Multimodal neuroimaging signatures of early cART-treated paediatric HIV

Distinguishing perinatally HIV-infected 7-year-old children from uninfected controls

Prepared by:
Isaac Lebogang Khobo (HW)

Supervised by:
Dr Frances Robertson
Dr Marcin Jankiewicz
Dr Martha Holmes
Faculty of Health Sciences
Dept. of Human Biology—BME division

February 7, 2020

Submitted to the Department of Human Biology at the University of Cape Town in partial fulfilment of the academic requirements for an MSc (Med) in Biomedical Engineering.

DECLARATION

I, ISAAC LEBOGANG KHOBO, hereby declare that the work on which this thesis is based is my own original work (except where acknowledgements indicate otherwise) and that neither the whole, nor any part of it, has been, or is to be submitted for any other degree in this or any other University.

I empower the University to reproduce, for the purpose of research, either the whole or part of the contents of this thesis in any manner.

Signed by candidate

Signature

7 February 2020

To Lobisa Dimakatso:
You have loved me since I was born;
Who will love me now that you are gone?

---Your son, Hydrophilic Wonder

SYNOPSIS

Introduction: HIV-related brain alterations can be identified using neuroimaging modalities such as proton magnetic resonance spectroscopy (^{1}H-MRS), structural magnetic resonance imaging (sMRI), diffusion tensor imaging (DTI), and functional MRI (fMRI). However, few studies have combined multiple MRI measures/features to identify a multivariate neuroimaging signature that typifies HIV infection.

Elastic net (EN) regularisation uses penalised regression to perform variable selection, shrinking the weighting of unimportant variables to zero. We chose to use the embedded feature selection of EN logistic regression to identify a set of neuroimaging features characteristic of paediatric HIV infection. We aimed to determine 1) the most useful features across MRI modalities to separate HIV+ children from HIV- controls and 2) whether better classification performance is obtained by combining multimodal MRI features rather than using features from a single modality.

Methods: The study sample comprised 72 HIV+ 7-year-old children from the Children with HIV Early Antiretroviral Therapy (CHER) trial in Cape Town, who initiated combination antiretroviral therapy (cART) in infancy and had their viral loads suppressed from a young age, and 55 HIV- control children.

Neuroimaging features were extracted to generate 7 MRI-derived sets. For sMRI, 42 regional brain volumes (1st set), mean cortical thickness and gyrification in 68 brain regions (2nd and 3rd set) were used. For DTI data: radial (RD), axial (AD), mean (MD) diffusivities, and fractional anisotropy (FA) in each of 20 atlas regions were extracted for a total of 80 DTI features (4th set). For ^{1}H-MRS, concentrations of 14 metabolites and their ratios to creatine in the basal ganglia, peritrigonal white matter, and midfrontal gray matter voxels (5th, 6th and 7th set) were considered. A logistic EN regression model with repeated 10-fold cross validation (CV) was implemented in R, initially on each feature set separately. Sex, age and total intracranial volume (TIV) were included as confounders with no shrinkage penalty. For each model, the classification performance for HIV+ vs HIV- was assessed by computing accuracy, specificity, sensitivity, and mean area under the receiver operator characteristic curve (AUC) across 10 CV folds and 100 iterations.

To combine feature sets, the best performing set was concatenated with each of the other sets and further EN regressions were run. The combination giving the largest AUC was combined with each of the remaining sets until there was no further increase in AUC. Two concatenation techniques were explored: nested and non-nested modelling. All models were assessed for their goodness of fit using χ^2 likelihood ratio tests for non-nested models and Akaike information criterion (AIC) for nested models.

To identify features most useful in distinguishing HIV infection, the EN model was retrained on all the data, to find features with non-zero weights. Finally, multivariate imputation using chained equations (MICE) was explored to investigate the effect of increased sample size on classification and feature selection.

Results: The best performing modality in the single modality analysis was sMRI volumes ($AUC = 0.71$) followed by DTI ($AUC = 0.62$). EN models on spectroscopic, gyrification and cortical thickness measures showed no class discrimination capability ($AUC \approx 0.5$). Adding DTI measures to sMRI volumes increased AUC to 0.78 and adding ^{1}H-MRS in basal ganglia to this combination produced the highest classification performance($AUC = 0.80$). However, the selected spectroscopy features increased the likelihood of information loss by our multimodal prediction model. Therefore, the final multimodal signature consisted of 22 DTI and sMRI volume features of cART-treated paediatric HIV infection, including reduced volume of the bilateral globus pallidus and amygdala, as well as reduced MD and RD in corticospinal tracts. Furthermore, AUC of imputed feature sets were equivalent to that of the non-imputed set, but signatures of imputation were inconsistent with the model that had the smaller sample size.

Discussion: Volumes obtained from sMRI on their own provide reasonable discrimination between HIV+ and HIV- children. The addition of measures from DTI and ^{1}H-MRS in basal ganglia present only marginal improvements to classification accuracy. While the ability to predict HIV-status based on neuroimaging is not clinically advantageous particularly in resource-limited settings, demonstration of its feasibility could help identify neuroimaging signatures that are common to treated HIV infection and offer a separate yet complementary dimension to the clinical diagnosis.

PREFACE

The work presented in this dissertation was undertaken at the Medical Imaging Research Unit of the University of Cape Town under the expert guidance of my supervisor and mentor Dr Frances Robertson. Methods were approved by Faculty of Health Science Human Research Ethics committee (HREC REF: 123/2019 sub-study linked to 448/2011). Participants were scanned at Cape Universities Body Imaging Centre following appropriate protocols.

Funding for this research was through grant R01-HD071664 from the U.S. National Institute of Health, University of Cape Town international conference travel scholarships for MSc degree research, and National Research Foundation grant for South African Research Chairs Initiative of the Department of Science and Technology.

The only way I could see this far is because I first stood on the shoulders of Dr Frances Robertson; I am delightfully indebted to you for years to come. I also appreciate the unwavering support and contribution to this work made by Dr Marcin Jankiewicz and Dr Martha Holmes as well as the statistical consultations I received from Prof Francesca Little, and structured programming solutions from Dr Keri Woods.

I am hopeful this dissertation will be a valuable reference to the Meintjes MRI Lab, research community, educators, and healthcare professionals.

—Isaac Lebogang

TABLE OF CONTENTS

Title page | Declaration | Dedication | Synopsis | Preface | Illustrations |
Works Cited | Appendix | Ethics

INTRODUCTION

Problem definition and dissertation expectactions — Page 1

BACKGROUND

HIV, cART and essentials of MR neuroimaging — 5

LITERATURE REVIEW

Feature selection and multimodal neuroimaging — 19

METHODOLOGY

Identifying multimodal HIV signatures — 30

FINDINGS

Classification performance and feature selection — 48

DISCUSSION & CONCLUSION

Signification of results, limitations and recommendations — 66

ILLUSTRATIONS

List of tables

	Page
Table 2.1| HIV prevalence in the top six countries	6
Table 3.1| Literature review matrix of feature selection	24
Table 4.1| Sample characteristics of the study population	32
Table 5.1| A list of participants with useable data in each modality	49
Table 5.2| Common participants across modalities	49
Table 5.3| Classification performance evaluation metrics of the 7-MRI derived feature sets	50
Table 5.4| Selected features for 1H-MRS in pwm classifier	53
Table 5.5| Selected features for ^{1}H-MRS in mfgm classifier	53
Table 5.6| Selected features for ^{1}H-MRS in bg classifier	54
Table 5.7|Selected features for sMRI cortical thickness classifier	54
Table 5.8| Selected features for sMRI gyrification classifier	55
Table 5.9| Multimodal classification performance in steps of improving AUC	56
Table 5.10| Classification performance metrics of non-nested multimodal analyses	56
Table 5.11| Selected features of sMRI volumes + DTI + ^{1}H-MRS bg	58
Table 5.12| Comparison of non-nested multimodal classifiers with AIC	59
Table 5.13| Classification performance of nested models	60
Table 5.14| Classification performance metrics of best performing nested sequence	60
Table 5.15| Likelihood ratio test of nested models	61
Table 5.16| Comparison of nested and non-nested models with AIC models	62
Table 6.17| Classification performance of imputed feature set with non-imputed model	64
Table 6.18| Selected features of best performing model that are also selected	65

***

List of figures

	Page		Page
Figure 2.1| Blood-brain barrier with three main neuronal supporting structures	7	Figure 2.2| Multimodal MR Neuroimaging (MR contrasts)	7
Figure 2.3A| Proton precession around B0 & mathematic relationship	10	Figure 2.3B| FID after RF pulses are switched off	10
Figure 2.4| In vivo ^{1}H-MRS neurochemical profile	11	Figure 2.5| 2D MR images of brain tissue visualised in axial, sagittal, and coronal planes	11
Figure 2.6| Diffusion tensor oriented by eigenvectors and shaped by eigenvalues	15	Figure 3.1| Wrappers vs Filters	15
Figure 4.1| Treatment arms for HIV+ asymptomatic infants who were part of the CHER trial	31	Figure 4.2| Voxel placement shown in sagittal, coronal, and axial planes	31
Figure 4.3| Overview of the generation of feature sets	35	Figure 4.4| 10-fold cross validation scheme	35
Figure 4.5| A logistic EN regression with 100 iterations of 10-fold CV	39	Figure 4.6| Feature selection with CV optimised parameters	39
Figure 4.7| Evaluation of selected features during 100 repeat cycle	41	Figure 4.8| Creation of nested models with sMRI volumes as first set	41
Figure 5.1| Selected volume structures in situ	51	Figure 5.3| Selected features of sMRI volumes + DTI in situ	51
Figure 5.2| Selected DTI tracts in situ	52	Figure 5.4| Feature selection of nested multimodal classification	52

***

Abbreviations and definitions

MR – magnetic resonance

DTI – diffusion tensor imaging

sMRI – structural MR imaging

^{1}H-MRS – proton MR spectroscopy

cART – combination antiretroviral therapy

AD – axial diffusivity

RD – radial diffusivity

MD – mean diffusivity

FA – fractional anisotropy

WM – white matter

ROC – receiver operating characteristic

ROI – region of interest

AUC – area under the ROC curve

LASSO – least absolute shrinkage and selection operator

EN – elastic net regularisation

HIV – human immunodeficiency virus

HIVE – HIV encephalopathy

HAND – HIV-associated neurological disorders

PML – progressive multifocal leukoencephalopathy

CNS – central nervous system

BBB – blood-brain barrier

CSF – cerebrospinal fluid

bg – basal ganglia

pwm – peritrigonal white matter

mfgm – midfrontal gray matter

CPE – CNS penetration effectiveness score

CHER – children with early antiretroviral therapy

sd – standard deviation

***

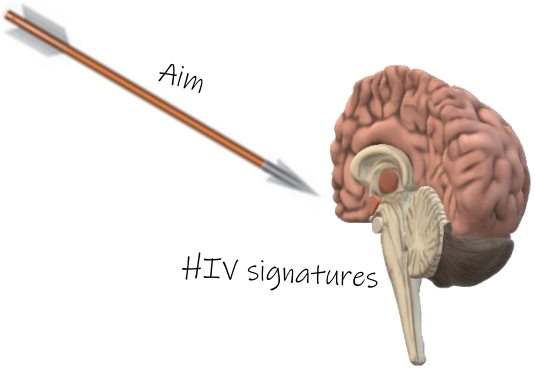
Biomarker or neurological marker – naturally occurring characteristic that identifies a physiological process

Feature selection – in machine learning and statistics, measured variables are either:

- Relevant: Informative measurements for the learning algorithm
- Irrelevant: Not informative measurements; no information is lost when removed
- Redundant: Not informative in the presence of relevant measurements

Variable/subset/attribute selection is the process of identifying relevant features

Signature – used in this thesis as a biomarker consisting of a combination of features that is characteristic of a pathology e.g. HIV or cancer



1

Introduction: problem definition and dissertation expectations

Children may be perinatally infected with human immunodeficiency virus (HIV) via transplacental exposure *in utero*, exposure to maternal blood during delivery, or ingestion of the virus during breast feeding. HIV then attacks and kills CD4+ T cells compromising the ability of the body to fight other infections and cancerous cells. When left untreated, HIV-infection results in poor health, and children commonly succumb to opportunistic infections. However, current management of HIV through combination antiretroviral therapy (cART), comprising drugs that interfere with HIV binding, fusing, and entering host cells, has improved clinical outcomes of children worldwide.

