 years).

7.2.2 Image analysis

Image processing was conducted via the two methods outlined in Chapter 3. As per section 4.3, Sequence 1 produced a higher contrast-to-background ratio compared to Sequence 2. CBR for Sequence 1 and 2 were normalised by averaging CBR of healthy on both sequences and multiplying each individual in Sequence 2 with the ratio of Sequence 1 to Sequence 2.

Group comparisons between patients and controls were performed using t-test analysis. Non-parametric tests were used for group comparisons in the case of non-normal distribution. Values were specified as mean ± SD unless stated otherwise; significance was defined at $P < 0.05$. Comparisons between gender (male vs female) and hemispheres (left vs right) of CBR and volume in SN were also performed either via univariate or multivariate analyses of variance with least-squares differences or t-tests. Disease severity effect and duration on SN intensity were analysed using linear regression controlling for age.

7.3 Results

7.3.1 PD Disease effect

Quantitatively, both measurement methods displayed a marked decrease in SN signal intensity in PD compared to healthy controls. Dorsal SN normalised CBR for the automated method were 0.163 ± 0.020 and 0.149 ± 0.033 for healthy controls and PD, respectively. While for semiautomated method, dorsal SN normalised CBR were 0.258 ± 0.033 and 0.231 ± 0.036 for

healthy controls and PD, respectively. For ventral SN, the template-based method reported normalised CBR of 0.093 ± 0.026 for healthy controls and 0.057 ± 0.029 for PD. Meanwhile, the manual drawing method reported normalised CBR for the healthy cohort and PD patients were 0.204 ± 0.041 and 0.158 ± 0.023, respectively. All the aforementioned normalised CBR were significant at $P < 0.0001$. Boxplot in Figure 36 visualised group comparison of both processing methods according to subregions. Results obtained are also summarised in Table 14 below:

Table 14. Overview of SNpc subregions signal intensity measured in the normalised contrast-to-background ratio (CBR) in healthy control and Parkinson patient analysed by automated template-based and semiautomated manual drawing method. P-values indicate statistically significant comparisons between the mean values of both groups.

Method/ Subregion	Healthy control	Parkinson disease	P-value
Automated template-based			
Dorsal SNpc	0.163 ± 0.020	0.149 ± 0.033	< 0.0001
Ventral SNpc	0.093 ± 0.026	0.057 ± 0.029	< 0.0001
Semiautomated manual drawing			
Dorsal SNpc	0.258 ± 0.033	0.231 ± 0.036	< 0.0001
Ventral SNpc	0.204 ± 0.041	0.158 ± 0.023	< 0.0001

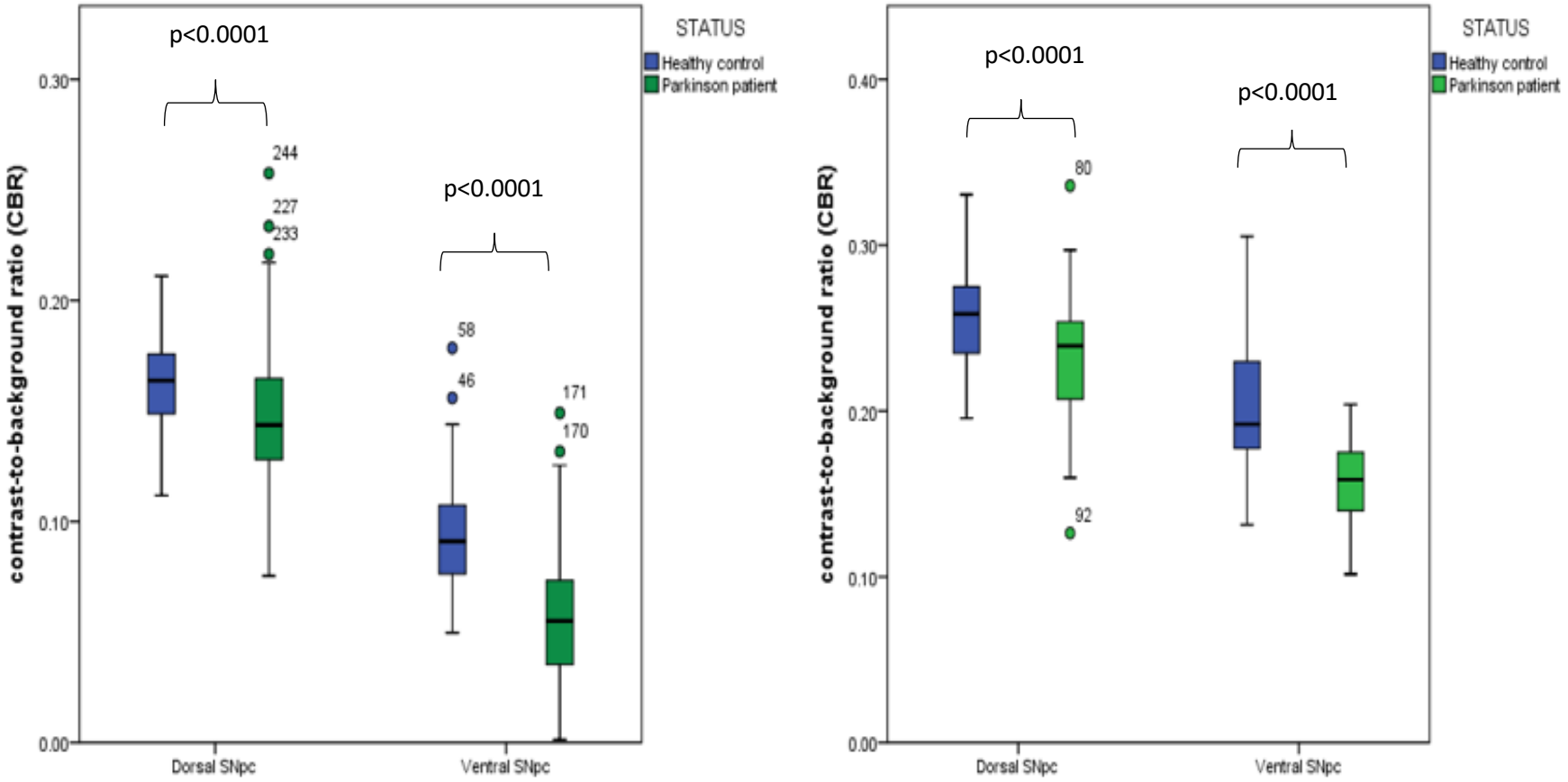


Figure 36. Box plot of CBR by subregion, dorsal and ventral SNpc in healthy controls vs Parkinson patients measured by automated and semiautomated methods.

7.3.2 Laterality and sex effects

No differences in laterality were found for both groups, HC and PD, for either measurement method. Among normal subjects, the template-based approach had balanced left and right SN CBR (left SNpc, 0.131±0.02; right SNpc, 0.125 ± 0.02; P = 0.096) while the drawing approach revealed bilateral CBRs (left, 0.224 ± 0.029; right, 0.238 ± 0.043; P = 0.157). Among the PD group, no differences were observed on both sides of SNpc in the template-based method (left, 0.104 ± 0.02; right, 0.10 ± 0.02; P = 0.366) and the drawing method (left SNpc, 0.192 ± 0.034; right SNpc, 0.198 ± 0.029; P = 0.363). Table 15 depicts the laterality effects.

Furthermore, no subregional differences in gender were observed within subgroups based on the two methods. Among normal subjects, the automated method found dorsal SNpc (male SNpc, 0.158 ± 0.02; females SNpc, 0.167 ± 0.02; P = 0.778) and ventral SNpc (male SNpc, 0.093 ± 0.03; females SNpc, 0.094 ± 0.03; P = 0.073) to be equal between genders. Likewise, for semiautomated method, dorsal SNpc (male SNpc, 0.257±0.03; females SNpc, 0.259 ± 0.03; P = 0.879) and ventral SNpc (male SNpc, 0.203±0.03; females SNpc, 0.204±0.04; P = 0.962), were equal between gender.

As for the PD group, both male and female dorsal SNpc (male SNpc, 0.152 ± 0.03; females SNpc, 0.144 ± 0.03; P = 0.228) and ventral SNpc (male SNpc, 0.055 ± 0.03; females SNpc, 0.059 ± 0.02; P = 0.296) displayed no difference in the automated technique. The semiautomated technique also

displayed no difference in dorsal SNpc (male SNpc, 0.229 ± 0.04; females SNpc, 0.235 ± 0.02; P = 0.408) and ventral SNpc (male SNpc, 0.155 ± 0.02; females SNpc, 0.162 ± 0.02; P = 0.267) between gender. Effects of gender are summarised in Table 16.

Table 15. Summary of laterality effects in healthy control and Parkinson patients for automated and semiautomated processing methods.

Method	**Healthy control**		P-value	**Parkinson disease**		P-value
	Left SNpc	Right SNpc		Left SNpc	Right SNpc	
Automated template-based	0.131±0.02	0.125±0.02	0.096	0.104±0.02	0.10±0.02	0.366
Semiautomated manual drawing	0.224±0.029	0.238±0.043	0.157	0.192±0.034	0.198±0.029	0.363

Table 16. Summary of sex effects in healthy control and Parkinson patients for automated and semiautomated processing methods.

Method/ Subregion	**Healthy control**		P-value	**Parkinson disease**		P-value
	Male	Female		Male	Female	
Automated template-based						
Dorsal SNpc	0.158±0.02	0.167±0.02	0.778	0.152±0.03	0.144±0.03	0.228
Ventral SNpc	0.093±0.03	0.094±0.03	0.073	0.055±0.03	0.059±0.02	0.296
Semiautomated manual drawing						
Dorsal SNpc	0.257±0.03	0.259±0.03	0.879	0.229±0.04	0.235±0.02	0.408
Ventral SNpc	0.203±0.03	0.204±0.04	0.962	0.155±0.02	0.162±0.02	0.267

7.3.3 Diagnostic performance

Accuracy among this cohort of NM-MRI study was also examined to diagnose Parkinson from the control group by performing a receiver operator

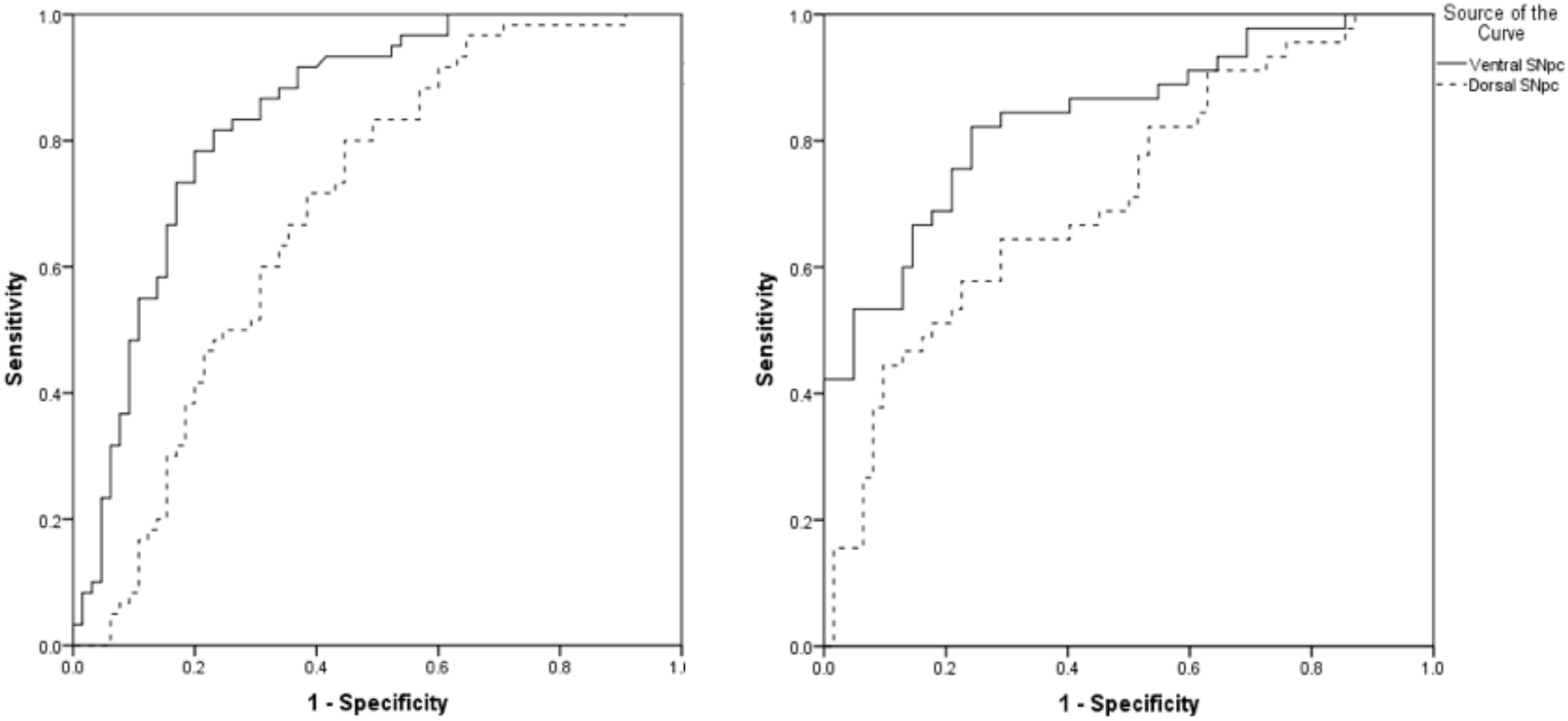
characteristics (ROC) analysis of the SN signal and subregions. The diagnostic performance is presented in Table 17.

Table 17. Summary of diagnostic performance (AUC) including optimal cut-off and sensitivity and specificity to diagnose Parkinson disease on two CBR measurement methods.

Method/ Subregion	AUC ± Standard error	Cut-off	Sensitivity (%)	Specificity (%)
Automated template-based				
Dorsal SNpc	0.691±0.048	0.152	0.717	0.615
Ventral SNpc	0.842± 0.036	0.074	0.817	0.769
Semiautomated manual drawing				
Dorsal SNpc	0.715±0.050	0.243	0.667	0.597
Ventral SNpc	0.837±0.400	0.175	0.822	0.758

The area under the curve demonstrated that ventral SNpc recorded the best diagnostic accuracy in distinguishing between healthy control and PD patients, with values of 0.842 and 0.837 for automated and semiautomated processing methods, respectively (the ROC plot is illustrated below in Figure 37). The optimal cut off for each subregion was 0.152 and 0.243 for dorsal SNpc along with 0.074 and 0.175 respectively for ventral SNpc.

Sensitivity and specificity were also high among the regions with dorsal SNpc generating sensitivities of 0.717 and 0.667, and specificities of 0.615 and 0.597 for the template-based and manual drawing techniques, respectively. For ventral SNpc, the automated approach yielded a sensitivity of 0.817 and specificity of 0.769. Meanwhile, the semi-automated approach recorded a sensitivity of 0.822 and specificity of 0.758.



Automated template-based method

Semiautomated drawing method

Figure 37. ROC plot of SNpc contrast-to-background ratio to diagnose Parkinson. The analysis compared diagnostic accuracy of the dorsal SNpc and ventral SNpc.

7.3.4 Disease progression and disease severity effect

Linear regression analysis of pigmented dopaminergic neurons signal with disease duration, UPDRS III and Hoehn and Yahr score yielded significant regression when controlled for age. Regressing out for age was considered as the previous chapter revealed NM age effects from adulthood to old age (Chapter 5). Hence, 92.5% of the data over the age of 63 were included in this study. For the automated technique, disease duration and UPDRS III revealed significant regression in ventral SNpc (Figure 38). The regression equation for disease duration was $F(2,61) = 3.495$, $P = 0.037$ with an $R2$ of 0.103. While regression equation for UPDRS III was $F(2,60) = 3.263$, $P = 0.045$ with an $R2$ of 0.098. However, no significant regression was found for Hoehn and Yahr with $P = 0.055$ for ventral SNpc and $P = 0.335$ for dorsal SNpc. Besides, dorsal SNpc reflected no regression in disease duration ($P = 0.25$) and UPDRS III ($P = 0.14$).

Interestingly, for semiautomated technique, no significant regression was found between any of the measured SNpc and disease duration (ventral, $P = 0.087$; dorsal, $P = 0.485$), Hoehn and Yahr (dorsal, $P = 0.996$; ventral, $P = 0.657$) and motor UPDRS (dorsal, $P = 0.854$ and ventral, $P = 0.692$).

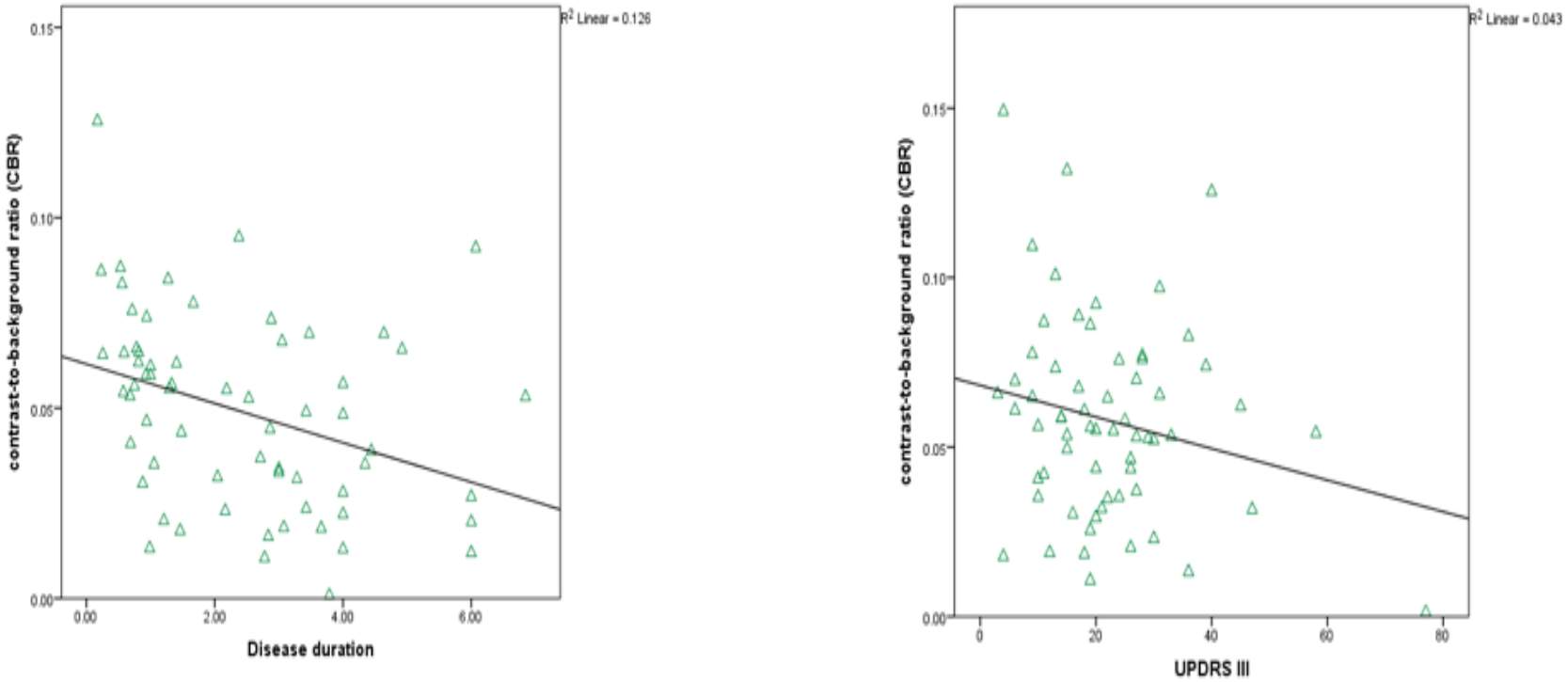


Figure 38. Scatter plot of disease duration and motor UPDRS vs dorsal SNpc CBR of the Parkinson patient with linear fit line illustrating significant linear regression with equations; disease duration, F (2,61) = 3.495, P=0.037 with an R^2 of 0.103 while UPDRS III, F (2,60) = 3.263, P=0.045 with an R^2 of 0.098.

7.4 Discussion

The research work showcased the standardisation method of two dedicated NM MRI sequences while describing the loss of pigmented neuron CBR in SN based on measures of two processing methods within a moderately large cohort with the potential to become a diagnostic biomarker for Parkinson's. Further validation of diagnostic performance demonstrated that patients with PD and healthy control were best distinguished by ventral SNpc, with high AUC for both processing methods (0.842 ± 0.036 vs 0.837 ± 0.400 for automatic and semiautomatic, respectively). Preferential signal loss in the ventral region is consistent with histological findings of PD induced neurodegeneration in the lateral ventral tier (Fearnley & Lees, 1991).
The utility of the NM-MRI to detect disease severity and progression effect was also assessed. Interestingly, it was found that the automated method had significant linear regression with disease duration and UPDRS III in ventral SNpc after regressed out the age. No linear regression was found for the semiautomated method neither at the dorsal SNpc nor ventral SNpc. The recently published longitudinal study (Gaurav et al., 2021; Xing et al., 2022) and previous studies in line described increased NM signal intensity loss with higher advancement of the disease, as measured through disease duration (Kashihara et al., 2011b), motor UPDRS (S. T. Schwarz et al., 2017), and Hoehn and Yahr stages (Schwarz et al., 2011). In contrast, previous observations from the use of an automated SNpc segmentation algorithm did not yield a correlation between substantia nigra volume, motor UPDRS and

disease duration (Castellanos et al., 2015). Another observation showed no difference in the contrast ratio between early and advanced Hoehn and Yahr stages (Chigumi Ohtsuka et al., 2013). This study justifies the inconsistencies in previous studies by attributing the finding to differences in imaging and measurement methods or technical differences, as discussed by Schwarz et al. (2017). The present study also highlighted that automated template-based methods were more sensitive to disease severity and progression than the semiautomated drawing method.

One major limitation of this study is a lack of gold standard post-mortem diagnosis. The study involved patients that were diagnosed based on UK Brain Bank Criteria, which may have a 15% tendency to be misclassified.

The unique challenge of MRI quantitative analysis due to its intensities are non-standardized and highly reliance on manufacturer, sequence type, and acquisition parameters (Simmons et al., 1994). The post-processing ratio-based normalisation method applied in this work are population-specific and prone to bias towards outlier. Standardized pre-processing intensity normalization method is proposed for future multi-sequence or multi-centric study (Carré et al., 2020).

Another limitation is that a few of the data were removed due to quality criteria assessment. Therefore, implying room for improvement to better develop the NM-MRI protocol, especially one with a fast acquisition that does not jeopardise image quality. This is important given that PD patients are highly susceptible to motion artefacts despite being medicated during scanning.

A potential barrier to clinical implementation is the time and energy-consuming nature of the semiautomated method, especially for large cohorts. Nevertheless, this limitation was subdued by the fast nature of the automated method though it is not fully automated.

7.5 Conclusion

In conclusion, this case-control study measured using normalised CBR of two NM-weighted sequences corroborated findings from a previously pooled study detailed in Chapter 6.1 by revealing PD-induced degeneration of the SNpc with excellent diagnostic performance. The automated method displayed better sensitivity in detecting disease severity and progression compared to the semiautomated method. More longitudinal study is proposed to further assess stage-dependent loss of NM concentration. Since this chapter recruited a cohort who had already been diagnosed with PD, there is also a need to explore NM-MRI application among a population of patients with clinically uncertain parkinsonism, which the next chapter (Chapter 8) aims to answer.

8. Can NM MRI detect nigral depigmentation in Clinically Uncertain Parkinsonism: a N3iPD sub study

8.1 Introduction

Clinically uncertain parkinsonian syndrome (CUPS) is a term used to define a group of patients suspected to have PD but present no full parkinsonism cardinal features like bradykinesia or other features like tremor and rigidity in the absence of bradykinesia. CUPS also includes patients with poor response to levodopa and portrays no disease progression (Catafau et al., 2004). It is estimated that one in five suspected Parkinson's patients fall under the CUPS category. Underlying presynaptic degeneration could explain the cause of the symptoms, determine the course of the disease, and enable planning for prognosis and therapy (Catafau & Tolosa, 2004).

DaTSCAN is a tool established to identify patients with CUPS. It is used for single photon emission computed tomography (SPECT) brain imaging to help identify functional dopaminergic neuron terminals loss. When injected, it binds to the dopamine transporter membrane located on presynaptic terminals of dopaminergic neurons in the striatum (caudate and putamen). Dopaminergic activity is then detected using the SPECT brain imaging (Booij et al., 2001).

DaTSCAN is currently the only imaging tool approved to diagnose CUPS. Since its approval for clinical usage, DaTSCAN which is mainly used to characterise non-degenerative and degenerative parkinsonism (Benamer et al., 2000) has assisted in optimising diagnostic accuracy and providing information for improved clinical management (Catafau & Tolosa, 2004).

A normal scan indicates a higher likelihood of non-degenerative parkinsonism such as drug-induced, psychogenic, vascular, essential tremor

or dystonic tremor (Bega et al., 2021). An abnormal scan however suggests nigrostriatal neurodegenerative parkinsonian disorder, such as PD, multiple system atrophy (MSA), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD) or Lewy body disease (Vlaar et al., 2007).

Nonetheless, its potential has been hindered by a few limitations. DaTSCAN alone cannot be used to diagnose PD or CUPS, it is not readily available worldwide, expensive due to shipping and transportation costs, and requires a high level of expertise in nuclear imaging to interpret the image. On the other hand, MRI is widely available, non-irradiated, and has an edge over DaTSCAN. Therefore, this study aims to explore whether NM-MRI could potentially characterise clinically uncertain PD patients in line with the clinically accepted standard, DaTSCAN. This project also attempts to assess whether CUPS is an intermediary phenotype as classified by NM MR imaging measured by CBR between healthy control and PD patients. This work is a sub-study of a multicentre project called nigrosomal iron imaging in Parkinson's disease (N3iPD) led by Prof Dorothee Auer.

8.2 Materials and methods

8.2.1 Participants

The N3iPD study recruited participants from three centres: the University of Nottingham, Imperial College London, and Royal Derby Hospital. However, this analysis was only confined to participants scanned at our centre. Study coordination, recruitment, governance, and data management were headed by Dr Saadnah Naidhu and assisted by Dr Yue Xing. The author

of this thesis contributed to managing logistic requirements, preparation of volunteers e.g., checking for MRI contraindication for data acquisition, as well as image processing and statistical analysis. The study was approved by the research ethical committee with reference: 16/EM/0229 (further details on the approval are attached in Appendix 4).

Participants' Eligibility criteria

The participants recruited were different from the rest of the studies reported in this thesis. Specifically, this study involved participants with suspected and/or clinical uncertainty of PD diagnosis. Therefore, specific inclusion and exclusion criteria were devised as listed below:

Inclusion criteria

- Provided informed consent
- Aged between 21 and 90 years
- Suspected of PD but clinically uncertain for a definite diagnosis such as:
    - symptoms manifestation partially fulfilled criteria for PD diagnosis according to UK Brain Bank diagnostic criteria
    - clinical features that were typically not associated with PD but with a different type of movement disorder
- referred for a DaTSCAN to enable clinical diagnosis as part of the National Health Service (NHS) or for this research study to investigate suspected parkinsonian-type disease

Exclusion criteria

- Presented contraindications to MRI as outlined in Chapter 3
- Had a significant head tremor
- Pregnant
- Medically unfit for MRI scan due to back pain, claustrophobia, acute sickness and may also include patients with signs of impaired temperature regulation such as an extremely high fever
- Presented contraindication to DaTSCAN due to severe allergy to iodide compounds or thyroid gland dysfunction

A total of 40 participants were recruited but four were excluded due to the unavailability of DaTSCAN diagnostic data. A final total of seven clinically uncertain PD patients with negative DaTSCAN (four males, three females) and 29 clinically uncertain Parkinson's with positive DaTSCAN (16 males, 13 females) were included. Their age ranged from 55 to 88 (mean ± SD; 70.71 ± 10.45) years and 54 to 80 (mean ± SD; 70.76 ± 7.57) years, respectively. DaTSCAN analysis was conducted by two researchers, Prof Paola Piccini with 20 years of experience and Dr Antonio Martin-Bastida with six years of experience.

8.2.2 Image Acquisition and processing

NM-weighted MRIs were acquired using Sequence 1 while DaTSCAN were acquired and diagnosed by local nuclear medicine specialists, Dr Jim

Birchall and Dr Alan Perkins. For this study, the diagnostic outcome was obtained as positive or negative DaTSCAN. NM-MRI images were processed utilising Methods 1 and 2. As the sequence and processing methods were identical to the previous case-control study of PaMIR (Chapter 7), the findings of this analysis were compared to the healthy control and PD patients. An example of a DaTSCAN image is presented in Figure 39.