A standard diagnosis of HIV infection in a child is obtained through a comprehensive laboratory blood test which reveals how well the immune system is functioning (CD4 count), the rate of rapid replication of the virus (viral load), and whether they have a drug resistant strain of HIV. But the test does not provide information about the possible harm that HIV has on the brain, nor is a prognostic tool for subsequent neurocognitive deficits. Unmanaged HIV does have severe effects on brains of children, which include calcification of the basal ganglia, cortical and cerebral atrophy, and reactive gliosis (George et al., 2009; Kieck et al., 2004). There is literature suggesting that neurocognitive and behavioural profile of children on cART is still within the low average range (e.g. Nichols et al., 2016, 2015; Garvie et al., 2014; Laughton et al., 2012), and there are alterations on neuroimaging compared to the general HIV-uninfected (HIV-) paediatric population. Up to 50% of HIV-infected (HIV+) adult patients continue to show milder forms of HIV-associated neurocognitive disorders (HAND) despite being on cART (Saylor et al., 2016). This is in part because not all cART drugs cross the blood-brain barrier (BBB) equally well.

Because of advances in neuroimaging, researchers are now able to examine *in vivo* alterations in cortical and subcortical brain structures that result from neurological diseases. These alterations include differences in measures of brain volume, cortical thickness and folding, diffusion of water molecules in white matter tracts, connectivity between gray matter (GM) regions, network organisation, and concentrations of neurochemicals. Different studies have found various neuroimaging alterations in cART-treated HIV, but it remains unclear which is most characteristic of HIV or whether there is a pattern of combined neuroimaging features that typify HIV infection. Neuroimaging can also be a powerful tool to supplement clinical

examinations, such as neurocognitive testing of HIV+ children, and may offer information about possible future outcomes of these patients.

Previous studies in paediatric HIV have looked at neuroimaging modalities separately and analysed the data using traditional statistical models, for example, analyses of variance and t-tests. However, these models often do not accurately discriminate between HIV-infected and uninfected children. No study to date has taken a combined approach of looking at several MRI modalities together and employing multivariate classification to discriminate cART-treated children from uninfected controls. Therefore, the aim of this dissertation was to investigate the effectiveness of multimodal[1] neuroimaging to predict HIV-status, thereby identifying characteristic patterns of neuroimaging features (signatures) of cART-treated paediatric HIV.

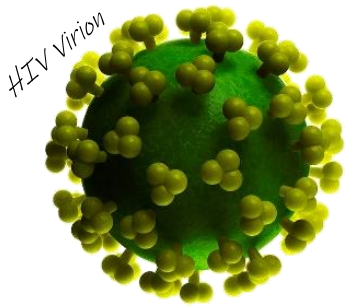
To achieve this aim, we used neuroimaging data including structural MRI (sMRI), diffusion tensor imaging (DTI) and magnetic resonance spectroscopy (^{1}H-MRS) from 7-year-old children from the Children with HIV Early Antiretroviral Therapy (CHER) trial on cART whose viral loads were suppressed from between 6 and 76 weeks of age. In precise terms, we asked these questions:

- What are signatures of cART-treated paediatric HIV in each individual MRI modality?
- How well do signatures from a single MRI modality discriminate between HIV+ and HIV- groups?
- What are signatures of cART-treated paediatric HIV in multimodal combinations of neuroimaging features?
- Does multimodal neuroimaging improve classification performance compared to single modality MRI?
- What can multimodal neuroimaging signatures of cART-treated paediatric HIV tell us about the effects of HIV on the brain of these children?

[1] Multimodal neuroimaging here is used in the sense of combining a spectrum of contrasts from one system, namely a magnetic resonance system with different MR contrasts. This is counter to the sense of hybrid imaging where different types of scans, such as MRI, ultrasound (US), positron emission tomography (PET), single photon emission computed tomography (SPECT) are merged, for example US-MRI, PET-MRI, PET-CT, SPECT-CT.

We hypothesised that a multivariate, multimodal approach discriminates better between HIV+ children on cART and uninfected controls than single modality analyses **(hypothesis I)**. And that a subset of MR neuroimaging measurements that are characteristic of treated perinatally acquired HIV are indeed identifiable at this age **(hypothesis II)**.

The succeeding chapters of the dissertation are structured as follows: Chapter 2 presents relevant background for HIV, cART, principles of MR neuroimaging, and previous findings in paediatric HIV. Chapter 3 describes feature selection and multimodal neuroimaging, and how they are applied to improve classification performance and identify signatures of various diseases. Chapter 4 details the implementation of classification and selection algorithm used during training and validation of our prediction models. Chapter 5 presents our findings. And the final chapter discusses the significance and implications of the identified signatures of cART-treated paediatric HIV, limitations of the dissertation, and makes recommendations for future work.



2

Background: HIV, cART & essentials of MR neuroimaging

Contents

2.1. Reach of the human immunodeficiency virus (HIV) 6
2.2. HIV-related neurophysiology 7
2.3. Neuroimaging: magnetic resonance (MR) 8
2.3.1. Proton magnetic resonance spectroscopy (^{1}H-MRS) 10
2.3.2. Structural imaging (sMRI) 12
2.3.3. Diffusion tensor imaging (DTI) 13
2.4. Neuroimaging analysis 16
2.4.1. Single modality analyses of MRI data in perinatal HIV infection 16
2.4.2. Multimodal MRI analyses 17

Chapter 2 Overview

- HIV is a public health crisis in Sub-Saharan Africa. It is an epidemic in South Africa, and highly prevalent in countries bordering it.
- Combination antiretroviral treatment has extended lifespans and improved cognitive outcomes of patients.
- ^{1}H-MRS determines relative concentrations and physical properties of a variety of biochemicals, and it can be used to study metabolic activity in brain tissue.
- DTI measures diffusion of water in the CNS, and extensively maps white matter tracts.
- sMRI provides anatomical and pathological overview of the brain; it allows gyrification indices, cortical thicknesses, and subcortical volumes to be measured.
- Multimodal MR studies are gaining traction in the neuroimaging field.

2.1. Reach of the human immunodeficiency virus (HIV)

Since its first recorded cases in the 1980s, HIV has become a far-reaching public health concern in continental Africa, particularly in south Sub-Saharan Africa where it is the most prevalent (UNAIDS, 2018). The Kingdom of Eswatini (Swaziland) is the leading country with HIV prevalence among the southern countries (top six nations are listed in Table 2.1), however the Republic of South Africa (RSA), population estimated at 57.9 million people, has the largest HIV epidemic with over 7.5 million people living with the illness (Stats SA, 2018).

Table 2.1| HIV prevalence in the top six countries (UNAIDS, 2018)

Country	Prevalence	People
Swaziland	27.2%	220 000
Lesotho	25.0%	330 000
Botswana	21.9%	360 000
Namibia	13.8%	230 000
Zimbabwe	13.5%	1 300 000
RSA	13.1%	7 562 630

As a result, 90% of perinatally HIV-infected children reside in these countries although new infections are rapidly declining (UNICEF, 2018). Statistics South Africa (2018) reports a perinatal prevalence of 3.1% in the country: a population of 230 000 children aged 0—14 years. About 85% of HIV cases in children are vertically transmitted from mothers, while others are due to blood transfusion, for example, children with haemophilia. Therefore, there exists an increased risk of infection for children born to HIV+ mothers, but due to proliferation and success of programs that prevent mother-to-child transmissions, many of these HIV-exposed children are themselves HIV- and tend to be healthier than their HIV+ counterparts.

With the advent of combination antiretroviral therapy (cART) as well as changes in guidelines and access to treatment, HIV is no longer a death sentence. The risk of mother-to-child transmission can be reduced to 2% with a comprehensive-treatment regimen. For the past ten years, cART has been introduced in infancy and early childhood to suppress viral load and slow down progression of the virus to the fatal acquired immunodeficiency syndrome—in most cases, patients live to national average life expectancy. Despite the total number of children receiving cART doubling from 355 000 to 740 000 worldwide between 2009 and 2013,

the World Health Organisation (WHO, 2017) estimates only less than a quarter (21—25%) of children in low-and middle-income countries are on treatment.

Untreated HIV is a detriment to the physical and mental wellbeing of children worldwide and may affect their development and acquisition of critical skills. As such, access and adherence to treatment is of vital importance. cART has minimised poor health and the threat of children succumbing to opportunistic infections. Understanding the mechanism HIV and its impact on the central nervous system (CNS) is pivotal to effective synthesis, delivery and effectiveness of cART drugs.

2.2. HIV-related neurophysiology

The human brain has a diffusion boundary, the blood-brain barrier (BBB), to impede an influx of compounds from blood to brain tissue and cerebrospinal fluid. In short, it is composed of three main neuronal supporting cell structures: astrocyte end-feet, BBB-endothelial cells, and pericytes (Figure 2.1), which work in tandem to exclude entry of blood-borne substances into the brain (Ballabh et al., 2004b). According to one 'trojan horse' theory, HIV-1 can cross the BBB inside circulating monocytes which are then activated and transformed into macrophages releasing virions into brain tissue (Izquierdo-Useros et al., 2010; Verani et al., 2005; Sharer et al., 1990). BBB disruption may lead to HIV-related encephalitis (HIVE) and HIV-associated dementia (HAD), which are two major CNS complications in individuals with HIV infection (Buckner et al., 2006; Toborek et al., 2005).

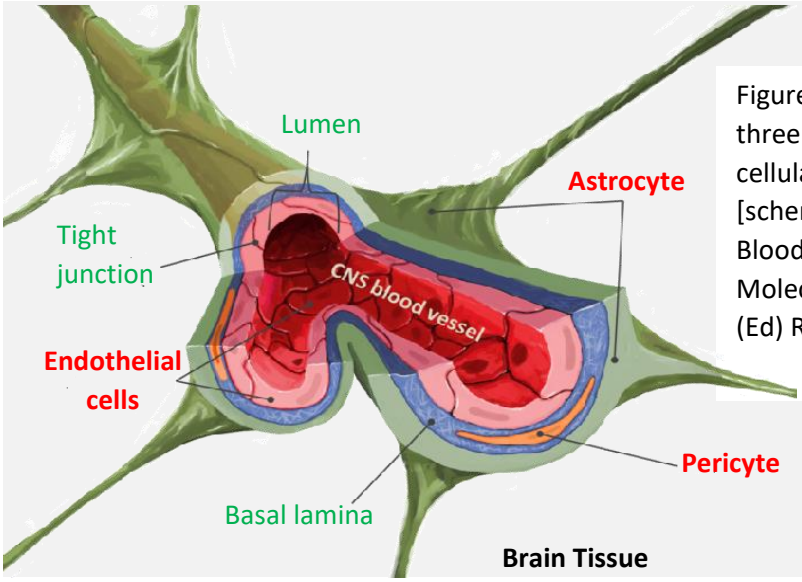


Figure 2.1| Blood-brain barrier with three main neuronal supporting cellular structures (in red)— [schematic adapted from: The Blood-Brain Barrier Cellular and Molecular Biology, Pardrige, W.M. (Ed) Raven Press]

The measure for the capacity of antiretroviral drugs to penetrate the BBB and inhibit viral replication is given by their CNS penetration effectiveness scores (CPE). For example, Zidovudine and Indinavir-r have a high CPE, while Nelfinar and Efavirenz have low scores (Caniglia et al., 2014; Letendre et al., 2008). But typically, large compounds are prevented entry. Therefore, even with a cART regimen that maintains reasonable peripheral viral-load suppression, there is still some form of viral build-up in the brain. With improved cART and development of regimens with higher CPE, HIVE has shifted from a subacute to a chronic condition, while HAND remains common among many patients (Kovalevich et al., 2012).