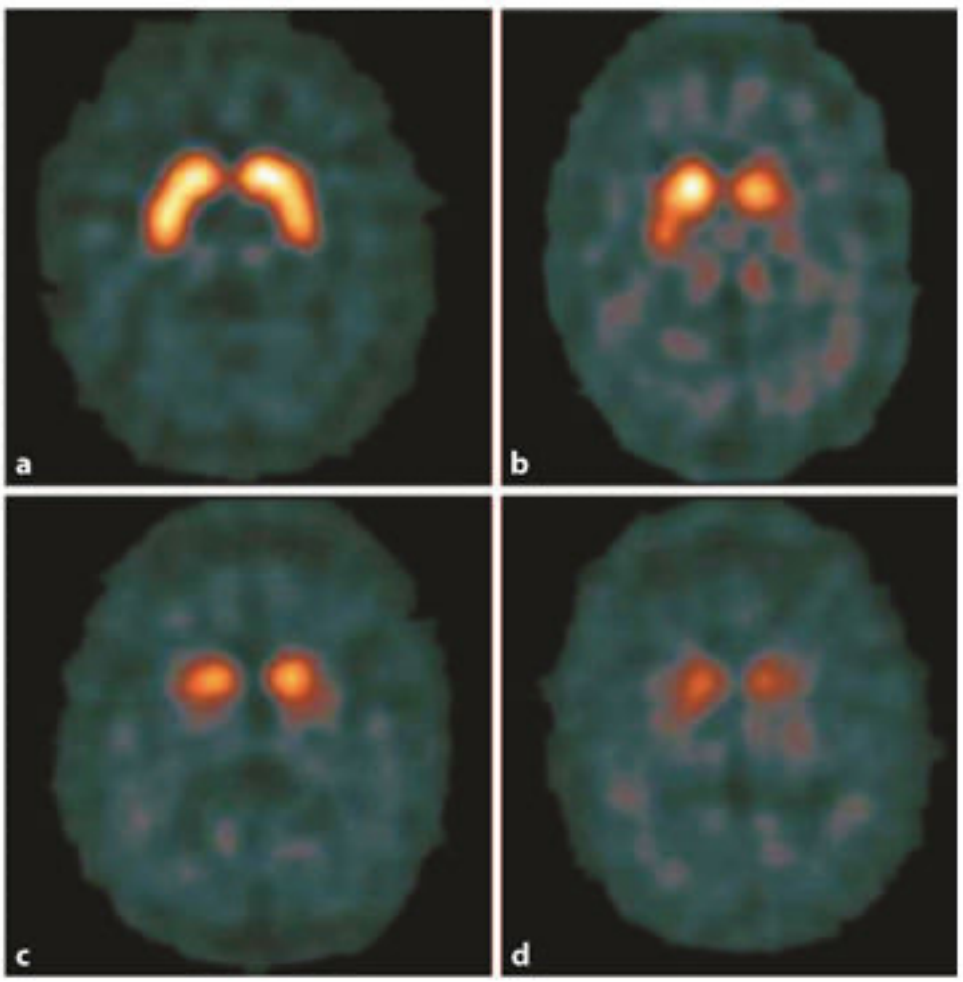


Figure 39 Normal and abnormal DaTSCAN images. a Normal DaTSCAN imaga. b–d Abnormal DaTSCAN images fall into at least one of the following three categories (all are considered abnormal): abnormal DaTSCAN image type 1(b), abnormal DaTSCAN image type 2 (c) or abnormal DaTSCAN image type 3 (d).(Kupsch et al., 2013)

The image was taken from the study conducted by Kupsch et al. (2013) entitled “Changes in clinical management and diagnosis following DaTscan SPECT imaging in patients with clinically uncertain parkinsonian syndromes: a 12-week follow-up study”. People with PD will typically have a smaller signal in a part of the brain called the striatum, where the ends of the dopamine neurons are meant to be. The image depicts a normal scan (a) with two distinct “comma” shaped features or comma-like or crescent shapes, indicating a healthy dopamine system. Next to the normal scan is an abnormal scan (b-d) with two “period” shaped features or period-like or oval shapes, or a combination of period and comma shapes, indicating a reduced uptake of DaTSCAN by an unhealthy dopamine system (Kupsch et al., 2013).

8.2.3 Statistical analysis

The data were checked for normality using the Shapiro Wilk test. All the data were normally distributed, hence parametric analysis: independent t-test was performed to compare the means between the positive and negative DaTSCAN, as well as between healthy control and PD patients (from Chapter 7). Meanwhile, a one-way Analysis of Variance (ANOVA) was conducted to detect significant difference between all the groups. This analysis was undertaken to explore whether clinically uncertain patients are intermediary phenotypes between PD patients and healthy controls. Upon detecting a significant difference between the groups, multiple comparisons were performed using the Bonferroni post-hoc test. A P-value of < 0.05 was considered significant.

8.3 Results

8.3.1 Comparisons of SN CBR between positive and negative DATSCAN using the automated method

This study aims to explore whether NM-MRI could potentially characterise clinically uncertain PD patients in line with the clinically accepted DaTSCAN. Table 18 depicts the SN CBR between the positive and negative DaTSCAN upon utilising the automated method. For both ventral and dorsal SNpc, the SN CBR was significantly lower in the positive DaTSCAN compared to the negative DaTSCAN (P = 0.001).

A significant decrease (P = 0.001) in SN CBR was also detected in the PD patients (0.057 ± 0.029) compared to the healthy controls (0.093 ± 0.026) in the ventral SNpc (Table 19). Likewise, the PD patients demonstrated a significant (P = 0.001) lower SN CBR (0.149 ± 0.033) compared to the healthy controls (0.163 ± 0.020) in the dorsal SNpc. (Outcomes from Chapter 7)

8.3.2 Comparisons of SN CBR between healthy controls, clinically uncertain patients, and PD patients using the automated method

Table 20 presents the comparisons between the healthy controls, clinically uncertain patients, and PD patients for the automated method. The initial one-way ANOVA revealed that the SN CBR was significantly different between the groups. Further multiple comparisons using the Bonferroni test indicated that SN CBR in the ventral SNpc was significantly higher in the healthy controls (0.093 ± 0.026) compared to the clinically uncertain patients with positive DaTSCAN (0.065 ± 0.030; P=0.001) and PD patients (0.057 ±

0.029; P=0.001), but similar to those of negative DaTSCAN (0.092 ± 0.024; P=1.00).

In dorsal SNpc, the automated approach also found a significantly higher CBR loss in the PD patients (0.149 ± 0.033) compared to the healthy controls (0.163 ± 0.019; P=0.014) and clinically uncertain patients with negative DaTSCAN (0.181 ± 0.033; P=0.014). In contrast, no significant difference was detected in the CBR loss between the PD and clinically uncertain patients with positive DaTSCAN (0.156±0.021; P=1.00).

8.3.3 Comparisons of SN CBR positive and negative DATSCAN using the semi-automated method

The semi-automated also found that the clinically uncertain patients with positive DaTSCAN recorded higher CBR loss compared to their negative counterparts (Table 21). Table 22 presents the SN CBR quantified by the semi-automated method in the healthy controls and PD patients in the ventral and dorsal SNpc. A significant CBR loss (P = 0.001) was observed in the PD patients in comparison to the healthy controls in the ventral (0.158 ± 0.023 vs 0.204 ± 0.040) and dorsal (0.231 ± 0.036 vs 0.258 ± 0.033) SNpc. (Outcomes from Chapter 7)

8.3.4 Comparisons of SN CBR between healthy controls, clinically uncertain patients, and PD patients using the semi-automated method

In the ventral SNpc, statistically significant differences were detected in the CBR loss between the groups based on the one-way ANOVA test. Multiple

comparisons using the Bonferroni post-hoc test are presented in Table 23. Specifically, in the ventral SNpc, the PD patients recorded a significantly higher CBR loss compared to the healthy controls (0.204 ± 0.040; P = 0.001) and clinically uncertain patients with negative DaTSCAN (0.195 ± 0.028; P = 0.021). The comparisons between PD and positive DaTSCAN patients (0.168 ± 0.031), and between healthy controls and clinically uncertain patients with negative DaTSCAN were not statistically significant (P = 0.782 vs P = 1.00).

In the dorsal SNpc, a significant CBR loss was also detected between the groups based on the one-way ANOVA test results. Multiple comparisons reflected that CBR loss in the dorsal SNpc was significantly higher in the PD patients than the healthy controls(P=0.001), but no significant difference was detected when compared to the clinically uncertain patients with negative (0.257 ± 0.038; P = 0.326) and positive DaTSCAN (0.237 ± 0.025; P = 1.00). In other words, the semi-automated method revealed that healthy controls had significantly higher CBR compared to PD patients (P = 0.001) and those with positive DaTSCAN (P = 0.049). Meanwhile, no significant difference was found between the healthy participants and clinically uncertain patients with negative DaTSCAN (P = 1.00).

Table 18. Summary of mean ± SD CBR in ventral SNpc and dorsal SNpc of clinically uncertain patients with negative and positive DaTSCAN measured by automated template-based method

Subregion	Ventral SNpc CBR (mean ± SD)	P-value	Dorsal SNpc CBR (mean ± SD)	P-value
Negative DaTSCAN	0.092±0.024	0.029	0.181±0.033	0.014
Positive DaTSCAN	0.065±0.030		0.156±0.021	

Table 19. Summary of mean ± SD CBR in ventral SNpc and dorsal SNpc of healthy controls and PD patients measured by automated template-based method

Subregion	Ventral SNpc CBR (mean ± SD)	P-value	Dorsal SNpc CBR (mean ± SD)	P-value
Healthy controls	0.093±0.026	0.001	0.163±0.019	0.001
PD patients	0.057±0.029		0.149±0.033	

Table 20. One-way ANOVA and multiple comparisons of CBR (mean ± SD) in ventral and dorsal SNpc of different groups measured by the automated template-based method

Subregion	Ventral SNpc CBR (mean ± SD)	Dorsal SNpc CBR (mean ± SD)
Healthy controls	$0.093±0.026^{a}$	$0.163±0.019^{a}$
Negative DaTSCAN	$0.092±0.023^{ab}$	$0.181±0.033^{a}$
Positive DaTSCAN	$0.065±0.028^{bc}$	$0.156±0.021^{ab}$
PD patients	$0.057±0.029^{c}$	$0.149±0.033^{b}$

Note: CBR is presented in mean ± SD, comparisons are along each column and values with completely different superscripts (letters) are statistically different at P < 0.05. Multiple comparisons were conducted using One-way ANOVA and Bonferroni post-hoc test.

Table 21. Summary of mean ± SD CBR in ventral SNpc and dorsal SNpc in clinically uncertain patients with negative and positive DaTSCAN measured by semi-automated manual drawing method

Subregion	Ventral SNpc CBR (mean ± SD)	P-value	Dorsal SNpc CBR (mean ± SD)	P-value
Negative DaTSCAN	0.195±0.028	0.047	0.257±0.038	0.231
Positive DaTSCAN	0.169±0.031		0.237±0.025	

Table 22. Summary of mean ± SD CBR in ventral SNpc and dorsal SNpc of healthy controls and PD patients measured by semi-automated manual drawing method

Subregion	Ventral SNpc CBR (mean ± SD)	P-value	Dorsal SNpc CBR (mean ± SD)	P-value
Healthy controls	0.204±0.041	0.001	0.258±0.033	0.001
PD patients	0.158±0.022		0.231±0.036	

Table 23. One-way ANOVA and multiple comparisons of CBR (mean ± SD) in ventral and dorsal SNpc of different groups measured by the semi-automated manual drawing method

Subregion	Ventral SNpc CBR (mean ± SD)	Dorsal SNpc CBR (mean ± SD)
Healthy controls	0.204±0.041^{a}	0.258±0.033^{a}
Negative DaTSCAN	0.195±0.028^{ab}	0.257±0.038^{acd}
Positive DaTSCAN	0.168±0.031^{bc}	0.237±0.025^{bcd}
PD patients	0.158±0.022^{c}	0.231±0.036^{d}

Note: CBR is presented in mean ± SD, comparisons are along each column and values with completely different superscripts (letters) are statistically different at P<0.05. Multiple comparisons were conducted using One-way ANOVA and Bonferroni post-hoc test.

The between-group comparisons and their corresponding p-values are further presented in Figures 39 and 40 for the automated and semiautomated techniques, respectively.

Method 1: Automated template-base method

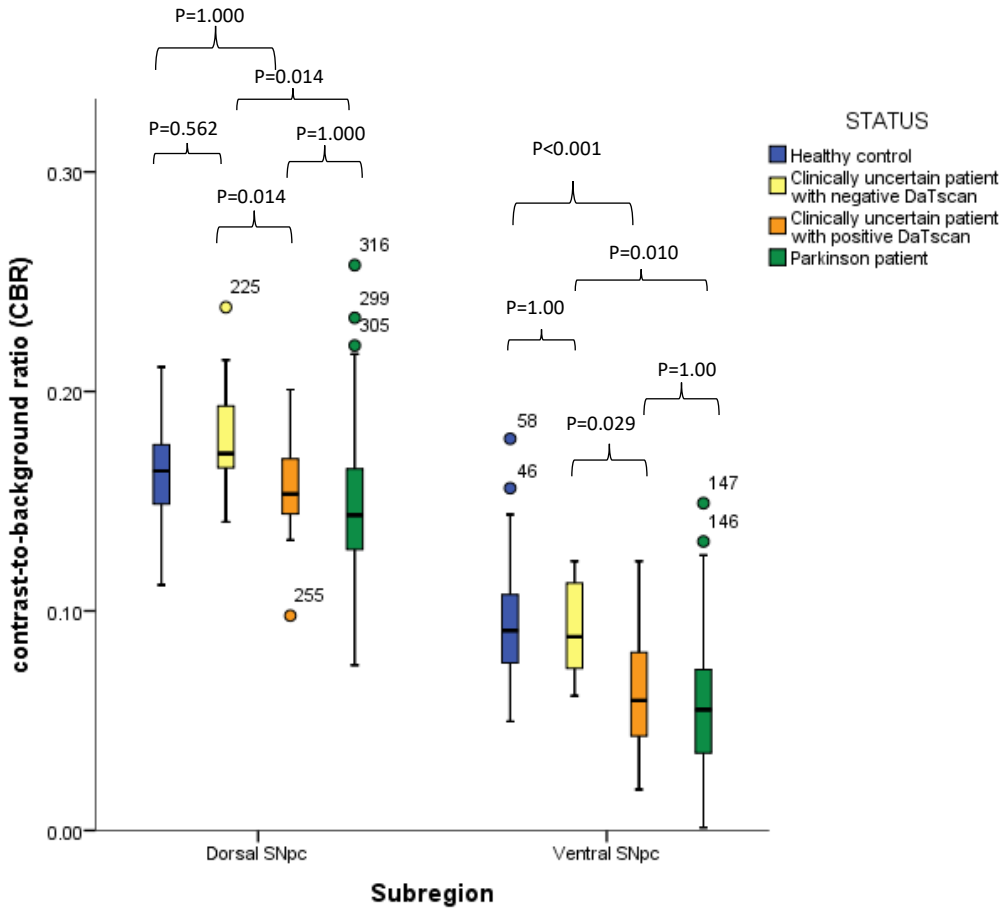


Figure 40. Box plot of CBRs measured using automated method in distinguishing different cohorts in two SNpc subregions: dorsal and ventral SNpc.

Method 2: Semiautomated manual drawing method

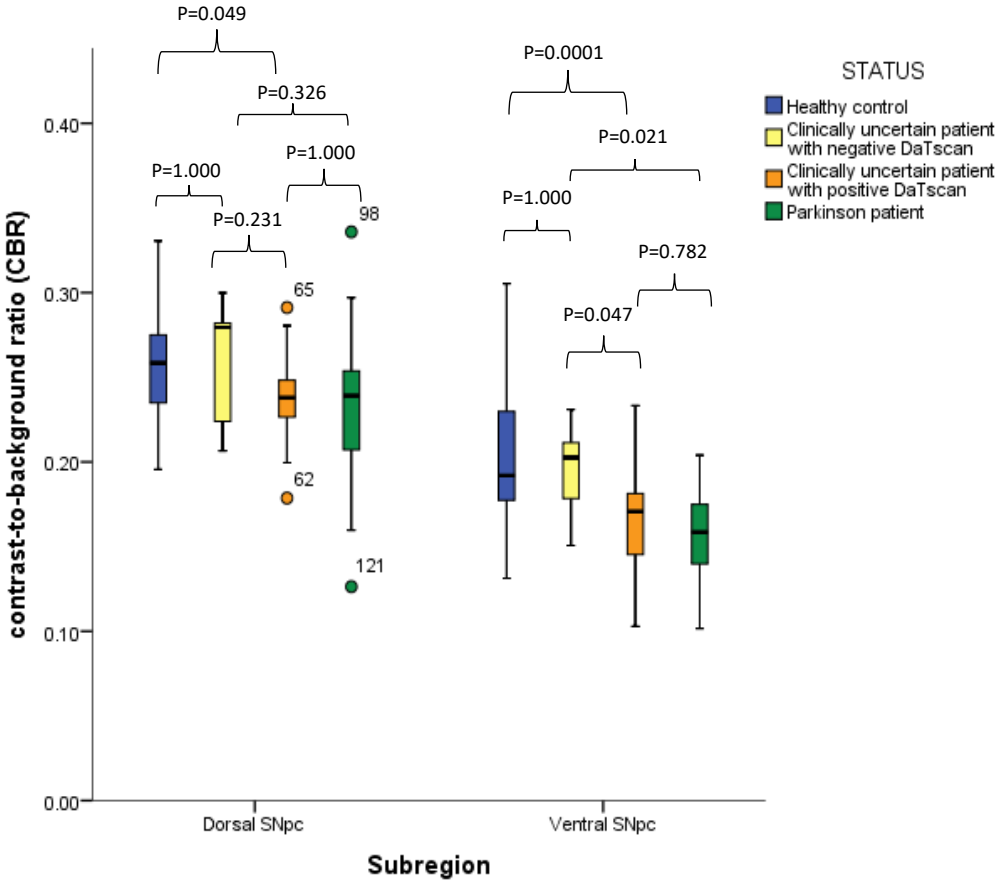


Figure 41. Box plot of CBRs measured using semiautomated method in distinguishing different cohorts in two SNpc subregions: dorsal and ventral SNpc.

8.4 Discussion

This preliminary study suggested that NM-MRI could potentially be an alternative to DaTSCAN in diagnosing CUPS. The automated and semiautomated image processing methods found reduced ventral SNpc CBR in clinically uncertain patients with positive DaTSCAN compared to clinically uncertain patients with negative DaTSCAN. The automated approach also reflected CBR reduction in dorsal SNpc but the finding was not significant with the semiautomated approach.

This study also attempted to distinguish negative and positive DaTSCAN groups from healthy control and PD patient using cohorts from a previous chapter (Chapter 7). Nevertheless, no difference was found to distinguish between positive DaTSCAN vs PD patients and between negative DaTSCAN against healthy control. One of the main limitations of this study is the small sample size as only seven participants were recruited in the negative DaTSCAN group. Diagnostic accuracy was also not analysed given that small sample sizes are associated with low statistical power. However, this study managed to achieve its intent of replicating the function of DaTSCAN, which is to differentiate non-degenerative and degenerative parkinsonism (Benamer et al., 2000).

One of the disadvantages of DaTSCAN is that the method alone is not sufficient to diagnose or differentiate PD from other neurodegenerative parkinsonian syndromes, such as atypical parkinsonism (Kalia & Lang, 2015). The application of NM-MRI in atypical parkinsonism was not tested in this thesis; however, the systematic review in Chapter 6.2 highlighted preliminary

studies differentiating atypical parkinsonian disorders, such as multiple system atrophy (MSA), progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD) (Biondetti et al., 2020; Kashihara et al., 2011a; Ohtsuka et al., 2014; Nadya Pyatigorskaya, Yahia-Cherif, et al., 2020). A recent review by Porter et al., 2020 suggested multimodal combinations of NM-MRI and iron imaging with DAT SPECT or positron emission tomography (PET) to assist in the clinical diagnosis with potential to identify PD subtypes and clearly distinguish PD from atypical Parkinsonism.

8.5 Conclusion

This research highlights the potential of NM-MRI in distinguishing non-neurodegenerative and neurodegenerative parkinsonism, which warrants further studies to become a biomarker similar to DaTSCAN. Future studies with a larger sample size may be worthwhile to understand and substantiate this relationship.

9. General discussion and Conclusion

9.1 Introduction

This final chapter presents a general discussion on NM MRI in PD patients. Discussion for each experiment is detailed in the relevant chapters. This chapter begins with a summary of findings covering several perspectives of NM-weighted imaging. Firstly, an assessment of image acquisition and image processing methods to detect SN signal intensity was highlighted. This was followed by discussing how the image processing methods were applied towards gaining further insight on risk factors, besides serving as a potential new biomarker for PD in clinical settings. Subsequently, limitations, future directions, and thesis conclusions are discussed.

9.2 NM intensity detection and estimation method

The earlier chapters of this thesis assessed two measurement methods in order to seek their reliability, consistency, and agreement before they were employed to analyse two NM-MRI protocols developed by our research group. This was particularly important to ensure that changes in the NM signal observed resulted from alterations related to the disease and not from intrinsic factors relevant to the machine or image processing methods.

Two SN intensity measurement methods; semiautomated manual drawing and automated template-based method showed high reliability, consistency, and agreement. This was evident as intrarater, interrater, short term rescan, and medium-term rescan analyses for all subregional SN CBR were within the upper and lower boundaries of the limit of agreement with minimal bias. Lower bias and narrower LOA observed in the automated

template-based method than in the semiautomated drawing method indicated better reliability, consistency, and agreement.

These two methods were then applied to two developed NM-MRI protocols involving short-term rescan without repositioning of individuals for SN detection amid obvious technical differences. BA plot illustrated almost all mean CBR in healthy and PD patients were within the upper and lower LOA. Therefore, implying that both protocols and methods of processing were reliable and were in high agreement for analysis of both healthy participants and PD patients. Scatterplots revealed that the sequence with the faster acquisition (Sequence 1) steadily produced higher SN intensity compared to the other sequence, where all CBR leaned towards one side and were away from the unity line.

Despite previous research applying the CBR as an imaging processing technique which detecting background signals in various brain structures (Isaias et al., 2016; Ohtsuka et al., 2013; Matsuura et al., 2016), the calculated CBR employed in this study has its drawbacks. In comparison to CNR, the CBR method may result in a lesser ratio due to the background anatomy signals. CNR produces more consistent background signals outside the brain resulting from constant ROI selection and well-calibrated MRI extrinsic noise, especially when image acquisition is undertaken in an MRI facility. Thus, CNR may depict the true inter-individual differences in SN signal rather than background variations. On the other hand, the background signal for CBR could vary following individuals' ROI, in this case, the cerebral cri. Vascular supply in the midbrain e.g., basilar artery may also cause artifacts, which may

cause increased BG signal in CBR (Ruchalski & Hathout, 2012). This event may affect the CBR measurement, hence might be a limiting factor for this method. It is proposed that future study to analyse both methods of background selection using similar data to explore further this relationship.

A limitation of *in-vivo* studies is the lack of true measurement or concentration, which can only be obtained from post-mortem specimens or synthetic phantom studies. Future research should consider postmortem samples or NM synthetic phantom. Hence, deviation of obtained NM-MRI measure from its true measurement through histological morphological measurement can be observed. Additionally, utilising postmortem specimens or synthetic phantom assists in eradicating artefacts that may arise due to movement among patients. This also offers opportunities for researchers to discover new protocols, besides allowing reliability and reproducibility studies to be conducted based on various acquisition parameters within a single session. Consequently, phantom and specimen are highly mobilised, thus facilitating their transportation from one facility to another to test for standardisation and harmonisation of measurement as proposed by meta-analysis in Chapter 6.1.

9.3 Understanding risk factor and pathogenesis of Parkinson

Extensive research has established ageing as an important risk factor for PD. In-vitro human lifespan trajectory of SN NM pigment began several decades ago, with contradicting results due to the differences in counting methods (Tanner & Goldman, 1996; F A Zucca et al., 2015). Our research

team was the first to investigate the in-vivo lifespan trajectory of NM in substantia nigra using NM-sensitive imaging.

Analysis of NM signal intensity among adults aged 20 years old up until old age yielded an inverted U-shaped trajectory. NM concentrations were found to accumulate linearly until middle age, about 55 years old before it slowly declined up to old age. The automated measurement method used in this chapter presented an enhanced subregional delineation compared to our recent similarly published semiautomated method.

The real function of NM pigment remains debatable and substantial evidence is required to confirm this NM-iron binding relationship. Besides, PD pathophysiology of NM reduction was associated with iron accumulation. Though a study combining NM-MRI and T2* sequence in healthy control and early PD patients did not reveal any correlation between neuromelanin signal and iron content in early Parkinson's. Hence, suggesting further studies are required to elucidate this pathological process (Reimao et al., 2016).

Another interesting area for discussion is neuroprotective factors such as nicotine or cigarette smoking, which has been inversely associated with the risk of PD. A recent meta-analysis denoted protective effects were most significant in men and presented a strong association with dose-response (Li et al., 2015). This relationship could be explained by a separate neurochemical study that detected the binding of nicotine to synthetic NM (Haining et al., 2016). Exploratory research may be worthwhile to better understand the underlying aetiology of the inverse relationship between smoking and Parkinson's, as well as the neuroprotective role of nicotine in

slowing down disease progression (Li et al., 2015; Nicholatos et al., 2018; Quik et al., 2012).

9.4 Role as a Parkinson disease biomarker

Any mechanism or tool must be extensively investigated before clinical implementation. This study subsequently explored the application of NM-MRI in clinical settings through several steps. Firstly, a systematic pooling of published studies was conducted to investigate its ability in diagnosing patients with PD from healthy control, its ability to distinguish between other parkinsonism syndromes, and whether it could detect the effect of disease severity and progression. Next, validated processing methods and developed sequences were implemented through NM imaging of the cohort of healthy control in this study, PD patients, and clinically uncertain parkinsonism.

Meta-analysis and systematic review consistently revealed support for the diagnostic ability of NM-MRI for PD against healthy participants, with significantly large disease size effect and high diagnostic performance. Analysis of heterogeneity revealed measurement metrics played a role in the result obtained, hence standardisation and harmonisation were proposed. Furthermore, the systematic review discussed the use of NM-MRI in distinguishing PD motor subtypes, differentiating atypical parkinsonian disorders, and detecting the effects of disease progression and severity. Meta-analyses could not be conducted for a few studies with small sample size, nevertheless, such studies could serve as preliminary evidence to depict other abilities of NM-MRI in PD within the neuroimaging field.

The moderately large case-control study reported in Chapter 7 utilised data scanned from two different sequences and the results were consistent with the findings from the meta-analysis. Significant reduction of SN signal intensity in PD patients upon normalisation was observed alongside high diagnostic performance. The study also highlighted that the automated method was more sensitive in detecting the effects of disease progression and severity cross-sectionally compared to the semiautomated method. Results from the automated method aligned with those from prior systematically pooled studies, which found SN intensity loss to be parallel with disease severity and progression. This is supported with three longitudinal studies (Gaurav et al., 2021; Matsuura et al., 2016; Xing et al., 2022). Longitudinal studies are vital to detect the association between individual losses of pigmented neurons and disease progression, thus substantiating its role as a disease progression marker. A longitudinal study is also warranted among normal participants for the best lifespan trajectory and full characterisation of NM pigment.

One limitation of the current study was that the UK Brain Bank criteria or MDS Parkinson Criteria and DaTSCAN were not the gold standards for diagnosis of PD and parkinsonism. Instead, histological study with postmortem specimens served as the gold standard for PD. Therefore, highlighting the possibility of false-positive or false-negative results. Nevertheless, based on meta-analysis and prospective study, NM-MRI consistently defended its role in distinguishing PD patients from healthy participants.

An equally important aspect of PD is detecting early parkinsonism during the preclinical stage to reduce the social and economic impact, as well as enhance patient management. To date, there is no screening tool or technique for the population with a high risk of PD or normal population. The systematic review also revealed the role of NM-MRI in rapid eye movement (REM) sleep behaviour disorder (RBD). It will be worthwhile to further prove and observe other prodromal factors, which could contribute to the development of Parkinson's.

The final study reported in this thesis recruited CUPS patients to determine if NM-weighted imaging could play a similar role as DaTSCAN, with the ability to differentiate degenerative from non-degenerative parkinsonism. This novel study found that NM-MRI could distinguish clinically uncertain parkinsonism in parallel with positive DaTSCAN and negative DaTSCAN results. However, a larger cohort study was proposed to explore this relationship and substantiate the practicality of using NM-MRI within the field.

This thesis also discussed the application of two other potential tools in neuroimaging of PD; (1) free-water imaging as another potentially useful biomarker, and (2) DaTSCAN, a well-established tool similar to that of NM-MRI with almost similar function. Different imaging tools revealed distinct information but combining these tools might be beneficial in terms of facilitating diagnosis or differentiating between PD and other types of parkinsonism. A review by Feraco et al. (2021) in which a multimodal MRI-PD

protocol is suggested to assist clinicians and neurologists in the differential diagnosis of neurogenerative disorders.

Retinal optical coherence tomography (OCT) measurement is another potential PD biomarker that may be worth exploring alongside NM-MRI. OCT is a fast, non-invasive, and cost-efficient method. A recent meta-analysis found a robust association between dopaminergic neurons and ganglion cell loss in the human retina and neurodegeneration in the human brain, which includes substantia nigra (Zhou et al., 2020).

9.5 Conclusion

The primary aim of this research was to assess and validate NM-rich area signal intensity estimation methods and test various applications of NM-MRI in PD. Two processing methods and two acquisition protocols with high reliability were presented. The automated processing approach proved to employ fewer human resources and a faster acquisition protocol with higher image contrast, thus suggesting its suitability for busy clinical settings.

Upon proving its usability to estimate SN signal intensity, meta-analysis and systematic review further implored NM-MRI relevant literature to determine its clinical ability to diagnose PD related disorders per the second aim of this research. All the studies consistently reported favourable results despite the observed heterogeneity observed, which was further supported by the subsequent large case-control study. A combination of these findings increased the possibility for NM-sensitive imaging to be adopted into clinical practice.

Studies on atypical parkinsonian disorders and clinically uncertain PD remain poorly understood. The systematic study and prospective study demonstrated the potential of NM-MRI in differentiating PD related disorders. However, its application could not be further confirmed given the low sample size in some of the reviewed studies. A larger cohort study was therefore proposed.

Conclusively, this research expands current knowledge of PD neuroimaging, proving the usability of NM-MRI not only as a diagnostic biomarker but also to monitor disease progression and severity effects, aside from clinically uncertain parkinsonism.