Moreover, Shah and colleagues (2016) reported that neurotoxic effects of cART itself may lead to neuronal damage. It is yet to be fully known whether there are long-term neurological effects even if cART is started early in the disease course, or how the presence of HIV-1 in brain reservoirs affects cognitive processes such as attention, language, memory, and fine motor skills. The question remains unsettled whether cART reverses the harmful effects of HIV on brain development of children. Neuroimaging (e.g. Ackermann et al., 2019; Robertson et al., 2018; Jankiewicz et al., 2017) and neurocognitive (e.g. Hoare et al., 2012; Laughton et al., 2012) studies are attempting to answer these questions, which are key for monitoring brain development in the presence of HIV infection and cART.

2.3. Neuroimaging: magnetic resonance (MR)

Neuropsychological (NP) testing is a standard approach in the examination of brain functionality and effects of neurological disorders such as Parkinson's disease and dementia including the spectrum from mild cognitive impairment to Alzheimer's disease (Zucchella et al., 2018). Neurocognitive tests have similarly been used to classify HIV-associated neurological disorders in adults into 3 stages of severity (Saylor et al., 2016; Clifford & Ances, 2013): asymptomatic neurological impairment—initial stage of infection; mild neurocognitive disorder—progression stage; and HAD, the most severe stage characterised by general loss of memory. However, NP testing require inferences of clinicians about personality and state of mind of the patient. Administering NP tests in very young children is a challenging task, and the outcomes themselves do not provide insight into pathological brain changes that underlie poor test performance.

Neuroimaging techniques such as positron emission tomography (PET), computerised tomography (CT), single-photon emission computed tomography (SPECT), brain ultrasonography (US), magnetoencephalography (MEG), and magnetic resonance (MR) imaging (Schouten et al., 2016; Blakemore, 2012; Johnson et al., 2012; Zhang et al., 2011) are becoming more widely used to examine alterations in brain structure and function. The latter modality is commonly used because of its non-invasiveness, high signal-to-noise ratio (SNR), and high-resolution.

With its broad spectrum of contrasts (Figure 2.2), MR neuroimaging (brain MRI) allows different properties of tissue to be examined, using for example, diffusion tensor imaging to look at white matter microstructure and its connectivity, MR spectroscopy to examine local brain metabolism, resting-state and task-based functional imaging (rs/tr- fMRI) to study brain function, and anatomical (T_1-weighted, structural) imaging to investigate brain morphometry. This dissertation uses spectroscopic, structural, and diffusion MR neuroimaging of the brain of HIV+ and uninfected controls. These imaging techniques are described briefly below.

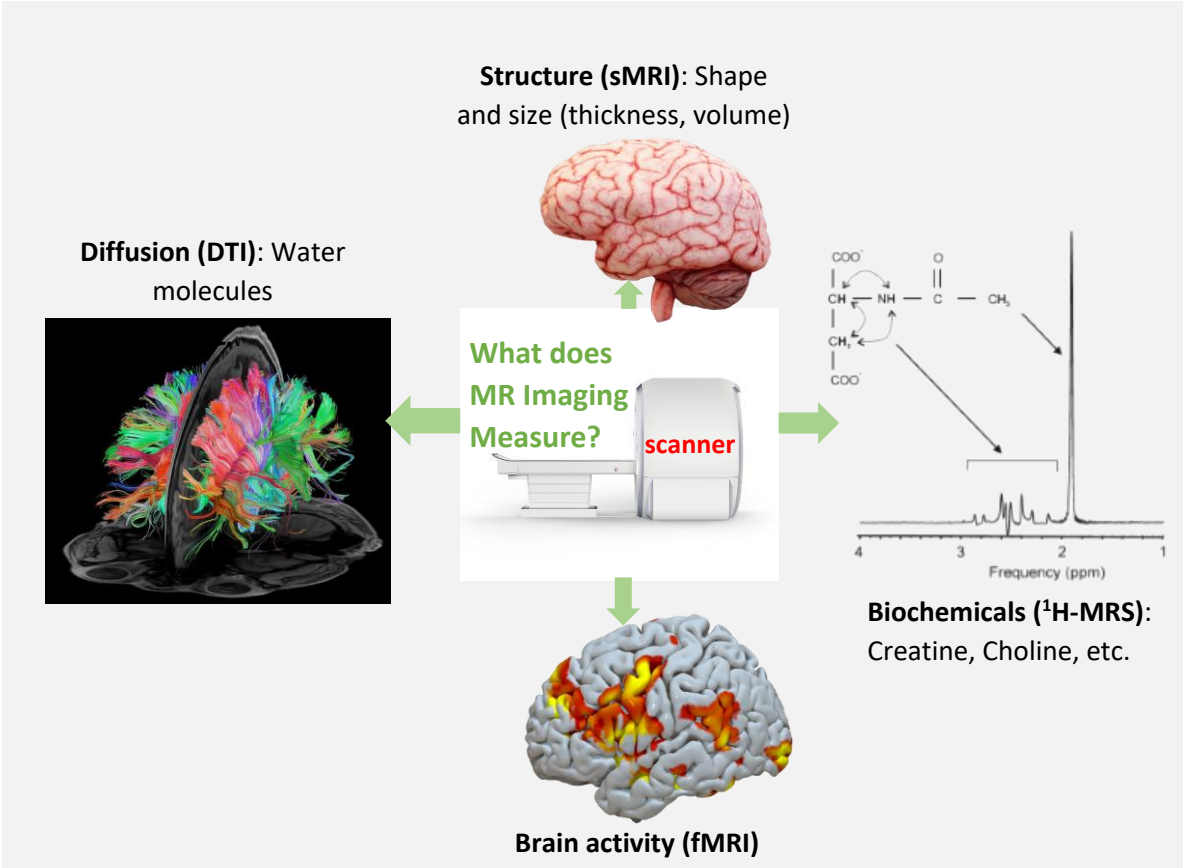


Figure 2.2| Multimodal MR neuroimaging (MR contrasts)

2.3.1. Proton magnetic resonance spectroscopy (^{1}H-MRS)

The proton (^{1}H) is abundant in living tissue because of its presence in water molecules, therefore highly sensitive to detect. Its dipole moment (μ) precesses around an external magnetic field (B_0) at an angular frequency (the Larmor frequency, ω_0) that is dependent on the strength of B_0 and chemical environment (σ_{cs}) which shields it from the field (Figure 2.3A). If radiofrequency (RF, ω_f) pulses are applied to the proton at Larmor frequency i.e. $\omega_0 = \omega_f$, it absorbs the energy (E) and moves from ground-to-excited state. This is known as magnetic resonance, and is the basis for all MR neuroimaging (Drost et al., 2002; Sorenson, 1990). ^{1}H gradually relaxes back to equilibrium when pulses are switched off, thereafter emits a time-domain free induction decay (FID) signal that can be Fourier transformed (FT) to a frequency-domain absorption spectrum (Figure 2.3B) in a process known as MR spectroscopy.

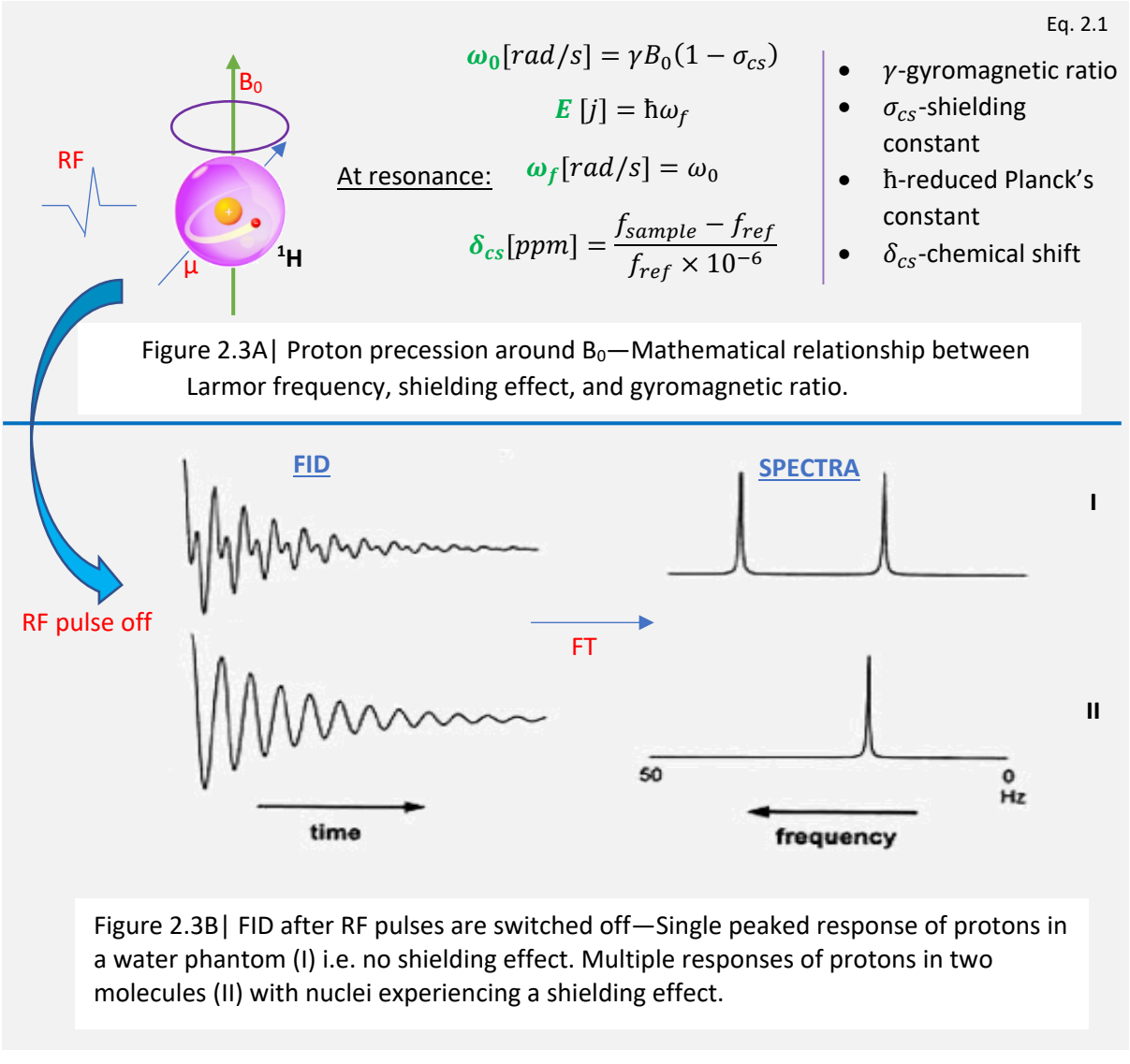


Figure 2.3A| Proton precession around B_0—Mathematical relationship between Larmor frequency, shielding effect, and gyromagnetic ratio.

Figure 2.3B| FID after RF pulses are switched off—Single peaked response of protons in a water phantom (I) i.e. no shielding effect. Multiple responses of protons in two molecules (II) with nuclei experiencing a shielding effect.

Chemical shifts (δ_{cs}) are resonant frequencies of ^{1}H nuclei relative to a standard in the presence of B_0. ^{1}H-MRS employs δ_{cs} to estimate the relative concentrations of ^{1}H-containing metabolites within brain tissue. By selectively exciting ^{1}H within a region of interest, a neurochemical profile (Figure 2.4) of that region can be obtained. Water at 4.7ppm is often chosen as the reference frequency (f_{ref}) relative to which the frequencies (f_{sample}) of other metabolites are shifted (Blüml & Panigrahy, 2013). Changes in metabolite concentration level have clinical significance—their decrease/increase can be of use in examining metabolic effects of HIV/cART as listed in Figure 2.4.

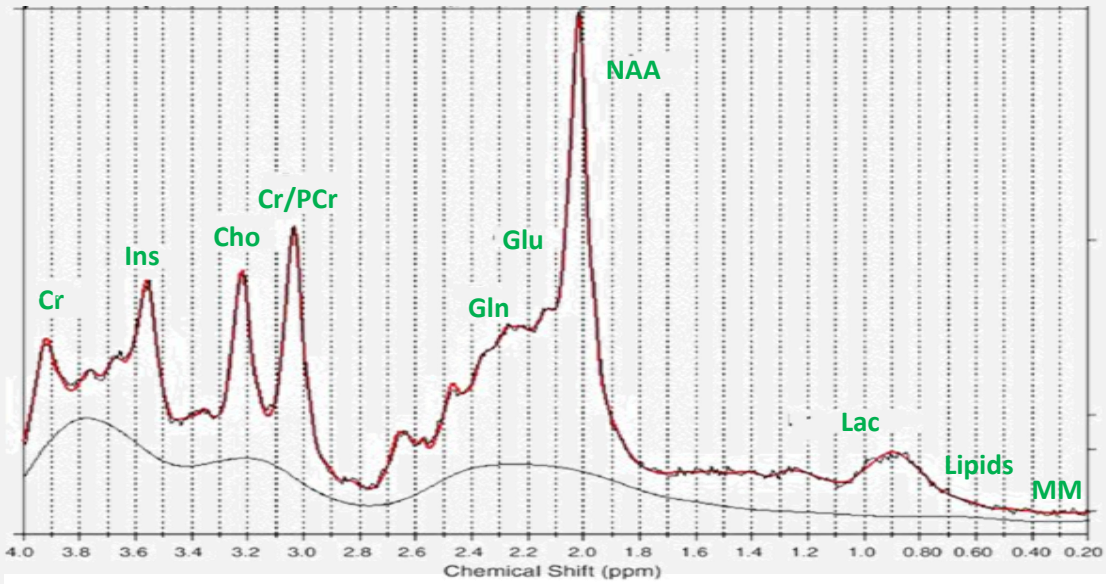


Figure 2.4| In vivo ^{1}H-MRS neurochemical profile—Shows chemical shifts at which target metabolites resonate [Spectra taken from LCModel]. Below is what their concentrations typically signify as reviewed by Bertholdo (2013), unless specified otherwise.

Creatine (Cr)	Categorised as the most stable cerebral metabolite, thereby used to calculate ratios of other metabolites.
Macromolecules (MM)	Biases accurate quantification of other metabolites. MM changes may possess clinical relevance (Považan et al., 2015)
Choline (Cho)	Increases in concentration during inflammation and infarction (tissue death due to inadequate blood/oxygen supply).
Lipids (Lip)	Seen in a brain with necrosis or primary malignant tumours.
Myo-inositol (mI)	Indicates proliferation of glial cells (gliosis) and myelin degradation. An increase of mI occurs in inflamed tissue.
Alanine (Ala)	Increased concentration signifies presence of meningioma and defects in oxidative metabolism.
Lactate (Lac)	Elevated concentration in cystic tumours, macrophage accumulation as well as normal pressure hydrocephalus.
N-acetylaspartate (NAA)	Indicates neuronal and axonal density. Loss of neurons are signalled by a decreased concentration.
Glutamate/Glutamine (Glu/Gln)	Elevated concentration often found in hepatic encephalopathy and alluded to in effects of HIV

2.3.2. Structural imaging (sMRI)

Using magnetic field gradients for spatial encoding, the measured MR signal can be converted using a FT to a matrix of pixels with a range of intensity levels to produce detailed anatomical MRI images (Figure 2.5). Different types of images are created by varying the time between successive RF pulses (repetition time, TR) and time between their delivery to reception of the echo signal (time to echo, TE). The longitudinal and transverse relaxation times ($T1/T2$) of tissue dictate the rate of decay of the MR signal as phase coherence is lost and ^{1}H relaxes to equilibrium. Relaxation times are especially important in anatomical MR neuroimaging as they make it possible to visually distinguish various kinds of soft tissue (Sorenson, 1990), for example, white matter ^{1}H relaxes faster (short $T1$) than cerebrospinal fluid (CSF, long $T1$) therefore has brighter pixels on T1-weighted images (short TR, short TE), compared to those of CSF which are dark.

Once MR images have been acquired in a certain orientation, they can be registered to a template derived from a pool of subjects in the study or a brain atlas (spatial normalisation), segmented, labelled, and used for brain morphometry. Morphometric measures that can be obtained from sMRI are cortical thickness, gyrification or cortical folding, and volumes of subcortical structures. There is a vast body of literature showing these measures are highly correlated with neurodevelopmental and neurodegenerative conditions and their subsequent effects on cognitive function, for example, fine motor skill (dexterity) and preclinical dementia (Spalleta & Fabrizio, 2018; Kotov, 2017; Trefler et al., 2016). Studies have reported associations of cortical and subcortical morphometry (e.g. Kallianpur et al., 2016; Thompson et al., 2005; Mitchell, 2001) summarised in Figure 2.5.

Furthermore, sMRI measures work in tandem and therefore difficult to isolate individual markers of cognitive function. White (2002) did a morphological study of twins and found cortical gyrification to be primarily determined by environmental factors whereas subcortical volumes are highly determined by genetic variations.

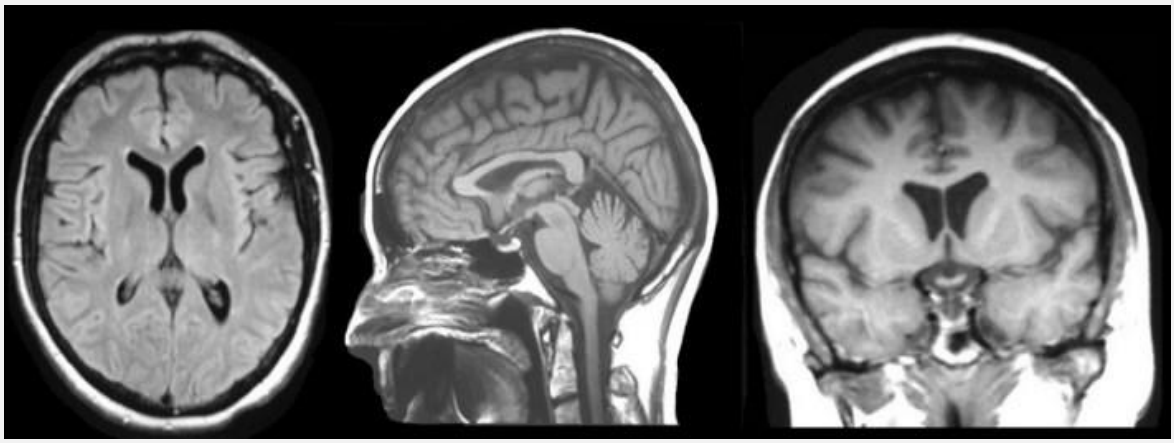


Figure 2.5| 2D MR images of brain tissue visualised in axial, sagittal, and coronal planes—Pixel intensity is characterised by longitudinal relaxation time ($i.e. T_1\text{-}weighted$); CSF (dark pixels), gray matter (intermediate), and WM (bright). [FreeSurfer]

Cortical Thickness (CT)	Closely associated with motor control: Cortical thinning highly correlates with immune system deterioration (Thompson et al., 2005) and psychomotor speed (Kallianpur et al., 2012)
Gyrification indices (LGI)	LGIs are positively associated with fine-tuning process of neural circuitry or maturation (White et al., 2002)
Subcortical volumes (SV)	A direct measurement of neurodegeneration (Racosta & Kimpinski, 2019), and linked with regional and neural atrophy and viral load (Mitchell, 2001)

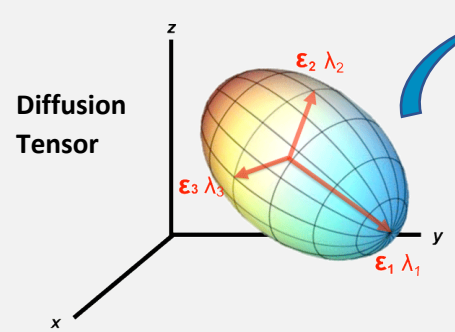
2.3.3. Diffusion tensor imaging (DTI)

The extent to which early cART can prevent or limit HIV-associated white matter (WM) damage can be investigated by DTI *in vivo*. Diffusion of water molecules in fibrous tissues including WM is anisotropic, highly hindered in directions perpendicular to fibres and relatively unimpeded parallel to the fibre orientation. While the primary barrier for free (isotropic) water diffusion in WM is myelin sheath, for an infected brain cellular swelling or increased cellular density increases diffusion in all directions.

Anisotropic diffusion behaviour can be described by a diffusion tensor (Basser et al., 1994), the covariance matrix of a 3D-gaussian distribution of spin displacements, which is symbolically visualised as an ellipsoid (see illustration in Figure 2.6). The MR signal is sensitive to diffusion due to dephasing from Brownian motion in the presence of a pair of magnetic field gradient (Hahn, 1950). In the MR scanner, a diffusion tensor is measured by imaging diffusion in at least six non-collinear individual gradient directions i.e. D_{xx}, D_{xy}, D_{xz}, and so on; the exact relationship between MR signal and spin diffusion was initially explained in depth by Stejskal and Tanner (1965) with the introduction of pulsed gradient spin-echo (PGSE).

To date, the extensive use of diffusion microstructure imaging has been in examining WM integrity in the presence of various neurological diseases. Both degree of anisotropy and local fibre direction can be mapped using DTI. Standard anisotropy measures are fractional anisotropy (FA)—a measure of degree of directional dependence of the tensor ranging from 0 (isotropy) to 1 (maximum anisotropy); axial diffusivity (AD)—diffusion along the primary orientation axis of the tensor; radial diffusivity (RD)—diffusion perpendicular to the primary orientation axis; and mean diffusivity (MD)—average rate of diffusion over all directions. These measures can reflect axon density, axonal diameters, intra/extra-cellular exchange times, and properties of the extra-cellular space (Tournier et al., 2019, 2011)—some common interpretations are listed in Figure 2.6. The orientation of the diffusion tensor may also be used to describe the paths of WM tracts in the brain, in a technique known as tractography. This allows non-invasive investigation of brain connectivity.

DTI has been used to study neurodevelopment, however, the extent to which WM changes with age during childhood and adolescence appears to depend on the used analysis method—namely, voxel-based (e.g. Barnea-Goraly et al., 2005; Schmithorst et al., 2002) or ROI-based analysis (e.g. Schneider et al., 2004). Voxel-based methods make voxel-by-voxel statistical comparisons throughout the brain, while an ROI method defines an area of interest in the brain to make measurements (Snook et al., 2007, 2005).



Eq. 2.2

$$\boldsymbol{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$$

Well-established anisotropy **measures**:

$$\boldsymbol{AD} = \lambda_1$$

$$\boldsymbol{RD} = \frac{1}{2}(\lambda_2 + \lambda_3)$$

$$\boldsymbol{MD} = \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3)$$

$$\boldsymbol{FA} = \sqrt{\frac{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}$$

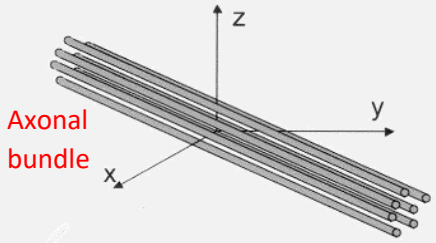


Figure 2.6| Diffusion tensor oriented by eigenvectors (ε_1, ε_2, ε_3), and shaped by eigenvalues (λ_1, λ_2, λ_3). [Tournier, 2019]. Mathematical relationship between anisotropy measures

Fractional Anisotropy (FA)	Increases with age and is associated with an increase in myelination, neural density and axonal count. (Brouwer et al., 2010 ; Lentz et al., 2014)
Mean diffusivity (MD)	Measure of organisation: Lower MD values are associated with dense axonal packing and well-organised structure. (Feldman et al., 2010)
Radial diffusivity (RD)	Decreases with age and is associated with increased myelination and/or densely packed axons. (Lebel and Beaulieu, 2011)
Axial diffusivity (AD)	Increases with age and associated with fibre coherence as axons straighten and merge within the bundle (ibid.)

***

Beyond the scope of this discussion are acquisition sequences referred to in data acquisition (Chapter 4.2) which include point resolved spectroscopy sequence (PRESS, Bottomley, 1984), chemical shift selective saturation (CHESS) for water suppression (Haase et al., 1985), twice-refocused spin echo sequence (TRSE), echo planar imaging (EPI), and multi-echo magnetisation prepared rapid gradient echo (MEMPRAGE) sequences (van der Kouwe et al., 2008). If need be, all sequences are referenced for further reading.