References

Albi, A., Pasternak, O., Minati, L., Marizzoni, M., Bartrés-Faz, D., Bargalló, N., Bosch, B., Rossini, P. M., Marra, C., Müller, B., Fiedler, U., Wiltfang, J., Roccatagliata, L., Picco, A., Nobili, F. M., Blin, O., Sein, J., Ranjeva, J. P., Didic, M., … Jovicich, J. (2017). Free water elimination improves test–retest reproducibility of diffusion tensor imaging indices in the brain: A longitudinal multisite study of healthy elderly subjects. *Human Brain Mapping*, *38*(1), 12–26. https://doi.org/10.1002/hbm.23350

Alladi, P. A., Mahadevan, A., Yasha, T. C., Raju, T. R., Shankar, S. K., & Muthane, U. (2009). Absence of age-related changes in nigral dopaminergic neurons of Asian Indians: Relevance to lower incidence of Parkinson's disease. *Neuroscience*, *159*(1), 236–245. https://doi.org/https://doi.org/10.1016/j.neuroscience.2008.11.051

Ariz, M., Abad, R. C., Castellanos, G., Martinez, M., Munoz-Barrutia, A., Fernandez-Seara, M. A., Pastor, P., Pastor, M. A., & Ortiz-De-Solorzano, C. (2019). Dynamic atlas-based segmentation and quantification of neuromelanin-rich brainstem structures in Parkinson disease. *IEEE Transactions on Medical Imaging*, *38*(3), 813–823. https://doi.org/10.1109/TMI.2018.2872852

Arribarat, G., Pasternak, O., De Barros, A., Galitzky, M., Rascol, O., & Péran, P. (2019). Substantia nigra locations of iron-content, free-water and mean diffusivity abnormalities in moderate stage Parkinson's disease. *Parkinsonism and Related Disorders*, *65*, 146–152. https://doi.org/10.1016/j.parkreldis.2019.05.033

Athauda, D., & Foltynie, T. (2015). The ongoing pursuit of neuroprotective therapies in Parkinson disease. *Nat Rev Neurol*, *11*(1), 25–40. https://doi.org/10.1038/nrneurol.2014.226

Bartzokis, G., Tishler, T. A., Lu, P. H., Villablanca, P., Altshuler, L. L., Carter, M., Huang, D., Edwards, N., & Mintz, J. (2007). Brain ferritin iron may influence age- and gender-related risks of neurodegeneration. *Neurobiology of Aging*, *28*(3), 414–423. https://doi.org/http://dx.doi.org/10.1016/j.neurobiolaging.2006.02.005

Bega, D., Kuo, P. H., Chalkidou, A., Grzeda, M. T., Macmillan, T., Brand, C., Sheikh, Z. H., & Antonini, A. (2021). Clinical utility of DaTscan in patients with suspected Parkinsonian syndrome: a systematic review and meta-analysis. *NPJ Parkinson's Disease*, *7*(1), 43. https://doi.org/10.1038/s41531-021-00185-8

Benamer, H. T. S., Patterson, J., Grosset, D. G., Booij, J., De Bruin, K., Van Royen, E., Speelman, J. D., Horstink, M. H. I. M., Sips, H. J. W. A., Dierckx, R. A., Versijpt, J., Decoo, D., Van Der Linden, C., Hadley, D. M., Doder, M., Lees, A. J., Costa, D. C., Gacinovic, S., Oertel, W. H., … Ries, V. (2000). Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: The [123I]-FP-CIT study group. *Movement Disorders*, *15*(3), 503–510. https://doi.org/10.1002/1531-8257(200005)15:3<503::AID-

MDS1013>3.0.CO;2-V

Berg, D, Lang, A. E., Postuma, R. B., Maetzler, W., Deuschl, G., Gasser, T., Siderowf, A., Schapira, A. H., Oertel, W., Obeso, J. A., Olanow, C. W., Poewe, W., & Stern, M. (2013). Changing the research criteria for the diagnosis of Parkinson's disease: obstacles and opportunities. *Lancet Neurol*, *12*(5), 514–524. https://doi.org/10.1016/s1474-4422(13)70047-4

Berg, Daniela, Postuma, R. B., Adler, C. H., Bloem, B. R., Chan, P., Dubois, B., Gasser, T., Goetz, C. G., Halliday, G., Joseph, L., Lang, A. E., Liepelt-Scarfone, I., Litvan, I., Marek, K., Obeso, J., Oertel, W., Olanow, C. W., Poewe, W., Stern, M., & Deuschl, G. (2015). MDS research criteria for prodromal Parkinson's disease. *Movement Disorders*, *30*(12), 1600–1611. https://doi.org/10.1002/mds.26431

Bolzoni, F., Giraudo, S., Lopiano, L., Bergamasco, B., Fasano, M., & Crippa, P. R. (2002). Magnetic investigations of human mesencephalic neuromelanin. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1586*(2), 210–218. https://doi.org/10.1016/S0925-4439(01)00099-0

Booij, J., Speelman, J. D., Horstink, M. W. I. M., & Wolters, E. C. (2001). The clinical benefit of imaging striatal dopamine transporters with [123I]FP-CIT SPET in differentiating patients with presynaptic parkinsonism from those with other forms of parkinsonism. *European Journal of Nuclear Medicine*, *28*(3), 266–272. https://doi.org/10.1007/s002590000460

Braak, H., Del Tredici, K., Rub, U., de Vos, R. A., Jansen Steur, E. N., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging*, *24*(2), 197–211. http://ac.els-cdn.com/S0197458002000659/1-s2.0-S0197458002000659-main.pdf?_tid=a8a5fe28-f82c-11e6-ba38-00000aacb35d&acdnat=1487678317_74dc73705b94a1ad50c11bbce5aae6de

Burciu, R. G., Ofori, E., Archer, D. B., Wu, S. S., Pasternak, O., McFarland, N. R., Okun, M. S., & Vaillancourt, D. E. (2017). *Progression marker of Parkinson's disease: A 4-year multi-site imaging study*. Brain. https://doi.org/10.1093/brain/awx146

Burn, D. J., Rowan, E. N., Allan, L. M., Molloy, S., O'Brien, J. T., & McKeith, I. G. (2006). Motor subtype and cognitive decline in Parkinson's disease, Parkinson's disease with dementia, and dementia with Lewy bodies. *J Neurol Neurosurg Psychiatry*, *77*(5), 585–589. https://doi.org/10.1136/jnnp.2005.081711

Cabello, C. R., Thune, J. J., Pakkenberg, H., & Pakkenberg, B. (2002). Ageing of substantia nigra in humans: cell loss may be compensated by hypertrophy. *Neuropathology and Applied Neurobiology*, *28*(4), 283–291. https://doi.org/10.1046/j.1365-2990.2002.00393.x

Carré, A., Klausner, G., Edjlali, M., Lerousseau, M., Briend-Diop, J., Sun, R., Ammari, S., Reuzé, S., Alvarez Andres, E., Estienne, T., Niyoteka, S., Battistella, E., Vakalopoulou, M., Dhermain, F., Paragios, N., Deutsch, E., Oppenheim, C., Pallud, J., & Robert, C. (2020). Standardization of brain

MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Scientific Reports 2020 10:1*, *10*(1), 1–15. https://doi.org/10.1038/s41598-020-69298-z

Cassidy, C. M., Zucca, F. A., Girgis, R. R., Baker, S. C., Weinstein, J. J., Sharp, M. E., Bellei, C., Valmadre, A., Vanegas, N., Kegeles, L. S., Brucato, G., Kang, U. J., Sulzer, D., Zecca, L., Abi-Dargham, A., & Horga, G. (2019). Neuromelanin-sensitive MRI as a noninvasive proxy measure of dopamine function in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(11), 5108–5117. https://doi.org/10.1073/pnas.1807983116

Castellanos, G., Fernandez-Seara, M. A., Lorenzo-Betancor, O., Ortega-Cubero, S., Puigvert, M., Uranga, J., Vidorreta, M., Irigoyen, J., Lorenzo, E., Munoz-Barrutia, A., Ortiz-de-Solorzano, C., Pastor, P., & Pastor, M. A. (2015). Automated neuromelanin imaging as a diagnostic biomarker for Parkinson's disease. *Mov Disord*, *30*(7), 945–952. https://doi.org/10.1002/mds.26201

Catafau, A. M., Tolosa, E., Laloux, P., Vander Borght, T., Van Zandijcke, M., De Geeter, F., Destee, A., Steinling, M., Lacomblez, L., Habert, M. O., Weiller, C., Clausen, M., Bogdahn, U., Haass, A., Kirsch, C. M., Antonini, A., Benti, R., Sorbi, S., Pupi, A., … Dudczak, R. (2004). Impact of dopamine transporter SPECT using 123I-Ioflupane on diagnosis and management of patients with clinically uncertain parkinsonian syndromes. *Movement Disorders*, *19*(10), 1175–1182. https://doi.org/10.1002/mds.20112

Cerri, S., Mus, L., & Blandini, F. (2019). Parkinson's Disease in Women and Men: What's the Difference? *Journal of Parkinson's Disease*, *9*(3), 501–515. https://doi.org/10.3233/JPD-191683

Chaudhuri, K. R., Healy, D. G., & Schapira, A. H. V. (2006). Non-motor symptoms of Parkinson's disease: Diagnosis and management. In *Lancet Neurology* (Vol. 5, Issue 3, pp. 235–245). https://doi.org/10.1016/S1474-4422(06)70373-8

Cho, S. J., Bae, Y. J., Kim, J. M., Kim, D., Baik, S. H., Sunwoo, L., Choi, B. S., & Kim, J. H. (2021). Diagnostic performance of neuromelanin-sensitive magnetic resonance imaging for patients with Parkinson's disease and factor analysis for its heterogeneity: a systematic review and meta-analysis. *European Radiology*, *31*(3), 1268–1280. https://doi.org/10.1007/s00330-020-07240-7

Cho, S. J., Bae, Y. J., Kim, J. M., Kim, H. J., Baik, S. H., Sunwoo, L., Choi, B. S., Jung, C., & Kim, J. H. (2021). Iron-sensitive magnetic resonance imaging in Parkinson's disease: a systematic review and meta-analysis. *Journal of Neurology*, *268*(12), 4721–4736. https://doi.org/10.1007/S00415-021-10582-X/TABLES/3

Clewett, D. V, Lee, T. H., Greening, S., Ponzio, A., Margalit, E., & Mather, M. (2016). Neuromelanin marks the spot: identifying a locus coeruleus biomarker of cognitive reserve in healthy aging. *Neurobiol Aging*, *37*, 117–126. https://doi.org/10.1016/j.neurobiolaging.2015.09.019

Conde, J. R., & Streit, W. J. (2006). Microglia in the aging brain. *J Neuropathol*

Exp Neurol, *65*(3), 199–203. https://doi.org/10.1097/01.jnen.0000202887.22082.63

Costa, K. M. (2014). The effects of aging on substantia nigra dopamine neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(46), 15133. https://doi.org/10.1523/JNEUROSCI.3739-14.2014

Cova, I., & Priori, A. (2018). Diagnostic biomarkers for Parkinson's disease at a glance: where are we? *Journal of Neural Transmission*, *125*(10), 1417–1432. https://doi.org/10.1007/s00702-018-1910-4

Cowen, D. (1986). The melanoneurons of the human cerebellum (nucleus pigmentosus cerebellaris) and homologues in the monkey. *J Neuropathol Exp Neurol*, *45*(3), 205–221. http://www.ncbi.nlm.nih.gov/pubmed/3958755

Damier, P., Hirsch, E. C., Agid, Y., & Graybiel, A. M. (1999). The substantia nigra of the human brain. II. Patterns of loss of dopamine-containing neurons in Parkinson's disease. *Brain*, *122 (Pt 8*, 1437–1448.

de Lau, L. M. L., & Breteler, M. M. B. (2006). Epidemiology of Parkinson's disease. *The Lancet Neurology*, *5*(6), 525–535. https://doi.org/10.1016/S1474-4422(06)70471-9

de Rijk, M. C., Launer, L. J., Berger, K., Breteler, M. M., Dartigues, J. F., Baldereschi, M., Fratiglioni, L., Lobo, A., Martinez-Lage, J., Trenkwalder, C., & Hofman, A. (2000). Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology*, *54*(11 Suppl 5), S21-3.

Di Lorenzo Alho, A. T., Suemoto, C. K., Polichiso, L., Tampellini, E., de Oliveira, K. C., Molina, M., Santos, G. A. B., Nascimento, C., Leite, R. E. P., de Lucena Ferreti-Rebustini, R. E., da Silva, A. V., Nitrini, R., Pasqualucci, C. A., Jacob-Filho, W., Heinsen, H., & Grinberg, L. T. (2016). Three-dimensional and stereological characterization of the human substantia nigra during aging. *Brain Structure and Function*, *221*(7), 3393–3403. https://doi.org/10.1007/s00429-015-1108-6

Dorsey, E. R., Elbaz, A., Nichols, E., Abd-Allah, F., Abdelalim, A., Adsuar, J. C., Ansha, M. G., Brayne, C., Choi, J.-Y. J., Collado-Mateo, D., Dahodwala, N., Do, H. P., Edessa, D., Endres, M., Fereshtehnejad, S.-M., Foreman, K. J., Gankpe, F. G., Gupta, R., Hankey, G. J., … Murray, C. J. L. (2018). Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study. *The Lancet Neurology*, *17*(11), 939–953. https://doi.org/10.1016/S1474-4422(18)30295-3

Dorsey, E. R., Sherer, T., Okun, M. S., & Bloem, B. R. (2018). The Emerging Evidence of the Parkinson Pandemic. *Journal of Parkinson's Disease*, *8*, S3–S8. https://doi.org/10.3233/JPD-181474

Double, K. L., Gerlach, M., Schünemann, V., Trautwein, A. X., Zecca, L., Gallorini, M., Youdim, M. B. H., Riederer, P., & Ben-Shachar, D. (2003). Iron-binding characteristics of neuromelanin of the human substantia nigra. *Biochemical Pharmacology*, *66*(3), 489–494. https://doi.org/10.1016/S0006-2952(03)00293-4

Engelen, M., Vanna, R., Bellei, C., Zucca, F. A., Wakamatsu, K., Monzani, E., Ito, S., Casella, L., & Zecca, L. (2012). Neuromelanins of Human Brain Have Soluble and Insoluble Components with Dolichols Attached to the Melanic Structure. *PLOS ONE*, *7*(11), e48490. https://doi.org/10.1371/JOURNAL.PONE.0048490

Enochs, W. S., Hyslop, W. B., Bennett, H. F., Brown 3rd, R. D., Koenig, S. H., & Swartz, H. M. (1989). Sources of the increased longitudinal relaxation rates observed in melanotic melanoma. An in vitro study of synthetic melanins. *Invest Radiol*, *24*(10), 794–804.