2.4. Neuroimaging analysis

Statistical models can be constructed from MR neuroimaging data for further analyses to quantify relationships between variables—for example, in classification, regression, compression, and data mining. The most common statistical approach to neuroimaging analysis is a 'mass univariate' one, where a single modality at a time is considered, usually at each voxel in the brain. This is the basic idea behind voxel-based morphometry (Ashburner & Friston, 2000; Kurth et al., 2015) and statistical parametric mapping (Rees, 2004).

2.4.1. Single modality analyses of MRI data in perinatal HIV infection

Uban and colleagues (2015) used linear-mixed effect modelling of DTI scans and found lower whole brain and regional FA in HIV+ children and adolescents on cART initiated after age 2. Likewise, Hoare et al. (2015) employed analysis of co-variance (ANCOVA) and found association between disease severity with AD and MD in forceps minor, including the corpus callosum, fornix, corona radiata, frontal and parietal WM, pre-/postcentral gyrus, and superior longitudinal fasciculus.

In spectroscopy, Pavlakis and colleagues (1995) used student t-tests to report reduced ratio of N-acetylaspartate to creatine in the basal ganglia of HIV+ children compared to controls. Other correlation analysis studies in sMRI found microcephaly and gradual loss of motor function skills in HIV+ children (Ellis et al., 2010; Mitchell, 2001). Other authors, for example Wade et al. (2019), used fixed effects multivariate linear regression analyses to model associations among the subcortical shape features and found that the basal ganglia structures were the most affected by perinatal HIV infection as well as that stable infection has only minor effects on brain structure.

Previous studies published by our group have examined group differences in various neuroimaging modalities separately using simple linear regression models. The cohort from the children with HIV early antiretroviral therapy (CHER) trial (described in Chapter 4.1) were scanned at ages 5, 7, 9, and 11 years. At age 7, findings from individual modalities in the cohort are summarised as follows:

^{1}H-MRS: No differences in basal ganglia neurometabolites concentrations between HIV+ and control group were found. (Robertson et al., 2018).

DTI: Lower FA and higher MD in the left inferior longitudinal fasciculus and fronto-occipital fasciculus were found in HIV+ children compared to controls. These changes were attributable to higher RD in HIV+ children. (Jankiewicz et al., 2017).

sMRI: Volumes of gray and white matter, bilateral hippocampus and putamen, and right thalamus were smaller in HIV-positive children compared to controls. A left inferior lateral occipital region showed thicker cortex and bilateral paracentral gyri showed lower gyrification in HIV-positive children compared to uninfected controls. (Nwosu et al., 2018).

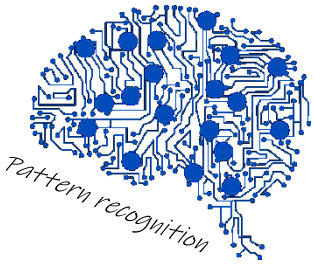
Although in such single modality studies differences in the mean value of brain properties may be found between groups—for example, fractional anisotropy measured using DTI in a region may be lower in HIV+ children than controls — a single measure usually discriminates inadequately between these groups. Furthermore, standard statistical models may result in inaccurate predictions for individual subjects and other study populations. In addition, the large number of measures generated in a typical neuroimaging study may make it difficult to identify the most salient neuroimaging features of HIV infection.

2.4.2. Multimodal MRI analyses

In its simplest form, multimodal imaging is using different contrasts or modalities, and mapping the complementary data of individual modalities into a common space for analysis and evaluation. It is a technique that is increasingly gaining steam in neurological investigations where measurements (features) from single modalities are combined and analysed simultaneously using advanced statistical learning. Recent research in Alzheimer's disease, mania, schizophrenia, and obsessive-compulsive disorder have shown that multimodal and multivariate techniques are useful and result in accurate class predictors (Cabral et al., 2016; Schouten et al., 2016; Ritter et al., 2015).

To our knowledge there are a handful of multimodal neuroimaging studies in adults with HIV. The first study combined structural imaging and a mechanoreception task during magnetoencephalography to probe patterns in older adults aged 50 to 70 years by using one-sample t-tests (Wilson et al., 2015). Another closely related approach evaluated basic motor control in HIV+ adults ($age = 38.2 \pm sd\ 10.7$ years) by combining sMRI and task-related fMRI (Zhou et al., 2017). Finally, a study by Samboju and colleagues (2018) combined DTI and resting-state fMRI ($age = 30 \pm\ sd\ 9.9$ years). The aim of these studies was to find links between structural and functional brain integrity in the early stages of HIV. In addition, two recent studies applied a multivariate pattern recognition approach within a single neuroimaging modality to predict HIV status based on respectively sMRI volumes (Adeli et al., 2018) and DTI metrics (Tang et al., 2017).

In HIV+ children, one study has been conducted to find whether quantitative MR techniques such as relaxometry and spectroscopy were more sensitive to early brain changes caused by HIV than conventional inspection of MRI images (Prado et al., 2011). However, no study to date has combined neuroimaging modalities or used multivariate pattern recognition approaches for classification of paediatric HIV infection. Examining neuroimaging features that are implicated in accurate HIV-status class predictions may provide information about which measures from various neuroimaging modalities are most relevant to paediatric HIV.



3 Literature review: feature selection and multimodal neuroimaging

Contents

3.1. Feature selection 20
3.1.1. Filter techniques 21
3.1.2. Wrapper techniques 21
3.1.3. Embedded techniques 22
3.2. Application of feature selection algorithms 22
3.3. Feature selection in multimodal neuroimaging 25
3.4. The elastic net regularisation 27

Chapter 3 Overview

- MRI images contain from 100 000 to a million voxels depending on voxel size.
- High ratio of *p* to *n* increases the tendency of a prediction model to fit to the noise of the sample, and a poor fit when applied to a new dataset.
- In contrast to dimensionality reduction techniques, feature selection preserves the interpretation of variables.
- The price for high quality of feature selection algorithms is training time.
- Elastic-net and LASSO outperform wrapper and filter techniques.
- There is a correlation between the total number of features selected by a model and stability of the classification model.
- Feature selection techniques can be extended to multimodal models.
- Multimodal feature sets improve classification performance measures such as accuracy, sensitivity, specificity and AUC.
- K-fold cross validation is among the best validation techniques to ensure a classification model performs well on independent datasets.

3.1. Feature selection

Consortia and projects are collecting large amounts of neuroimaging data. MRI-based examples include: Alzheimer's Disease Neuroimaging initiative (ADNI, Jack et al., 2008), Adolescent Brain Cognitive Development study (ABCD), International Consortium for Brain Mapping (ICBM, Mazziotta et al., 2009), IMAGEN consortium (Schumann et al., 2010), Collaborative Initiative on Foetal Alcohol Spectrum Disorder (CIFASD, Moore et al., 2007), and OpenNeuro (Gorgolewski et al., 2017).

A neuroimaging dataset is high-dimensional in nature, since it often contains more predictor variables (voxels, measurements, or features) than observations (e.g. participants). As such, it is routine for neuroimaging analyses to confront the curse of dimensionality ($p >> n$, where *p* and *n* denote the number of variables and observations respectively). For example, a single MRI image can contain between 100 000 and a million voxels, which results in regression or classification models generalising poorly because of capturing irrelevant 'noisy' patterns in the sample (Jollans et al., 2019; Paulus et al., 2019). Neuroimaging-specific overfitting is discussed in depth by Whelan & Garavan (2014). Even if small regions of interest rather than voxels are selected, the resulting number of features remains large for multimodal neuroimaging datasets that are combined with other indications (e.g. cognitive data, demographics, other imaging contrasts or modalities). Ergo, dimension reduction is required in order to avoid overfitting and to construct a classification paradigm with good predictive power.

Dimension reduction can be achieved via various methods: either unsupervised feature projection techniques such as the data-driven principal and independent component analyses (PCA & ICA), domain-knowledge-driven coordinate-based meta-analysis (CBMA), or a supervised-learning approach like feature selection (Guyon & Elisseeff, 2003). However, taking the projection route will not necessarily preserve interpretability of the features because they are linearly or nonlinearly transformed into a space with fewer dimensions than the original. Understanding the neuroimaging features that are characteristic of HIV infection is an important part of our aim, accordingly, feature selection is a preferable course.

To develop an accurate predictive model from a neuroimaging dataset, only the most relevant attributes which accurately discriminate between classes should be considered; feature selection is an objective-function-driven process of identifying those attributes in the dataset material to the prediction problem. Consequently, this procedure provides fast training times; improves performance of a prediction model; and offers insight into the underlying process that generated the neuroimaging data (ibid.). Feature selection and supervised machine learning algorithms are increasingly applied to these high-dimensional neuroimaging datasets (e.g. Davatzikos, 2019; Jollans et al., 2019; Janssen et al., 2018; Woo et al., 2017; Wu et al., 2017).

Broadly, three feature selection algorithms exist: filter, wrapper, and embedded techniques (Khaire & Dhanalksmi, 2019; Mwangi et al., 2014; Chandrashekar & Sahin, 2014). These algorithms and their examples are summarised as follows:

3.1.1. Filter techniques

Filters assign a scoring to each feature using statistical descriptors, such as the mean and variance. The score is used to rank the feature in terms of its relevance to the outcome thereby determining whether it should be kept or removed from the dataset. The most basic example is the variance thresholding method: all features with zero-variance i.e. have the same value in both classes are removed. Other examples of filters include linear discriminant analysis (LDA), and information gain (IG). LDA finds a linear combination of features that minimises the spreading with the class and maximise the distance between the mean of each class. IG measures how much “information” each feature gives about the class; unrelated features should give zero information while features that perfectly partition between classes give maximal information.

3.1.2. Wrapper techniques

Wrappers use an objective function from a regression or classification model to rank the features. The selection is generally a search problem with forward and/or backward passes, random hill climbing algorithms, or best-fit searches for relevant features. In practice, the objective function for a wrapper can be a filter itself. A common method used for wrapping is recursive feature elimination (RFE). In RFE, all features are initially assigned weights (e.g. coefficients of a linear model), then the weakest coefficient is removed and the model refitted

to find the weakest coefficient of the new model. This process is repeated until all irrelevant features are exhausted.