Enochs, W. S., Nilges, M. J., & Swartz, H. M. (1993). Purified Human Neuromelanin, Synthetic Dopamine Melanin as a Potential Model Pigment, and the Normal Human Substantia Nigra: Characterization by Electron Paramagnetic Resonance Spectroscopy. *Journal of Neurochemistry*, *61*(1), 68–79. https://doi.org/10.1111/j.1471-4159.1993.tb03538.x

Enochs, W. S., Petherick, P., Bogdanova, A., Mohr, U., & Weissleder, R. (1997). Paramagnetic metal scavenging by melanin: MR imaging. *Radiology*, *204*(2), 417–423. https://doi.org/10.1148/radiology.204.2.9240529

Esiri, M. M. (2007). Ageing and the brain. In *Journal of Pathology* (Vol. 211, Issue 2, pp. 181–187). https://doi.org/10.1002/path.2089

Fabbri, M., Reimao, S., Carvalho, M., Nunes, R. G., Abreu, D., Guedes, L. C., Bouca, R., Lobo, P. P., Godinho, C., Coelho, M., Goncalves, N. C., Rosa, M. M., Antonini, A., Ferreira, J. J., Reimão, S., Carvalho, M., Nunes, R. G., Abreu, D., Guedes, L. C., … Ferreira, J. J. (2017). Substantia Nigra Neuromelanin as an Imaging Biomarker of Disease Progression in Parkinson's Disease. *J Parkinsons Dis*, *7*(3), 491–501. https://doi.org/10.3233/jpd-171135

Farrall, A. J., & Wardlaw, J. M. (2009). Blood-brain barrier: ageing and microvascular disease--systematic review and meta-analysis. *Neurobiol Aging*, *30*(3), 337–352. https://doi.org/10.1016/j.neurobiolaging.2007.07.015

Fearnley, J. M., & Lees, A. J. (1991). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain*, *114*(5), 2283–2301. https://doi.org/10.1093/brain/114.5.2283

Fedorow, H., Tribl, F., Halliday, G., Gerlach, M., Riederer, P., & Double, K. L. (2005). Neuromelanin in human dopamine neurons: Comparison with peripheral melanins and relevance to Parkinson's disease. *Progress in Neurobiology*, *75*(2), 109–124. https://doi.org/10.1016/j.pneurobio.2005.02.001

Ferri, C. P., Prince, M., Brayne, C., Brodaty, H., Fratiglioni, L., Ganguli, M., Hall, K., Hasegawa, K., Hendrie, H., Huang, Y., Jorm, A., Mathers, C., Menezes, P. R., Rimmer, E., & Scazufca, M. (2005). Global prevalence of dementia: A Delphi consensus study. *Lancet*, *366*(9503), 2112–2117. https://doi.org/10.1016/S0140-6736(05)67889-0

Fu, M. C., Buerba, R. A., Long, W. D., Blizzard, D. J., Lischuk, A. W., Haims, A. H., & Grauer, J. N. (2014). Interrater and intrarater agreements of

magnetic resonance imaging findings in the lumbar spine: Significant variability across degenerative conditions. *Spine Journal*, *14*(10), 2442–2448. https://doi.org/10.1016/J.SPINEE.2014.03.010

Gaurav, R., Yahia-Cherif, L., Pyatigorskaya, N., Mangone, G., Biondetti, E., Valabrègue, R., Ewenczyk, C., Hutchison, R. M., Cedarbaum, J. M., Corvol, J.-C., Vidailhet, M., & Lehéricy, S. (2021). Longitudinal Changes in Neuromelanin MRI Signal in Parkinson's Disease: A Progression Marker. *Movement Disorders : Official Journal of the Movement Disorder Society*. https://doi.org/10.1002/mds.28531

Gelb, D. J., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for Parkinson disease. *Archives of Neurology*, *56*(1), 33–39. https://doi.org/10.1001/archneur.56.1.33

German, D. C., Manaye, K., Smith, W. K., Woodward, D. J., & Saper, C. B. (1989). Midbrain dopaminergic cell loss in Parkinson's disease: computer visualization. *Ann Neurol*, *26*(4), 507–514. https://doi.org/10.1002/ana.410260403

Gibb, W. R. G., & Lees, A. J. (1991). Anatomy, pigmentation, ventral and dorsal subpopulations of the substantia nigra, and differential cell death in Parkinson's disease. *J Neurol Neurosurg Psychiatry*, *54*(5), 388–396. https://doi.org/10.1136/jnnp.54.5.388

Goetz, C. G., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stebbins, G. T., Stern, M. B., Tilley, B. C., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., Olanow, C. W., … LaPelle, N. (2007). Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Process, format, and clinimetric testing plan. *Movement Disorders*, *22*(1), 41–47. https://doi.org/10.1002/mds.21198

Grossman, R. I., Gomori, J. M., Ramer, K. N., Lexa, F. J., & Schnall, M. D. (1994). Magnetization transfer: theory and clinical applications in neuroradiology. *Https://Doi.Org/10.1148/Radiographics.14.2.8190954*, *14*(2), 279–290. https://doi.org/10.1148/RADIOGRAPHICS.14.2.8190954

Guttuso, T., Bergsland, N., Hagemeier, J., Lichter, D. G., Pasternak, O., & Zivadinov, R. (2018). Substantia nigra free water increases longitudinally in Parkinson disease. *American Journal of Neuroradiology*, *39*(3), 479–484. https://doi.org/10.3174/ajnr.A5545

Haining, R. L., Jones, T. M., & Hernandez, A. (2016). Saturation Binding of Nicotine to Synthetic Neuromelanin Demonstrated by Fluorescence Spectroscopy. *Neurochemical Research*, *41*(12), 3356–3363. https://doi.org/10.1007/s11064-016-2068-9

Halliday, G. M., Fedorow, H., Rickert, C. H., Gerlach, M., Riederer, P., & Double, K. L. (2006). Evidence for specific phases in the development of human neuromelanin. *Journal of Neural Transmission*, *113*(6), 721–728. https://doi.org/10.1007/s00702-006-0449-y

Halliday, G. M., McRitchie, D. A., Cartwright, H., Pamphlett, R., Hely, M. A., & Morris, J. G. (1996). Midbrain neuropathology in idiopathic Parkinson's disease and diffuse Lewy body disease. *J Clin Neurosci*, *3*(1), 52–60.

Hauser, R. A., Koller, W. C., Hubble, J. P., Malapira, T., Busenbark, K., &

Olanow, C. W. (2000). Time course of loss of clinical benefit following withdrawal of levodopa/carbidopa and bromocriptine in early Parkinson's disease. *Movement Disorders*, *15*(3), 485–489. https://doi.org/10.1002/1531-8257(200005)15:3<485::AID-MDS1010>3.0.CO;2-F

Henkelman, R. M., Stanisz, G. J., & Graham, S. J. (2001). Magnetization transfer in MRI: a review. *NMR in Biomedicine*, *14*(2), 57–64. https://doi.org/10.1002/NBM.683

Hirsch, E., Graybiel, A. M., & Agid, Y. A. (1988). Melanized dopaminergic neurons are differentially susceptible to degeneration in Parkinson's disease. *Nature*, *334*(6180), 345–348. https://doi.org/10.1038/334345a0

Hirtz, D., Thurman, D. J., Gwinn-Hardy, K., Mohamed, M., Chaudhuri, A. R., & Zalutsky, R. (2007). How common are the "common" neurologic disorders? *Neurology*, *68*(5), 326–337. https://doi.org/10.1212/01.wnl.0000252807.38124.a3

Hodaie, M., Neimat, J. S., & Lozano, A. M. (2007). The dopaminergic nigrostriatal system and Parkinson's disease: Molecular events in development, disease, and cell death, and new therapeutic strategies. *Neurosurgery*, *60*(1), 17–28. https://doi.org/10.1227/01.NEU.0000249209.11967.CB

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology*, *17*(5), 427–442. https://doi.org/10.1212/WNL.17.5.427

Holdorff, B. (2002). Friedrich Heinrich Lewy (1885-1950) and his work. *Journal of the History of the Neurosciences*, *11*(1), 19–28. https://doi.org/10.1076/jhin.11.1.19.9106

Hozo, S. P., Djulbegovic, B., & Hozo, I. (2005). Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology*, *5*(1), 13. https://doi.org/10.1186/1471-2288-5-13

Huddleston, D. E., Langley, J., Sedlacik, J., Boelmans, K., Factor, S. A., & Hu, X. P. (2017). In vivo detection of lateral-ventral tier nigral degeneration in Parkinson's disease. *Hum Brain Mapp*, *38*(5), 2627–2634. https://doi.org/10.1002/hbm.23547

Hughes, A. J., Daniel, S. E., Ben-Shlomo, Y., & Lees, A. J. (2002). The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain*, *125*(Pt 4), 861–870.

Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery & Psychiatry*, *55*(3), 181–184. https://doi.org/10.1136/jnnp.55.3.181

Iranzo, A., Fernández-Arcos, A., Tolosa, E., Serradell, M., Molinuevo, J. L., Valldeoriola, F., Gelpi, E., Vilaseca, I., Sánchez-Valle, R., Lladó, A., Gaig, C., & Santamaría, J. (2014). Neurodegenerative disorder risk in idiopathic REM sleep behavior disorder: Study in 174 patients. *PLoS ONE*, *9*(2), 89741. https://doi.org/10.1371/journal.pone.0089741

Isaias, I. U., Trujillo, P., Summers, P., Marotta, G., Mainardi, L., Pezzoli, G.,

Zecca, L., & Costa, A. (2016). Neuromelanin Imaging and Dopaminergic Loss in Parkinson's Disease. *Front Aging Neurosci*, *8*, 196. https://doi.org/10.3389/fnagi.2016.00196

Jellinger, K. A. (2012). Neuropathology of sporadic Parkinson's disease: evaluation and changes of concepts. *Mov Disord*, *27*(1), 8–30. https://doi.org/10.1002/mds.23795

Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. In *The Lancet* (Vol. 386, Issue 9996, pp. 896–912). Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(14)61393-3

Kashihara, K., Shinya, T., & Higaki, F. (2011a). Reduction of Neuromelanin-Positive Nigral Volume in Patients with MSA, PSP and CBD. *Internal Medicine*, *50*(16), 1683–1687. https://doi.org/10.2169/internalmedicine.50.5101

Kashihara, K., Shinya, T., & Higaki, F. (2011b). Neuromelanin magnetic resonance imaging of nigral volume loss in patients with Parkinson's disease. *Journal of Clinical Neuroscience*, *18*(8), 1093–1096. https://doi.org/10.1016/j.jocn.2010.08.043

Kitao, S., Matsusue, E., Fujii, S., Miyoshi, F., Kaminou, T., Kato, S., Ito, H., & Ogawa, T. (2013). Correlation between pathology and neuromelanin MR imaging in Parkinson's disease and dementia with Lewy bodies. *Neuroradiology*, *55*(8), 947–953. https://doi.org/10.1007/s00234-013-1199-9

Kordower, J. H., Olanow, C. W., Dodiya, H. B., Chu, Y., Beach, T. G., Adler, C. H., Halliday, G. M., & Bartus, R. T. (2013). Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. *Brain*, *136*(8), 2419–2431. https://doi.org/10.1093/brain/awt192

Kubis, N., Faucheux, B. A., Ransmayr, G., Damier, P., Duyckaerts, C., Henin, D., Forette, B., Le Charpentier, Y., Hauw, J. J., Agid, Y., & Hirsch, E. C. (2000). Preservation of midbrain catecholaminergic neurons in very old human subjects. *Brain*, *123 (Pt 2*, 366–373.

Kupsch, A., Bajaj, N., Weiland, F., Tartaglione, A., Klutmann, S., Copp, R., Sherwin, P., Tate, A., & Grachev, I. D. (2013). Changes in clinical management and diagnosis following DaTscan SPECT imaging in patients with clinically uncertain parkinsonian syndromes: a 12-week follow-up study. *Neuro-Degenerative Diseases*, *11*(1), 22–32. https://doi.org/10.1159/000337351

Lang, A. E. (2011). A critical appraisal of the premotor symptoms of Parkinson's disease: Potential usefulness in early diagnosis and design of neuroprotective trials. In *Movement Disorders* (Vol. 26, Issue 5, pp. 775–783). https://doi.org/10.1002/mds.23609

Langley, J., Huddleston, D. E., Crosson, B., Song, D. D., Factor, S. A., & Hu, X. (2020). Multimodal assessment of nigrosomal degeneration in Parkinson's disease. *Parkinsonism and Related Disorders*, *80*, 102–107. https://doi.org/10.1016/j.parkreldis.2020.09.021

Langley, J., Huddleston, D. E. D. E., Chen, X., Neuroimage, J. S.-, 2015, U., Sedlacik, J., Zachariah, N., & Hu, X. (2015). A multicontrast approach for comprehensive imaging of substantia nigra. *Neuroimage*, *112*, 7–13.

https://doi.org/10.1016/j.neuroimage.2015.02.045
Langley, J., Huddleston, D. E., Liu, C. J., & Hu, X. (2017). Reproducibility of locus coeruleus and substantia nigra imaging with neuromelanin sensitive MRI. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *30*(2), 121–125. https://doi.org/10.1007/s10334-016-0590-z
Le Berre, A., Kamagata, K., Otsuka, Y., Andica, C., Hatano, T., Saccenti, L., Ogawa, T., Takeshige-Amano, H., Wada, A., Suzuki, M., Hagiwara, A., Irie, R., Hori, M., Oyama, G., Shimo, Y., Umemura, A., Hattori, N., & Aoki, S. (2019). Convolutional neural network-based segmentation can help in assessing the substantia nigra in neuromelanin MRI. *Neuroradiology*, *61*(12), 1387–1395. https://doi.org/10.1007/s00234-019-02279-w
Li, X., Li, W., Liu, G., Shen, X., & Tang, Y. (2015). Association between cigarette smoking and Parkinson's disease: A meta-analysis. *Arch Gerontol Geriatr*, *61*(3), 510–516. https://doi.org/10.1016/j.archger.2015.08.004
Ma, Röytt, Collan, Rinne, Ma, S. Y., Röyttä, M., Collan, Y., & Rinne, J. O. (1999). Unbiased morphometrical measurements show loss of pigmented nigral neurones with ageing. *Neuropathology and Applied Neurobiology*, *25*(5), 394–399. https://doi.org/10.1046/j.1365-2990.1999.00202.x
Mann, D. M. (1983). The locus coeruleus and its possible role in ageing and degenerative disease of the human central nervous system. *Mech Ageing Dev*, *23*(1), 73–94.
Mann, D. M. A., & Yates, P. O. (1974). LIPOPROTEIN PIGMENTS—THEIR RELATIONSHIP TO AGEING IN THE HUMAN NERVOUS SYSTEMII. THE MELANIN CONTENT OF PIGMENTED NERVE CELLS. *Brain*, *97*(1), 489–498. https://doi.org/10.1093/brain/97.1.489
Marras, C, Beck, J. C., Bower, J. H., Roberts, E., Ritz, B., Ross, G. W., Abbott, R. D., Savica, R., Van Den Eeden, S. K., Willis, A. W., Tanner, C. M., & Group, on behalf of the P. F. P. (2018). Prevalence of Parkinson's disease across North America. *Npj Parkinson's Disease*, *4*(1), 21. https://doi.org/10.1038/s41531-018-0058-0
Marras, C, & Chaudhuri, K. R. (2016). Nonmotor features of Parkinson's disease subtypes. *Mov Disord*, *31*(8), 1095–1102. https://doi.org/10.1002/mds.26510
Marras, Connie, Rochon, P., & Lang, A. E. (2002). Predicting motor decline and disability in Parkinson disease: A systematic review. In *Archives of Neurology* (Vol. 59, Issue 11, pp. 1724–1728). American Medical Association. https://doi.org/10.1001/archneur.59.11.1724
Martin Bland, J., & Altman, D. G. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8
Matsuura, K., Maeda, M., Tabei, K. I., Umino, M., Kajikawa, H., Satoh, M., Kida, H., & Tomimoto, H. (2016). A longitudinal study of neuromelanin-sensitive magnetic resonance imaging in Parkinson's disease. *Neurosci Lett*, *633*, 112–117. https://doi.org/10.1016/j.neulet.2016.09.011
Matsuura, K., Maeda, M., Yata, K., Ichiba, Y., Yamaguchi, T., Hidekazu, K. K.,

Tomimoto, H., Kanamaru, K., & Tomimoto, H. (2013). Neuromelanin magnetic resonance imaging in Parkinson's disease and multiple system atrophy. *Eur Neurol*, *70*(1–2), 70–77. https://doi.org/10.1159/000350291

McGeer, P. L., McGeer, E. G., & Suzuki, J. S. (1977). Aging and extrapyramidal function. *Arch Neurol*, *34*(1), 33–35. https://doi.org/10.1001/archneur.1977.00500130053010

Miyoshi, F., Ogawa, T., Kitao, S. I., Kitayama, M., Shinohara, Y., Takasugi, M., Fujii, S., & Kaminou, T. (2013). Evaluation of Parkinson disease and Alzheimer disease with the use of neuromelanin MR imaging and (123)I-metaiodobenzylguanidine scintigraphy. *AJNR Am J Neuroradiol*, *34*(11), 2113–2118. https://doi.org/10.3174/ajnr.A3567