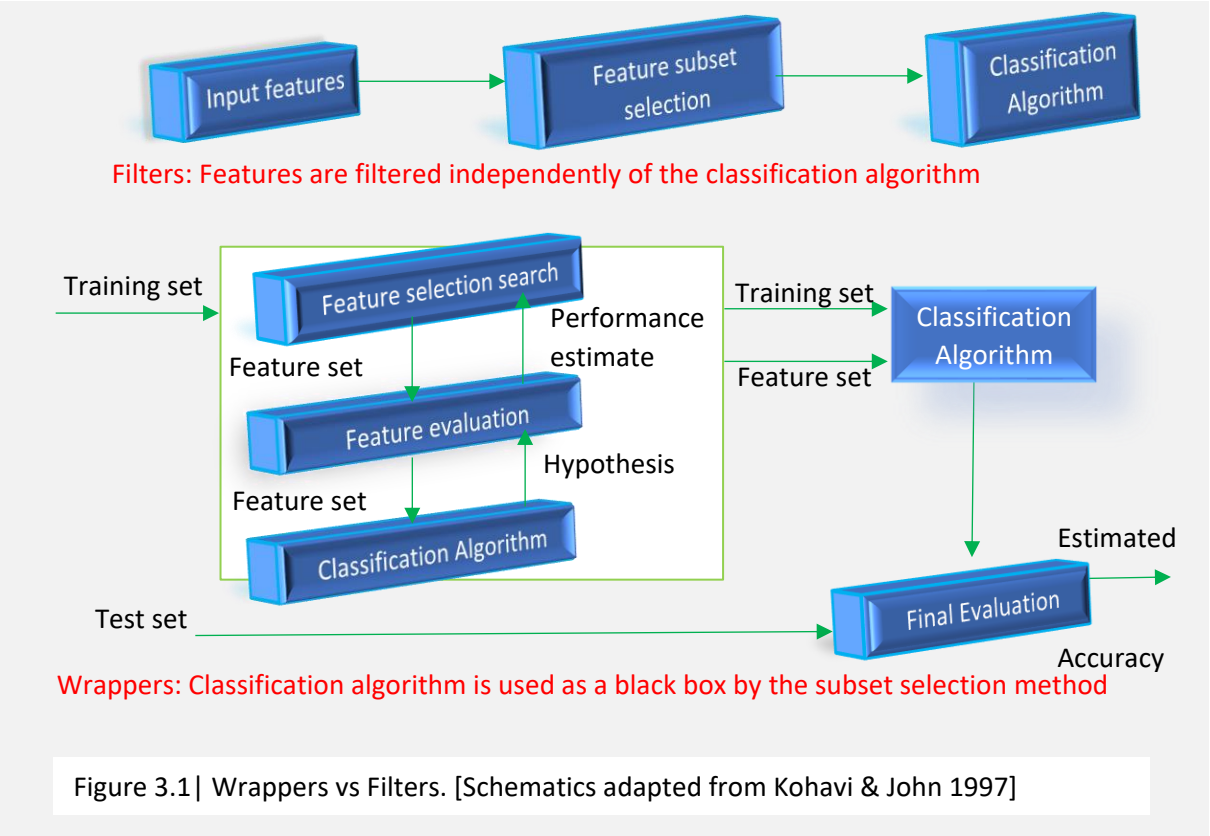


Figure 3.1| Wrappers vs Filters. [Schematics adapted from Kohavi & John 1997]

3.1.3. Embedded techniques

These algorithms select the relevant features by enforcing certain penalties while the model is being created—feature selection is part of the classification algorithm. Examples are regularisation methods: least absolute shrinkage and selection operator (LASSO), ridge regression, and elastic net (EN). LASSO penalises L^1-norm the coefficients of the regression variables shrinking some of them exactly to zero. Ridge method adds L^2-norm penalty making coefficients of irrelevant predictors approach zero. EN adds both vector norms penalties to select relevant features.

***

Despite being fast, filters are usually criticised because the feature set is not optimised for the classifier used (Figure 3.1); therefore, they are often used as a pre-processing step for other methods. Wrappers are universal and simple but require a large amount of computation

(training time). Moreover, their predictive power is measured by validation sets. Embedded methods are in the middle, taking advantage of predictive and computational power of both filters and wrappers. These methods are likely to identify relevant patterns and generalise well to new data.

In the study of dementia using sMRI (Tohka et al. & ADNI, 2016), embedded feature selection methods outperformed SVM-based wrappers and filters for distinguishing mild cognitive impairment (MCI) from normal function, but choice of feature selection algorithm did not matter for distinguishing Alzheimer's disease (AD) patients from controls. Elastic net and LASSO had a classification accuracy of 81—84%, while all SVM t-test filters had an accuracy of 80—82%. Tohka et al. (2016) also demonstrated that the stability of the feature set correlates with number of features selected. Feature sets identified using embedded feature selection were not as stable as those identified with filter-based methods.

3.2. Application of feature selection algorithms

In background, we alluded to two studies that used embedded methods within single modalities to classify HIV in adults (Adeli et al., 2018; Tang et al., 2017). Adeli and colleagues (2018) found that classifiers using all structural measures had an accuracy of 87.7% (87% AUC). When average thickness was removed the accuracy dropped to 79.6% (78% AUC). When mean curvature, surface area, and gray matter volume were removed so that only cortical measurements where left in the prediction model, the accuracy was 69.2% (75% AUC). Tang et al. (2017) could classify HIV with 83.1% accuracy and AUC of 0.91 and established that motor-related tracts like the corticospinal tract were the most impacted by HIV.

Feature selection techniques can be extended to multimodal neuroimaging studies to improve classification performance and stability. Given in Table 3.1 is a summary of previous studies investigating the effectiveness of embedded feature selection in multimodal neuroimaging to find a subset of relevant features i.e. biomarker of a disease. Furthermore, these studies provide a framework for combining modalities and assessing whether classification performance is improved satisfactorily when using embedded classifiers.

Author(s)	Disease condition	MRI modalities and other features	Best feature combination	Classification and feature selection algorithm	Model Evaluation	Classification Performance	Conclusion
Schouten et al. (2016)	Mild to moderate Alzheimer's Disease (AD)	• sMRI • DTI • rs-fMRI	• WMD • GMD • FA • MD	(Embedded classifier) (EN-VA) 10-fold cross validation (CV), process repeated 10 times (100 training samples)	ROC (AUC)	From unimodal AUC range: 0.743—0.915 to multimodal, 0.946—0.959	Multimodal procedure provides the better classification performance than unimodal techniques in mild to moderate AD
Bowman et al. (2014)	Neurodegenerative Parkinson's disease (PD)	• DTI • T1-MRI • rs-fMRI	24 neural manifestations of PD • 2 regional VBM measures • 1 SC measure • 21 FC features -- (Used AAL) --	EN-VA (and univariate screening) 2-fold cross validation, process repeated 100 times (200 training samples)	ROC (AUC)	Features selected with a bounded search region $\mathrm{AUC} \geq 0.9$	Approach demonstrates extremely high accuracy isolating implicated PD regions. rs-fMRI most predictive of measures
Cabral et al. (2016)	Classes of Schizophrenia (SZ)	• sMRI • rs-fMRI	• Age (Patients) • Age (illness)	PCA + L2-LR (L2-regularised logistic regression) & OOT (out-of-group training) 20-fold cross validation, membership score	ROC (Accuracy)	Dual modality reached 75% classification accuracy compared to 69.7% (sMRI) and 70.5 % (fMRI)	Fusing 2 imaging modalities achieves higher accuracy distinguishing HC from SZ patients
Ritter et al. (2015)	Mild cognitive impairment (MCI) to AD	• sMRI • BIO (CSF) • PET • NP Testing • DEMO • MEDHIST • BLSYMP	Predictive Feature ranking: 1. Functional Activities Questionnaire (NP) 2. AD Assessment Score (NP) 3. Volumes of Left and right hippocampi (MRI)	Forward/backward feature selection & SVM (C=1), F-score, and Random Forest. 10-fold cross validation	ROC (Accuracy)	Methodology offers classification accuracy in the range 61.48%—73.44%	Conversion from MCI to AD is predictable with high accuracy, balanced sensitivity and specificity. Nonlinear SVMs produce best classification results.
Sui et al. (2015)	Cognitive deficits in Schizophrenia (SZ), PANSS score	• rs-fMRI • sMRI • DTI	1. fALFF 2. GM Volume changes in regions: • Thalamus • Striatum • Hippocampus • Mid-occipital region	MCCA (multi-set canonical correlation analysis) Canonical Variants (CVs) score	Variance thresholding Only features with 99.5% of the variance were retained.	Performance not studied for unseen data.	FA changes were not distinct in increased or reduced cognitive performance. Linked functional and structural deficits in cortico-striato-thalamic circuits may be closed related to cognitive impairments in SZ
Colby et al. (2012)	Classes of Attention Deficit Hyperactivity Disorder (ADHD)	• sMRI • rs-f MRI • DEMO	• Gender • IQ And cortical features ranked: • Posterior cingulate • Bank of superior temporal sulcus (Thickness Average) • Superior temporal (Mean Curvature)	SVM-RFE (feature ranking) & RBF-SVM (generalisation of performance) 10-fold cross validation	ROC (Accuracy)	Increased classification accuracy to 55% from 39% chance level in the sample	The approach allows predictive structural and functional features into abnormal brain circuitry in ADHD
Tohka et al. (2016)	Dementia (Feature Selection techniques comparison)	Structural MRI (GMD, WMD, CSF)	NA	▪ EN-VA ▪ EN-05 ▪ EN-05STAB ▪ LASSO ▪ LASSOSTAB ▪ SVM-Fx ▪ SVM-ALL k= 10 folds and BEE	ROC (Accuracy)	Highest Test accuracies for FS methods: EN-VACV : 82.1—84.6% EN-VABEE : 81.5—84.1% LASSOCV : 81.3—81.7%	Embedded feature selection methods outperform SVM-based ones (MCI vs HC.) Stability of feature set correlates with number of features selected.
Kambeitz et al. (2017)	Depression	• sMRI • tr-fMRI • rs-fMRI • DTI	rs-fMRI and DTI measures (FA)	EN-VA 2-fold CV, leave-one-out CV & leave-one-subject-per-group-out CV	ROC (Sensitivity and Specificity)	Sensitivity/Specificity sMRI: 69.9%/71.1% tr-fMRI:74.1%/77.2% rs-fMRI: 85.4%/83% DTI: 88.7%/91.5% Combined 76.7%/77.7%	Feature selection technique reduced the performance of the individual modalities except sMRI in classifying subjects with Depression.

Abbr. MEDHIST – medical history, DEMO – demographics, NP – neuropsychological, BLYSMP – blood sample, SVM – support vector machine, rs-/tr – resting state/task related, EN-VA – elastic net with cross validation, EN-05STAB – elastic net, second alpha set at 0.5, with stability selection; PANSS – positive and negative syndrome scale, high correlation (+) and low correction (-); fALFF – fractional amplitude of low frequency fluctuations; BIO(CSF) – biological sample of cerebrospinal fluid; BEE- Bayesian error estimation

Table 3.1| Literature review matrix of feature selection & multimodal neuroimaging

A meta-analysis of multivariate pattern analysis for identifying neuroimaging biomarkers of depression (Kambeitz et al., 2017) found a mean sensitivity of 76.7% and 77.6% specificity respectively across four modalities: DTI, tr-fMRI, rs-fMRI, and sMRI (Kambeitz et al., 2017) considered separately. Resting-state fMRI (85% sensitivity, 83% specificity) and DTI (88% sensitivity, 92% specificity) had better performance than sMRI or task-based fMRI in predicting depression.

Many attempts have been made to distinguish between classes of dementia (mild cognitive impairment to Alzheimer's disease (Schouten et al., 2016)), Parkinson's disease (Bowman et al., 2016), attention deficit hyperactivity disorder (Colby et al., 2012), and depression (Kambeitz et al., 2017) using neuroimaging features to achieve classification accuracies in the 61.9—95.9 % range. From these studies, classification performance increased by 2—7 % when using a multimodal approach compared to single modalities.

Because it results in the largest classification improvement, more studies have used embedded feature selection: including elastic net regularisation (Bowman et al., 2016; Schouten et al., 2016; Tohka et al., 2016), LASSO (Tohka et al., 2016) and ridge regression (Cabral et al., 2016) than support vector machines in combination with forward/backward feature selection and recursive feature elimination (Ritter et al., 2015; Jie et al., 2014; Zhang & Wu, 2012; Colby et al., 2012). In addition, Cabral and colleagues (2016) suggest applying a feature reduction technique such as PCA before feature selection or canonical correlation analysis.

Schouten et al. (2016) combined anatomical, diffusion, and resting state fMRI features in the classification of mild to moderate AD from structural and functional changes in the brain using an elastic net classifier. The most relevant features were found to be white and gray matter densities (WMD, GMD), mean diffusion, and fractional anisotropy (FA). The same modalities were combined using an elastic net in the study of neurodegenerative Parkinson's disease (PD) and 24 features were found that predicted PD: 2 regional volume-based morphometry (VBM), 1 SC measure, and 21 FC features; rs-fMRI features were the most predictive (Bowman et al., 2016). Cabral and colleagues (2016) found that combining features from sMRI and rs-fMRI achieved higher classification accuracy in distinguishing schizophrenia patients from uninfected controls than either modality alone, although the most predictive measures were age of patients and age of onset. Another study of schizophrenia included diffusion features

(Sui et al., 2015), however the most predictive features were fractional amplitude of low frequency fluctuations (fALFF) and gray matter volume changes in the thalamus, striatum, hippocampus, and mid-occipital regions. In several studies, demographic data (age, gender), neuropsychological testing (functional activities questionnaires, cognitive scores, etc), and medical history are included as features in addition to neuroimaging, and have been shown to be better predictors in schizophrenia (ibid.) and MCI conversion to AD (Ritter et al., 2015) than neuroimaging features.