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

Moon, W. J., Park, J. Y., Yun, W. S., Jeon, J. Y., Moon, Y. S., Kim, H., Kwak, K. C., Lee, J. M., & Han, S. H. (2016). A Comparison of Substantia Nigra T1 Hyperintensity in Parkinson's Disease Dementia, Alzheimer's Disease and Age-Matched Controls: Volumetric Analysis of Neuromelanin Imaging. *Korean J Radiol*, *17*(5), 633–640. https://doi.org/10.3348/kjr.2016.17.5.633

Muthane, U., Yasha, T. C., & Shankar, S. K. (1998). Low numbers and no loss of melanized nigral neurons with increasing age in normal human brains from India. *Ann Neurol*, *43*(3), 283–287. https://doi.org/10.1002/ana.410430304

Nakane, T., Nihashi, T., Kawai, H., & Naganawa, S. (2008). Visualization of Neuromelanin in the Substantia Nigra and Locus Ceruleus at 1.5T Using a 3D-gradient Echo Sequence with Magnetization Transfer Contrast. *Magnetic Resonance in Medical Sciences*, *7*(4), 205–210. https://doi.org/10.2463/mrms.7.205

Nicholatos, J. W., Francisco, A. B., Bender, C. A., Yeh, T., Lugay, F. J., Salazar, J. E., Glorioso, C., & Libert, S. (2018). Nicotine promotes neuron survival and partially protects from Parkinson's disease by suppressing SIRT6. *Acta Neuropathologica Communications*, *6*(1), 120. https://doi.org/10.1186/s40478-018-0625-y

Noyce, A. J., Lees, A. J., & Schrag, A. E. (2016). The prediagnostic phase of Parkinson's disease. In *Journal of Neurology, Neurosurgery and Psychiatry* (Vol. 87, Issue 8, pp. 871–878). BMJ Publishing Group. https://doi.org/10.1136/jnnp-2015-311890

Ofori, E, Pasternak, O., Planetta, P. J., Burciu, R., Snyder, A., Febo, M., Golde, T. E., Okun, M. S., & Vaillancourt, D. E. (2015). Increased free water in the substantia nigra of Parkinson's disease: a single-site and multi-site study. *Neurobiol Aging*, *36*(2), 1097–1104. https://doi.org/10.1016/j.neurobiolaging.2014.10.029

Ofori, Edward, Krismer, F., Burciu, R. G., Pasternak, O., McCracken, J. L.,

Lewis, M. M., Du, G., McFarland, N. R., Okun, M. S., Poewe, W., Mueller, C., Gizewski, E. R., Schocke, M., Kremser, C., Li, H., Huang, X., Seppi, K., & Vaillancourt, D. E. (2017). Free water improves detection of changes in the substantia nigra in parkinsonism: A multisite study. *Movement Disorders*, *32*(10), 1457–1464. https://doi.org/10.1002/mds.27100

Ofori, Edward, Pasternak, O., Planetta, P. J., Li, H., Burciu, R. G., Snyder, A. F., Lai, S., Okun, M. S., & Vaillancourt, D. E. (2015). Longitudinal changes in free-water within the substantia nigra of Parkinson's disease. *Brain*, *138*(8), 2322–2331. https://doi.org/10.1093/brain/awv136

Ogisu, K., Kudo, K., Sasaki, M., Sakushima, K., Yabe, I., Sasaki, H., Terae, S., Nakanishi, M., & Shirato, H. (2013). 3D neuromelanin-sensitive magnetic resonance imaging with semi-automated volume measurement of the substantia nigra pars compacta for diagnosis of Parkinson's disease. *Neuroradiology*, *55*(6), 719–724. https://doi.org/10.1007/s00234-013-1171-8

Ohtsuka, C, Sasaki, M., Konno, K., Kato, K., Takahashi, J., Yamashita, F., & Terayama, Y. (2014). Differentiation of early-stage parkinsonisms using neuromelanin-sensitive magnetic resonance imaging. *Parkinsonism Relat Disord*, *20*(7), 755–760. https://doi.org/10.1016/j.parkreldis.2014.04.005

Ohtsuka, Chigumi, Sasaki, M., Konno, K., Koide, M., Kato, K., Takahashi, J., Takahashi, S., Kudo, K., Yamashita, F., & Terayama, Y. (2013). Changes in substantia nigra and locus coeruleus in patients with early-stage Parkinson's disease using neuromelanin-sensitive MR imaging. *Neuroscience Letters*, *541*, 93–98. https://doi.org/10.1016/j.neulet.2013.02.012

Okuzumi, A., Hatano, T., Kamagata, K., Hori, M., Mori, A., Oji, Y., Taniguchi, D., Daida, K., Shimo, Y., Yanagisawa, N., Nojiri, S., Aoki, S., & Hattori, N. (2019). Neuromelanin or DaT-SPECT: which is the better marker for discriminating advanced Parkinson's disease? *European Journal of Neurology*, *26*(11), 1408–1416. https://doi.org/10.1111/ENE.14009

Pakkenberg, B., Møller, A., Gundersen, H. J. G., Mouritzen Dam, A., Pakkenberg, H., Moller, A., Gundersen, H. J. G., Dam, A. M., & Pakkenberg, H. (1991). The absolute number of nerve cells in substantia nigra in normal subjects and in patients with Parkinson's disease estimated with an unbiased stereological method. *Journal of Neurology, Neurosurgery & Psychiatry*, *54*(1), 30–33. https://doi.org/10.1136/jnnp.54.1.30

Parent, M., & Parent, A. (2010). Substantia nigra and Parkinson's disease: A brief history of their long and intimate relationship. In *Canadian Journal of Neurological Sciences* (Vol. 37, Issue 3, pp. 313–319). Canadian Journal of Neurological Sciences. https://doi.org/10.1017/S0317167100010209

Parkinson, J. (2002). An essay on the shaking palsy. 1817. *J Neuropsychiatry Clin Neurosci*, *14*(2), 223–236; discussion 222. https://doi.org/10.1176/jnp.14.2.223

Pasternak, O., Sochen, N., Gur, Y., Intrator, N., & Assaf, Y. (2009). Free water

elimination and mapping from diffusion MRI. *Magnetic Resonance in Medicine*, *62*(3), 717–730. https://doi.org/10.1002/mrm.22055

Planetta, P. J., Ofori, E., Pasternak, O., Burciu, R. G., Shukla, P., DeSimone, J. C., Okun, M. S., McFarland, N. R., & Vaillancourt, D. E. (2016). Free-water imaging in Parkinson's disease and atypical parkinsonism. *Brain*, *139*(Pt 2), 495–508. https://doi.org/10.1093/brain/awv361

Porter, E., Roussakis, A. A., Lao-Kaim, N. P., & Piccini, P. (2020). Multimodal dopamine transporter (DAT) imaging and magnetic resonance imaging (MRI) to characterise early Parkinson's disease. In *Parkinsonism and Related Disorders* (Vol. 79, pp. 26–33). Elsevier Ltd. https://doi.org/10.1016/j.parkreldis.2020.08.010

Post, B., Merkus, M. P., de Bie, R. M. A., de Haan, R. J., & Speelman, J. D. (2005). Unified Parkinson's Disease Rating Scale motor examination: Are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable? *Movement Disorders*, *20*(12), 1577–1584. https://doi.org/10.1002/mds.20640

Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., Obeso, J., Marek, K., Litvan, I., Lang, A. E., Halliday, G., Goetz, C. G., Gasser, T., Dubois, B., Chan, P., Bloem, B. R., Adler, C. H., & Deuschl, G. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Movement Disorders*, *30*(12), 1591–1601. https://doi.org/10.1002/mds.26424

Postuma, R. B., Poewe, W., Litvan, I., Lewis, S., Lang, A. E., Halliday, G., Goetz, C. G., Chan, P., Slow, E., Seppi, K., Schaffer, E., Rios-Romenets, S., Mi, T., Maetzler, C., Li, Y., Heim, B., Bledsoe, I. O., & Berg, D. (2018). Validation of the MDS clinical diagnostic criteria for Parkinson's disease. *Movement Disorders : Official Journal of the Movement Disorder Society*, *33*(10), 1601–1608. https://doi.org/10.1002/mds.27362

Prasad, S., Stezin, A., Lenka, A., George, L., Saini, J., Yadav, R., & Pal, P. K. (2018). Three-dimensional neuromelanin-sensitive magnetic resonance imaging of the substantia nigra in Parkinson's disease. *Eur J Neurol*, *25*(4), 680–686. https://doi.org/10.1111/ene.13573

Pyatigorskaya, N., Magnin, B., Mongin, M., Yahia-Cherif, L., Valabregue, R., Arnaldi, D., Ewenczyk, C., Poupon, C., Vidailhet, M., & Lehéricy, X. S. (2018). Comparative study of MRI biomarkers in the substantia nigra to discriminate idiopathic Parkinson disease. *American Journal of Neuroradiology*, *39*(8), 1460–1467. https://doi.org/10.3174/ajnr.A5702

Pyatigorskaya, Nadya, Gaurav, R., Arnaldi, D., Leu-Semenescu, S., Yahia-Cherif, L., Valabregue, R., Vidailhet, M., Arnulf, I., & Lehéricy, S. (2017). *Magnetic Resonance Imaging Biomarkers to Assess Substantia Nigra Damage in Idiopathic Rapid Eye Movement Sleep Behavior Disorder*. Sleep. https://doi.org/10.1093/sleep/zsx149

Pyatigorskaya, Nadya, Yahia-Cherif, L., Gaurav, R., Ewenczyk, C., Gallea, C., Valabregue, R., Gargouri, F., Magnin, B., Degos, B., Roze, E., Bardinet, E., Poupon, C., Arnulf, I., Vidailhet, M., & Lehericy, S. (2020). Multimodal Magnetic Resonance Imaging Quantification of Brain Changes in Progressive Supranuclear Palsy. *Movement Disorders*, *35*(1), 161–170.

https://doi.org/10.1002/mds.27877
Quik, M., Perez, X. A., & Bordia, T. (2012). Nicotine as a potential neuroprotective agent for Parkinson's disease. *Movement Disorders : Official Journal of the Movement Disorder Society*, *27*(8), 947–957. https://doi.org/10.1002/mds.25028
Rao, S. S., Hofmann, L. A., & Shakil, A. (2006). Parkinson's disease: Diagnosis and treatment. *American Family Physician*, *74*(12).
Reimao, S., Ferreira, S., Nunes, R. G., Lobo, P. P., Neutel, D., Abreu, D., Goncalves, N., Campos, J., Ferreira, J. J., Pita Lobo, P., Neutel, D., Abreu, D., Goncalves, N., Campos, J., & Ferreira, J. J. (2016). Magnetic resonance correlation of iron content with neuromelanin in the substantia nigra of early-stage Parkinson's disease. *Eur J Neurol*, *23*(2), 368–374. https://doi.org/10.1111/ene.12838
Rizzo, G., Copetti, M., Arcuti, S., Martino, D., Fontana, A., & Logroscino, G. (2016). Accuracy of clinical diagnosis of Parkinson disease. *Neurology*, *86*(6), 566–576. https://doi.org/10.1212/WNL.0000000000002350
Rossi, C., Frosini, D., Volterrani, D., De Feo, P., Unti, E., Nicoletti, V., Kiferle, L., Bonuccelli, U., & Ceravolo, R. (2010). Differences in nigro-striatal impairment in clinical variants of early Parkinson's disease: Evidence from a FP-CIT SPECT study. *European Journal of Neurology*, *17*(4), 626–630. https://doi.org/10.1111/j.1468-1331.2009.02898.x
Ruchalski, K., & Hathout, G. M. (2012). A Medley of Midbrain Maladies: A Brief Review of Midbrain Anatomy and Syndromology for Radiologists. *Radiology Research and Practice*, *2012*, 1–11. https://doi.org/10.1155/2012/258524
Rudow, G., O'Brien, R., Savonenko, A. V, Resnick, S. M., Zonderman, A. B., Pletnikova, O., Marsh, L., Dawson, T. M., Crain, B. J., West, M. J., & Troncoso, J. C. (2008). Morphometry of the human substantia nigra in ageing and Parkinson's disease. *Acta Neuropathol*, *115*(4), 461–470. https://doi.org/10.1007/s00401-008-0352-8
Rutledge, J. N., Hilal, S. K., Silver, A. J., Defendini, R., & Fahn, S. (1987). Study of movement disorders and brain iron by MR. *American Journal of Roentgenology*, *149*(2), 365–379. https://doi.org/10.2214/ajr.149.2.365
Sasaki, M., Shibata, E., Tohyama, K., Takahashi, J., Otsuka, K., Tsuchiya, K., Takahashi, S., Ehara, S., Terayama, Y., & Sakai, A. (2006). Neuromelanin magnetic resonance imaging of locus ceruleus and substantia nigra in Parkinson's disease. *Neuroreport*, *17*(11), 1215–1218. https://doi.org/10.1097/01.wnr.0000227984.84927.a7
Schrag, A., Horsfall, L., Walters, K., Noyce, A., & Petersen, I. (2015). Prediagnostic presentations of Parkinson's disease in primary care: a case-control study. *Lancet Neurol*, *14*(1), 57–64. https://doi.org/10.1016/s1474-4422(14)70287-x
Schulz-Schaeffer, W. J. (2010). The synaptic pathology of alpha-synuclein aggregation in dementia with Lewy bodies, Parkinson's disease and Parkinson's disease dementia. *Acta Neuropathol*, *120*(2), 131–143. https://doi.org/10.1007/s00401-010-0711-0
Schwarz, S., Bajaj, N., Morgan, P., Reid, S., … P. G.-P. I. S. M. R., & 2013, U.

(2013). *Magnetisation transfer contrast to enhance detection of neuromelanin loss at 3T in Parkinson's disease*.

Schwarz, S. T., Rittman, T., Gontu, V., Morgan, P. S., Bajaj, N., & Auer, D. P. (2011). T1-weighted MRI shows stage-dependent substantia nigra signal loss in Parkinson's disease. *Mov Disord*, *26*(9), 1633–1638. https://doi.org/10.1002/mds.23722

Schwarz, S. T., Xing, Y., Tomar, P., Bajaj, N., & Auer, D. P. (2017). In Vivo Assessment of Brainstem Depigmentation in Parkinson Disease: Potential as a Severity Marker for Multicenter Studies. *Radiology*, *283*(3), 789–798. https://doi.org/10.1148/radiol.2016160662

Shibata, E., Sasaki, M., Tohyama, K., Kanbara, Y., Otsuka, K., Ehara, S., & Sakai, A. (2006). Age-related changes in locus ceruleus on neuromelanin magnetic resonance imaging at 3 Tesla. *Magn Reson Med Sci*, *5*(4), 197–200. https://doi.org/10.2463/mrms.5.197

Shinde, S., Prasad, S., Saboo, Y., Kaushick, R., Saini, J., Pal, P. K., & Ingalhalikar, M. (2019). Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI. *NeuroImage: Clinical*, *22*, 101748. https://doi.org/10.1016/j.nicl.2019.101748

Simmons, A., Tofts, P. S., Barker, G. J., & Arridge, S. R. (1994). Sources of intensity nonuniformity in spin echo images at 1.5 T. *Magnetic Resonance in Medicine*, *32*(1), 121–128. https://doi.org/10.1002/MRM.1910320117

Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., & Goedert, M. (1997). Alpha-synuclein in Lewy bodies. *Nature*, *388*(6645), 839–840. https://doi.org/10.1038/42166

Sulzer, D., Cassidy, C., Horga, G., Kang, U. J., Fahn, S., Casella, L., Pezzoli, G., Langley, J., Hu, X. P., Zucca, F. A., Isaias, I. U., & Zecca, L. (2018). Neuromelanin detection by magnetic resonance imaging (MRI) and its promise as a biomarker for Parkinson's disease. In *npj Parkinson's Disease* (Vol. 4, Issue 1, p. 11). Nature Publishing Group. https://doi.org/10.1038/s41531-018-0047-3

Sulzer, D., & Zecca, L. (2000). Intraneuronal dopamine-quinone synthesis: a review. *Neurotox Res*, *1*(3), 181–195. https://doi.org/10.1007/bf03033289

Takahashi, H, Watanabe, Y., Tanaka, H., Mihara, M., Mochizuki, H., Liu, T., Wang, Y., & Tomiyama, N. (2018). Quantifying changes in nigrosomes using quantitative susceptibility mapping and neuromelanin imaging for the diagnosis of early-stage Parkinson's disease. *Br J Radiol*, *91*(1086), 20180037. https://doi.org/10.1259/bjr.20180037

Takahashi, Hiroto, Watanabe, Y., Tanaka, H., Mochizuki, H., Kato, H., Hatazawa, J., & Tomiyama, N. (2019). Quantifying the Severity of Parkinson Disease by Use of Dopaminergic Neuroimaging. *AJR. American Journal of Roentgenology*, *213*(1), 163–168. https://doi.org/10.2214/AJR.18.20655

Taniguchi, D., Hatano, T., Kamagata, K., Okuzumi, A., Oji, Y., Mori, A., Hori, M., Aoki, S., & Hattori, N. (2018). Neuromelanin imaging and midbrain volumetry in progressive supranuclear palsy and Parkinson's disease.