For example, Colby and colleagues (2012) used sMRI, rs-fMRI, and demographic measures. This approach allowed the selection of structural and functional features and their association with abnormal brain circuitry in ADHD. The most predictive features were gender, intelligence quotient (IQ), and a ranking of cortical features, the highest being thickness of the posterior cingulate and bank of superior temporal sulcus as well as the mean curvature of the superior temporal region. Demographic measures could otherwise be used to validate the resulting classifications, offer clinical interpretations, or be treated as uninteresting features i.e. confounders.

To estimate how accurately a classification model performs on an independent data set, the next step is cross-validation (CV). Variations include k-fold CV, leave-one-out CV, and leave-one-subject-per-group-out CV (Kohavi, 1995; Geisser, 1993). Furthermore, in k-fold CV, k=10 is preferable in neuroimaging studies especially for relatively small *n* since it avoids overfitting and selection bias. Contrarily, Bowman et al. (2016) used a 2-fold CV and repeated the process 100 times to produce 200 training samples; the approach demonstrated an extremely high classification accuracy and was thereby able to isolate regions implicated in PD.

There is a general downside to cross validation: when dealing with small sample situations, an excess variability of error estimates is consequential to the classification accuracy (Glick, 1978). The solution is often Bayesian error estimation (BEE) to minimise the posterior expected loss function (Lehmann & Casella, 1998). Although the cut-off for small *n* is unclear, with a sample size $n = 100$, an elastic net classifier with cross validation (EN-CV) on sMRI features (GMD, WMD, and CSF) resulted in a better test accuracy than elastic net with Bayesian error estimation (EN-BEE) (Tohka et al., 2016). However, the accuracy for both techniques was improved on average by 0.023 % when the sample size is doubled ($N = 200$).

Stability selection like BEE is still in its infancy in neuroimaging (Tohka et al., 2016; Ye et al., 2012). It is based on combining the selection algorithms with subsampling to provide sample control for some error rates of false discoveries (Meinshausen & Bühlmann, 2010). Consequently, it is not a new feature selection technique but rather aimed at enhancing and improving existing methods e.g. LASSO, and EN. Lastly, in cases where a sample size should be shrunk because of missing data entries, oftentimes it is best to perform mean imputation and imputation by expectation-maximisation to replace missing datasets (Ritter et al., 2015).

These studies have shown that multimodal disease models formed by integrating different modalities may provide superior classification performance compared to each of these single data channels considered separately. However, further work needs to be undertaken to investigate factors determining the diagnostic performance of MRI-based biomarkers across the life span (Cabral et al., 2016), to investigate dependencies between features (Ritter et al., 2015), to examine interrelationships between cognition measures and abnormalities seen in chosen modalities via a multivariate method (Sui et al., 2015), to determine a biological interpretation of model parameters (Schouten et al., 2016), and to translate neuroimaging research into clinical practice (Mwangi et al., 2014).

Based on the literature, the embedded feature selection of the elastic net was chosen as the preferred approach for this dissertation. The elastic net was selected for its good discrimination capabilities on multimodal neuroimaging data in neurological disorders, as well as its preservation of the interpretability of the selected features.

3.3. The elastic net regularisation

One method of accomplishing a multimodal neuroimaging analysis is to use the different measures from MRI contrasts i.e. DTI, fMRI, sMRI, and ^{1}H-MRS, and cast them into an elastic-net embedded feature selection algorithm (Zhou & Hastie, 2005).

Elastic net regularisation, as well as the closely related LASSO method, is frequently selected for neuroimaging classification problems (Bowman et al., 2016; Schouten et al., 2016; Tohka et al., 2016). By identifying a small number of relevant predictors, these algorithms are well-suited for data with a small number of observations and large number of variables (small *n*

large *p*), and coefficients are more easily interpretable than many other methods.

Zhou and Hastie (2005) formulated EN regularisation to solve the limitations of LASSO and ridge regressions while including each as special case. Friedman et al. (2010) provided algorithms for fitting penalised logistic and multinomial regression models because this is a more natural approach for classification problems. In the case of a binary response variable e.g. HIV+ and HIV- outcomes, the fitting algorithms proposed by the authors maximise the penalised log-likelihood function (Equations 3.1):

$$\max_{\beta_0 \beta} \left[\frac{1}{N} \sum_{i=1}^{N} y_i \cdot (\beta_0 + x_i^T \beta) - \log\left(1 + e^{\beta_0 + x_i^T \beta}\right) - \lambda P_a(\beta) \right]$$

Where the penalty term:

$$P_\alpha(\beta) = \frac{1-\alpha}{2} \left|\left|\beta\right|\right|_2^2 + \alpha \left|\left|\beta\right|\right|_1$$
$$= \sum_{j=1}^{P} \left(\frac{1-\alpha}{2} \beta_j^2 + \alpha \left|\beta_j\right| \right)$$

EN is the same as L^1 penalty when α = 1, just as in LASSO, and L^2 penalty (ridge regression) for α=0. Therefore, the penalty term P incorporates both L^1 and L^2 norm regularisations.

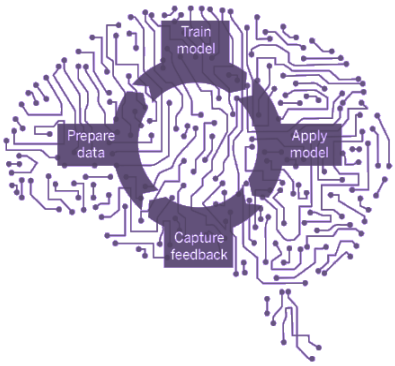
- N is the number of observations
- p is the number of predictors
- y_i is the response at observation i
- x_i is data, a vector of p values at observation i
- α is a mixing parameter between 0 and 1
- λ is a positive regularisation parameter
- β_0 is intercept of the logistic regression model
- β_j is the j^{th} estimated logistic regression coefficient

Eq 3.1

The solution is commonly found using the coordinate descent algorithm which is provided by Friedman and collegues (2010) in both R and MATLAB (The MathWorks Inc.). This regularisation technique identifies relevant predictors, produces shrinkage estimates with potentially less classification errors than ordinary least squares, and selects among redundant predictors. The quadratic part of the penalty counters any limitation on the number of

selected features and leads to a grouping effect for highly correlated variables (Mwangi et al., 2014; Ogutu et al., 2012; Bunea et al., 2011). This is useful if there are many correlated predictors which is often the case with neuroimaging data.

***



4

Methodology: identifying multimodal HIV signatures

Contents

4.1. Study design and participants 31
4.2. Neuroimaging acquisition 33
4.3. Processing and extraction of neuroimaging measures 34
4.4. Quality control criteria for feature sets and handling of missing data 35
4.5. A breakdown of objectives 36
4.6. Cross-validated elastic net models on individual feature sets 37
4.7. Feature selection for elastic-net models 40
4.8. Multimodal elastic-net models 41
4.8.1. Non-nested multimodal models 42
4.8.2. Nested multimodal models 43
4.9. Model comparisons 45
4.9.1. Likelihood-ratio $\chi 2$ tests 45
4.9.2. Akaike Information Criterion 46
4.10. Multivariate Imputation using Chained Equations (MICE) 47

Chapter 4 Overview

- The number of observations (sample size) is $n = 127$ (72 HIV+, 55 HIV-)
- 7-MRI derived feature sets were used in the multimodal analysis
- Feature (variable) selection and classification were produced from logistic elastic-net regression models
- EN models were formally compared with AUC and AIC to assess goodness of fit

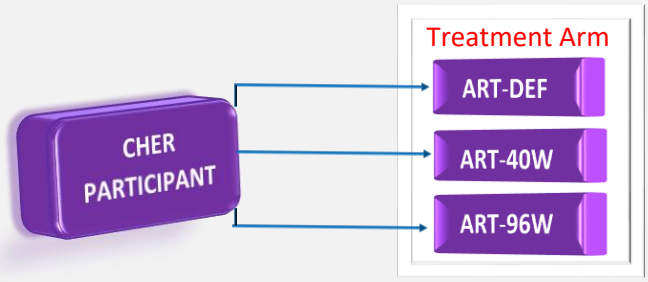
4.1. Study design and participants

The study population includes asymptomatic HIV+ minors from the Children with HIV Early Antiretroviral Therapy (CHER) trial conducted in Cape Town and Johannesburg, South Africa. This open label randomised controlled trial was carried out to determine when to start cART treatment for perinatally HIV-infected children.

Infants on the CHER trial with $CD4^{+}$ percentage of at least 25% were randomised to one of the three treatment arms shown in Figure 4.1. For one group treatment was deferred until clinically indicated, the other two groups began cART immediately, but had treatment interrupted after respectively 40 and 96 weeks. After interruption, cART was restarted upon observation of clinical evidence of disease progression. The cART regimen consisted of $Zidovudine\ (ZDV) + Lamivudine\ (3TC) + Lopinavir{\sim}Ritonavir\ (LPV/r)$. It was found that early treatment reduced mortality, provided better clinical outcomes, and correlated with better neuropsychological test performance at 18 months than late treatment (Cotton et al., 2013; Laughton et al., 2012; Violari et al., 2008).

In addition to 72 HIV+ children from the CHER trial, our study also included 55 HIV negative (HIV-) control children from a related vaccine trial (Madhi et al., 2010). The HIV- group comprised 34 unexposed (HU) and 21 exposed yet uninfected (HEU) children born to HIV+ mothers. All children were enrolled in a follow-up longitudinal neuroimaging study (Nwosu et al., 2018; Robertson et al., 2018; Jankiewicz et al., 2017; Holmes et al., 2017; Mbugua et al., 2016) and had MR neuroimaging every 2 years from the of age 5. Neuroimaging data available is at ages 5, 7, 9, and 11 years. This dissertation analyses cross-sectionally data at age 7 $(mean \pm sd = 7.22 \pm 0.13\ years)$ because this was the biggest available dataset with the highest number of subjects scanned at a single age.

Sample characteristics of the study population are given in Table 4.1.



$ART - DEF$: Antiretroviral therapy (ART) deferred until $CD4\% < 25\%$ in year 1, or $CD4\% < 20\%$ afterward

$ART - 40W$: ART started before 12 weeks of age and interrupted after 40 weeks.

$ART - 96W$: ART started before 12 weeks of age and interrupted after 96 weeks.

Figure 4.1| Treatment arms for HIV+ asymptomatic infants part of the CHER trial

Table 4.1| Sample characteristics of our study population

	HIV+	**Controls**
Demographics		
Observations (n)	72	55
Gender Female, n (%)	36 (50%)	24 (44%)
Age at scan	7.20 ± 0.13 years	7.24 ± 0.13 years
Clinical measures at baseline (6—8 weeks)		
CD4 count (cells/mm^3)	1 764 ± 869	
CD4%	33 ± 11	
CD4/CD8	1.28 ± 0.73	
CD8 count (cells/mm^3)	1 666 ± 1 097	
CD8%	31 ± 10	
High viral load (>750 000)	36	
Low viral load (400—750 000)	25	
Suppressed viral load (<400)	0	
Clinical measures at scan		
CD4 count (cells/mm^3)	1 128 ± 464	
CD4%	37 ± 6	
High viral load (>750 000)	0	
Low viral load (400—750 000)	3	
Suppressed viral load (<400)	58	
Treatment related measures		
cART initiation before 12 weeks	48	
Age at cART interruption (weeks)	71 ± 27	
Duration of cART interruption (weeks)	65 ± 89	

4.2. Neuroimaging acquisition

Children were scanned at age 7 years on a 3 Tesla Allegra MRI scanner (Siemens Erlangen, Germany) at the Cape Universities Brain Imaging Centre (CUBIC)—South Africa using a single channel head coil. Neuroimaging was performed without sedation in accordance with protocols approved by Human Research Ethics committees of the participating institutions (UCT, Stellenbosch University); parents or guardians provided written informed consent, and oral assent was provided by the children, who were first familiarised with the procedure on a mock scanner. Structural and diffusion tensor imaging were performed along with single-voxel ^{1}H-MRS. The high-resolution structural T_1-weighted acquisition used a 3-dimensional EPI navigated MEMPRAGE sequence with field of view (FOV) of $224 \times 224 \times 144\ mm^3$, voxel size of $1.3 \times 1.0 \times 1.0\ mm^3$, flip angle of 7°, echo times (TEs) of $1.53/3.19/4.86/6.53\ ms$, repetition time (TR) of $2530\ ms$, inversion time (TI) of $1\ 160\ ms$, bandwidth of $650\ Hz/px$, and 144 sagittal slices—scan time (SCT) of 5min and 20s.