Movement Disorders, *33*(9), 1488–1492. https://doi.org/10.1002/mds.27365

Tanner, C. M., & Goldman, S. M. (1996). Epidemiology of Parkinson's disease. *Neurol Clin*, *14*(2), 317–335.

Trujillo, P., Summers, P. E., Smith, A. K., Smith, S. A., Mainardi, L. T., Cerutti, S., Claassen, D. O., & Costa, A. (2017). Pool size ratio of the substantia nigra in Parkinson's disease derived from two different quantitative magnetization transfer approaches. *Neuroradiology*, *59*(12), 1251–1263. https://doi.org/10.1007/s00234-017-1911-2

Tysnes, O. B., & Storstein, A. (2017). Epidemiology of Parkinson's disease. *J Neural Transm (Vienna)*, *124*(8), 901–905. https://doi.org/10.1007/s00702-017-1686-y

Van De Loo, S., Walter, U., Behnke, S., Hagenah, J., Lorenz, M., Sitzer, M., Hilker, R., & Berg, D. (2010). Reproducibility and diagnostic accuracy of substantia nigra sonography for the diagnosis of Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, *81*(10), 1087–1092. https://doi.org/10.1136/jnnp.2009.196352

van der Heeden, J. F., Marinus, J., Martinez-Martin, P., Rodriguez-Blazquez, C., Geraedts, V. J., & van Hilten, J. J. (2016). Postural instability and gait are associated with severity and prognosis of Parkinson disease. *Neurology*, *86*(24), 2243–2250. https://doi.org/10.1212/wnl.0000000000002768

van der Pluijm, M., Cassidy, C., Zandstra, M., Wallert, E., de Bruin, K., Booij, J., de Haan, L., Horga, G., & van de Giessen, E. (2021). Reliability and Reproducibility of Neuromelanin-Sensitive Imaging of the Substantia Nigra: A Comparison of Three Different Sequences. *Journal of Magnetic Resonance Imaging*, *53*(3), 712. https://doi.org/10.1002/JMRI.27384

van Rooden, S. M., Colas, F., Martinez-Martin, P., Visser, M., Verbaan, D., Marinus, J., Chaudhuri, R. K., Kok, J. N., & van Hilten, J. J. (2011). Clinical subtypes of Parkinson's disease. *Mov Disord*, *26*(1), 51–58. https://doi.org/10.1002/mds.23346

Visser, M., Müller, D. M. J., van Duijn, R. J. M., Smits, M., Verburg, N., Hendriks, E. J., Nabuurs, R. J. A., Bot, J. C. J., Eijgelaar, R. S., Witte, M., van Herk, M. B., Barkhof, F., de Witt Hamer, P. C., & de Munck, J. C. (2019). Inter-rater agreement in glioma segmentations on longitudinal MRI. *NeuroImage: Clinical*, *22*, 101727. https://doi.org/10.1016/J.NICL.2019.101727

Vlaar, A. M. M., van Kroonenburgh, M. J. P. G., Kessels, A. G. H., & Weber, W. E. J. (2007). Meta-analysis of the literature on diagnostic accuracy of SPECT in parkinsonian syndromes. *BMC Neurology*, *7*. https://doi.org/10.1186/1471-2377-7-27

Wang, J., Li, Y., Huang, Z., Wan, W., Zhang, Y., Wang, C., Cheng, X., Ye, F., Liu, K., Fei, G., Zeng, M., & Jin, L. (2018). Neuromelanin-sensitive magnetic resonance imaging features of the substantia nigra and locus coeruleus in de novo Parkinson's disease and its phenotypes. *European Journal of Neurology*, *25*(7), 949–955. https://doi.org/10.1111/ene.13628

Wang, Jian, Huang, Z., Li, Y., Ye, F., Wang, C., Zhang, Y., Cheng, X., Fei, G., Liu,

K., Zeng, M., Zhong, C., & Jin, L. (2019). Neuromelanin-sensitive MRI of the substantia nigra: An imaging biomarker to differentiate essential tremor from tremor-dominant Parkinson's disease. *Parkinsonism and Related Disorders*, *58*, 3–8. https://doi.org/10.1016/j.parkreldis.2018.07.007

Wang, X., Zhang, Y., Zhu, C., Li, G., Kang, J., Chen, F., & Yang, L. (2019). The diagnostic value of SNpc using NM-MRI in Parkinson's disease: meta-analysis. *Neurological Sciences*, *40*(12), 2479–2489. https://doi.org/10.1007/S10072-019-04014-Y/FIGURES/5

Wang, Y., Ji, X., Leak, R. K., Chen, F., & Cao, G. (2017). Stem cell therapies in age-related neurodegenerative diseases and stroke. *Ageing Res Rev*, *34*, 39–50. https://doi.org/10.1016/j.arr.2016.11.002

Weir, S., Samnaliev, M., Kuo, T.-C., Tierney, T. S., Walleser Autiero, S., Taylor, R. S., & Schrag, A. (2018). Short- and long-term cost and utilization of health care resources in Parkinson's disease in the UK. *Movement Disorders*, *33*(6), 974–981. https://doi.org/doi:10.1002/mds.27302

Wengler, K., He, X., Abi-Dargham, A., & Horga, G. (2020). Reproducibility assessment of neuromelanin-sensitive magnetic resonance imaging protocols for region-of-interest and voxelwise analyses. *NeuroImage*, *208*, 116457. https://doi.org/10.1016/j.neuroimage.2019.116457

Xiang, Y. Y., Gong, T., Wu, J. W., Li, J. F., Chen, Y., Wang, Y. X., Li, S., Cong, L., Lin, Y. T., Han, Y. X., Yin, L., Wang, G. B., & Du, Y. F. (2017). Subtypes evaluation of motor dysfunction in Parkinson's disease using neuromelanin-sensitive magnetic resonance imaging. *Neurosci Lett*, *638*, 145–150. https://doi.org/10.1016/j.neulet.2016.12.036

Xing, Y., Sapuan, A., Dineen, R. A., & Auer, D. P. (2018). Life span pigmentation changes of the substantia nigra detected by neuromelanin-sensitive MRI. *Movement Disorders*, *33*(11), 1792–1799. https://doi.org/10.1002/mds.27502

Xing, Y., Sapuan, A. H., Martín-Bastida, A., Naidu, S., Tench, C., Evans, J., Sare, G., Schwarz, S. T., Al-bachari, S., Parkes, L. M., Kanavou, S., Raw, J., Silverdale, M., Bajaj, N., Pavese, N., Burn, D., Piccini, P., Grosset, D. G., & Auer, D. P. (2022). Neuromelanin-MRI to Quantify and Track Nigral Depigmentation in Parkinson's Disease: A Multicenter Longitudinal Study Using Template-Based Standardized Analysis. *Movement Disorders : Official Journal of the Movement Disorder Society*. https://doi.org/10.1002/MDS.28934

Yang, J., Archer, D. B., Burciu, R. G., Müller, M. L. T. M., Roy, A., Ofori, E., Bohnen, N. I., Albin, R. L., & Vaillancourt, D. E. (2019). Multimodal dopaminergic and free-water imaging in Parkinson's disease. *Parkinsonism and Related Disorders*, *62*, 10–15. https://doi.org/10.1016/j.parkreldis.2019.01.007

Zecca, L, Bellei, C., Costi, P., Albertini, A., Monzani, E., Casella, L., Gallorini, M., Bergamaschi, L., Moscatelli, A., Turro, N. J., Eisner, M., Crippa, P. R., Ito, S., Wakamatsu, K., Bush, W. D., Ward, W. C., Simon, J. D., & Zucca, F. A. (2008). New melanic pigments in the human brain that accumulate in aging and block environmental toxic metals. *Proc Natl Acad Sci U S A*,

105(45), 17567–17572. https://doi.org/10.1073/pnas.0808768105

Zecca, L, Pietra, R., Goj, C., Mecacci, C., Radice, D., & Sabbioni, E. (1994). Iron and other metals in neuromelanin, substantia nigra, and putamen of human brain. *J Neurochem*, *62*(3), 1097–1101.

Zecca, L, Shima, T., Stroppolo, A., Goj, C., Battiston, G. A., Gerbasi, R., Sarna, T., & Swartz, H. M. (1996). Interaction of neuromelanin and iron in substantia nigra and other areas of human brain. *Neuroscience*, *73*(2), 407–415. https://doi.org/https://doi.org/10.1016/0306-4522(96)00047-4

Zecca, L, Stroppolo, A., Gatti, A., Tampellini, D., Toscani, M., Gallorini, M., Giaveri, G., Arosio, P., Santambrogio, P., Fariello, R. G., Karatekin, E., Kleinman, M. H., Turro, N., Hornykiewicz, O., & Zucca, F. A. (2004). The role of iron and copper molecules in the neuronal vulnerability of locus coeruleus and substantia nigra during aging. *Proc Natl Acad Sci U S A*, *101*(26), 9843–9848. https://doi.org/10.1073/pnas.0403495101

Zecca, L, & Swartz, H. M. (1993). Total and paramagnetic metals in human substantia nigra and its neuromelanin. *J Neural Transm Park Dis Dement Sect*, *5*(3), 203–213.

Zecca, L, Tampellini, D., Gerlach, M., Riederer, P., Fariello, R. G., & Sulzer, D. (2001). Substantia nigra neuromelanin: structure, synthesis, and molecular behaviour. *Mol Pathol*, *54*(6), 414–418.

Zecca, L, Zucca, F. A., Wilms, H., & Sulzer, D. (2003). Neuromelanin of the substantia nigra: a neuronal black hole with protective and toxic characteristics. *Trends Neurosci*, *26*(11), 578–580. https://doi.org/10.1016/j.tins.2003.08.009

Zecca, Luigi, Casella, L., Albertini, A., Bellei, C., Zucca, F. A., Engelen, M., Zadlo, A., Szewczyk, G., Zareba, M., & Sarna, T. (2008). Neuromelanin can protect against iron-mediated oxidative damage in system modeling iron overload of brain aging and Parkinson's disease. *J Neurochem*, *106*(4), 1866–1875. https://doi.org/10.1111/j.1471-4159.2008.05541.x

Zecca, Luigi, Fariello, R., Riederer, P., Sulzer, D., Gatti, A., & Tampellini, D. (2002). The absolute concentration of nigral neuromelanin, assayed by a new sensitive method, increases throughout the life and is dramatically decreased in Parkinson's disease. *FEBS Lett*, *510*(3), 216–220. https://doi.org/http://dx.doi.org/10.1016/S0014-5793(01)03269-0

Zecca, Luigi, Gallorini, M., Schunemann, V., Trautwein, A. X., Gerlach, M., Riederer, P., Vezzoni, P., Tampellini, D., Schünemann, V., Trautwein, A. X., Gerlach, M., Riederer, P., Vezzoni, P., & Tampellini, D. (2001). Iron, neuromelanin and ferritin content in the substantia nigra of normal subjects at different ages: consequences for iron storage and neurodegenerative processes. *J Neurochem*, *76*(6), 1766–1773. https://doi.org/10.1046/j.1471-4159.2001.00186.x

Zhou, W., Tao, J., & Li, J. (2020). Optical coherence tomography measurements as potential imaging biomarkers for Parkinson's disease: A systematic review and meta-analysis. *European Journal of Neurology*, ene.14613. https://doi.org/10.1111/ene.14613

Zucca, F A, Giaveri, G., Gallorini, M., Albertini, A., Toscani, M., Pezzoli, G.,

Lucius, R., Wilms, H., Sulzer, D., Ito, S., Wakamatsu, K., & Zecca, L. (2004). The neuromelanin of human substantia nigra: physiological and pathogenic aspects. *Pigment Cell Res*, *17*(6), 610–617. https://doi.org/10.1111/j.1600-0749.2004.00201.x

Zucca, F A, Segura-Aguilar, J., Ferrari, E., Munoz, P., Paris, I., Sulzer, D., Sarna, T., Casella, L., & Zecca, L. (2015). Interactions of iron, dopamine and neuromelanin pathways in brain aging and Parkinson's disease. *Prog Neurobiol*. https://doi.org/10.1016/j.pneurobio.2015.09.012

Zucca, Fabio A., Basso, E., Cupaioli, F. A., Ferrari, E., Sulzer, D., Casella, L., & Zecca, L. (2014). Neuromelanin of the Human Substantia Nigra: An Update. *Neurotoxicity Research*, *25*(1), 13–23. https://doi.org/10.1007/s12640-013-9435-y

Zupan, G., Šuput, D., Pirtošek, Z., & Vovk, A. (2019). Semi-automatic signature-based segmentation method for quantification of neuromelanin in substantia Nigra. *Brain Sciences*, *9*(12). https://doi.org/10.3390/brainsci9120335

Appendices

Appendix 1

Main project Title : The Development and Optimisation of Novel MRI Techniques study

Sub study title : Substantia nigra Imaging in Healthy Human and Phantom'

Chief Investigator : Dr Robert A Dineen

Research Ethical Committee (REC) number: B12012012a-051

Appendix 2

Title : The Parkinson's MR Imaging Repository (PaMIR),

Chief Investigator : Prof Dorothee P Auer

Research Ethical Committee (REC) number: 14/EM/0061

Funded by Parkinson's UK, NHS

Appendix 3

Title : Brain Network Dysfunction in Multiple Sclerosis (BRaNDy-MS),

Chief Investigator : Dr. Robert A Dineen

Research Ethical Committee (REC) number: 14/EM/0064

Funded by the MS Society

Appendix 4

Title : Nigrosomal iron imaging in Parkinson's disease (N3iPD)

Primary Investigator: Prof Dorothee Auer

Research Ethical Committee (REC) number: 16/EM/0229

Approved by: Health Research Authority East Midlands – Derby Research Ethics Committee)

Since it involved radioactive substances, the project also obtained approval from the Administration of Radioactive Substances Administration Committee (ARSAC Ref 253/3629/35864).