Two diffusion-weighted (DW) datasets were acquired with a volumetric navigated (Alhamud et al., 2012) TRSE: $FOV = 224 \times 224 \times 144\ mm^3,\ voxel\ size = 2 \times 2 \times 2\ cm^3,\ TR = 2000\ ms, TE = 86\ ms, b_0 = 0\ s.mm^{-2},\ b_1 = 1000\ s.mm^{-2}, 30\ diffusion\ directions$ with opposite phase encodings for EPI distortion correction during processing—SCT of 13min and 20s . ^{1}H-MRS data was obtained from 3 voxels of interest (Figure 4.2)—in the basal ganglia (bg), midfrontal gray matter (mfgm) and peritrigonal white matter (pwm) using EPI-navigated PRESS: $TR = 2000\ ms, TE = 30ms, voxel\ size = 1.5 \times 1.5 \times 1.5\ cm^3$, $averages = 64$ with CHESS for water suppression—SCT of 2min and 16s per MRS voxel; an acquisition without water suppression in each voxel was performed for water referencing (SCT 12s).

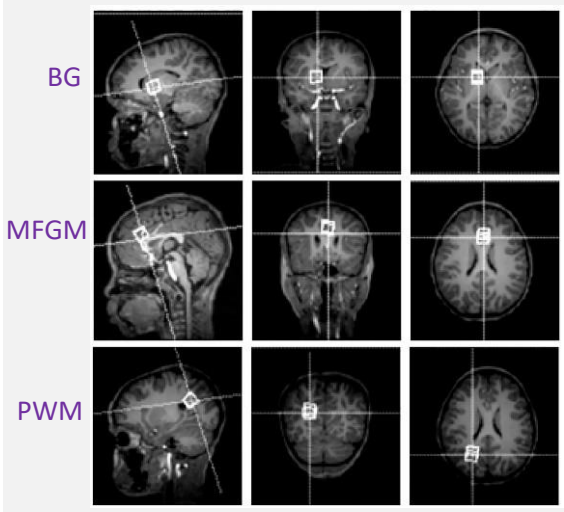


Figure 4.2| Voxel placement shown in sagittal, coronal, and axial planes—bg voxel is frontal limb of internal capsule and part of globus pallidus, putamen, and caudate nucleus; mfgm voxel is mainly gray matter, and pwm voxel is primarily white matter.

4.3. Processing and extraction of neuroimaging measures

For spectroscopy, processing was performed as described in (Robertson et al., 2018). Eddy current correction and calculation of absolute concentrations of metabolites and their ratios to creatine were performed with LCModel version 6.1 (Provencher, 2001). Segmentation of MEMPRAGE regions corresponding to single voxel spectroscopy (SVS) voxel into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) was performed with the SPM12 software, and for partial volume correction and absolute metabolite concentration estimation using the water scaling method (Kreis et al., 1993). Concentrations of 14 metabolites and their ratios to creatine including N-acetylaspartate (NAA), creatine (Cr), choline (Cho), myo-inositol (mI), lipids and macromolecules (MM) were obtained for each of the 3-SVS voxels.

For structural T1-weighted dataset, FreeSurfer version 6.0 was used for automated cortical reconstruction, skull-strip corrections (removal of extra-cerebral voxels) were performed if required. The automatic segmentation (Aseg) functionality of the software allowed for a measurement of total and regional brain volumes from 42 structures, and cortical thickness and local gyrification index (LGI) across cortical surface of the brain. Mean cortical thickness and LGI was extracted for each of the 34 regions of the Desikan-Kallany (DKT) atlas (Desikan et al., 2006) in each cerebral hemisphere (LH & RH).

Diffusion-weighted dataset motion correction were done within Tortoise v.2.5.2, and diffusion tensors and diffusivity measures computed in AFNI (Jankiewicz et al., 2017). Mean radial diffusivity (RD), axial diffusivity (AD), mean diffusivity (MD) and fractional anisotropy (FA) were calculated for each of the 20 regions of interest (ROI) in the John Hopkins University (JHU) atlas (Mori et al., 2005).

Features were extracted from the modalities to generate 7 MRI-derived feature sets: sMRI volumes (1st set), sMRI cortical thickness (2nd set), sMRI gyrification (3rd set), and DTI (4th set) ^{1}H-MRS bg (5th set), ^{1}H-MRS mfgm (6th set), ^{1}H-MRS pwm (7th set). A summary of the number of features and regions in each feature set is shown in Figure 4.3.

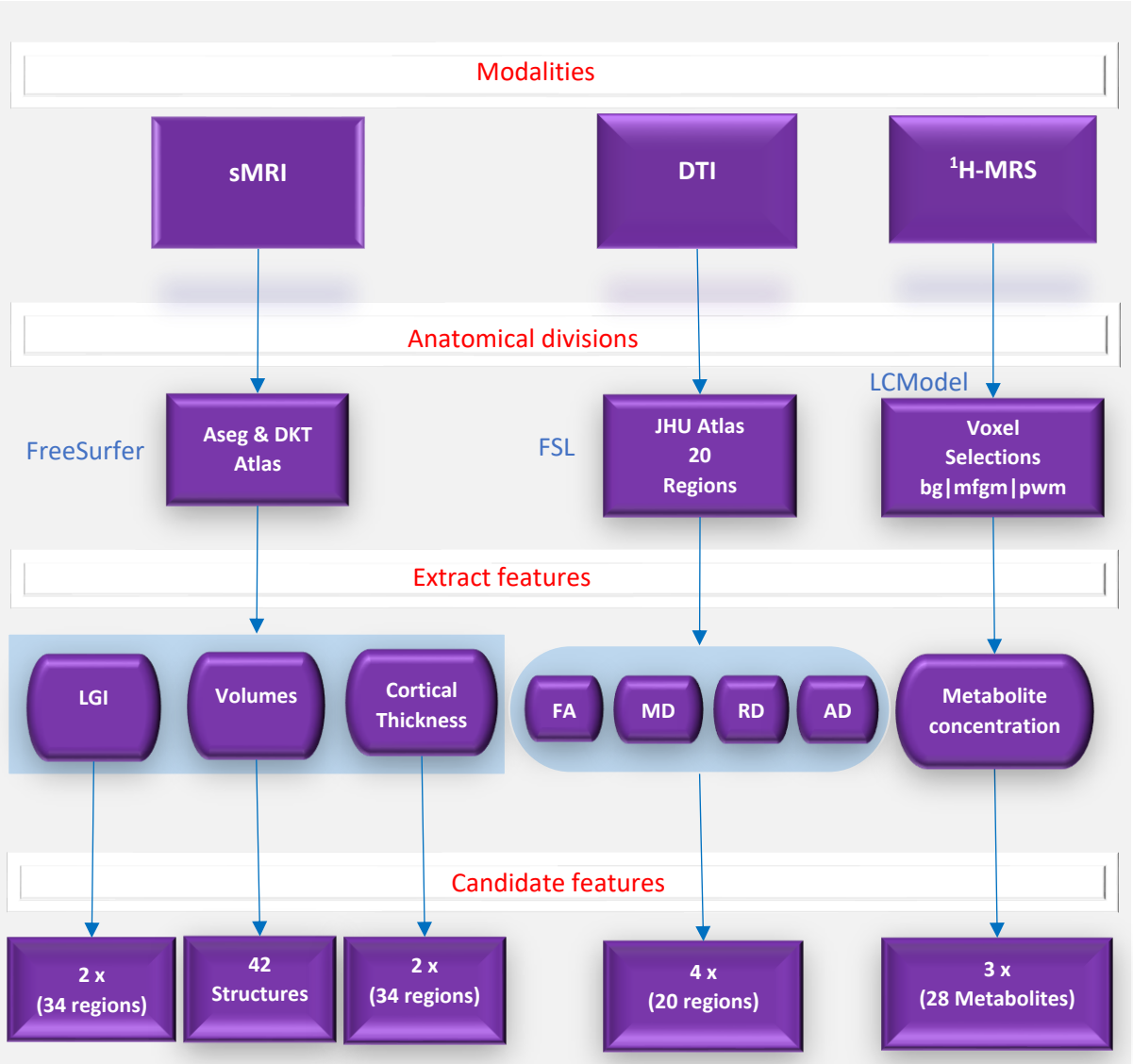


Figure 4.3| Overview of the generation of feature sets—First level shows the considered modalities; the second level—anatomical divisions; and the last level shows the total number of candidate neuroimaging features. Atlas regions and automatically segmented structures are listed in the appendix {List of features}.

4.4. Quality control criteria for feature sets and handling of missing data

Quality-control measures were applied to the 7 MRI-derived feature sets. Extreme outlier T1-weighted measures and DWI with large motion artefacts and dropout slices were eliminated from analysis. Datasets with failed alignment to JHU and DSK templates in standard space were excluded. Similarly, poor quality spectra were eliminated based on low SNR (SNR<6) or broad linewidth (FWHM>0.07 ppm) as calculated by LCModel.

Because some subjects did not complete the full scanning session and in some subjects not all modalities met quality control criteria, all 7 feature sets were not present for all subjects. Subjects were therefore excluded from analyses for which they did not have complete datasets. As an alternative to excluding subjects with missing data, another method for dealing with missing data was explored: multiple imputation of missing data, which is detailed in §4.10. This was completed to investigate whether change in *n*, or specifically a larger sample size, influences variable selection and classification performance.

4.5. A breakdown of objectives

First, the study participants were assigned labels or classes dependent on their HIV status: HIV+ group '1' and controls '0'. Then we sought patterns specific to each class. The actions (italicised) that were taken are formulated below.

Objective 1: Construct separate logistic elastic net prediction models for the following 7 feature sets and determine classification performance for each model using cross validation:

1. neurometabolites in basal ganglia (bg)
2. neurometabolites in midfrontal gray matter (mfmg)
3. neurometabolites in peritrigonal white matter (pwm)
4. FA, AD, MD, and RD of DTI
5. regional gyrification (LGI)
6. subcortical structure volumes
7. regional cortical thickness

Objective 2: Extract neuroimaging measurements within each modality that are most predictive of HIV status: The classifiers are linear combinations of individual measurements (neuroimaging features) in each modality, whose weighting in the linear model can be assessed.

Objective 3: Combine modality measures and determine classification performance for multimodal classifiers: To build multimodal classifiers beginning with one modality and adding the remaining 6 modalities in a stepwise manner, there are 5040 possible arrangements. Therefore, a method is required for deciding which feature sets should be combined in which order.

Objective 4: Formally compare models to measure their goodness of fit: For this action, the hypothesis is directly tested—as a formal complement to already computed model classification performances. The *null* hypothesis is that multimodal predictive features offer no improvement to discrimination capability of the single modality relevant features.

Objective 5: Extract and rank multimodal neuroimaging features of the best combination: An identical process as when extracting measurements from one modality is followed during this step. In this instance, however, there is a linear combination of multimodal features (measurements from combined feature sets) with different weights that determine whether the feature is relevant to be included in a multimodal classifier.

Objective 6: Investigate the feasibility of using imputed data for classification and feature selection: Subjects with predictors that do not meet quality control criteria are not excluded, rather the values are replaced with those estimated by multivariate imputation with chained equations (MICE).

4.6. Cross-validated elastic net models on individual feature sets

Logistic EN regression models were implemented in R, initially on the 7 MRI-derived feature sets separately using 10-fold cross validation (CV). For all elastic net models, variables were z-score normalised (Grus, 2016, ch.10) before the analyses.

The subjects were randomised into folds to initiate a cross-validation loop. In the first iteration: Fold 1 was marked as the validation set (VS) and excluded from the model-building process but used to assess the performance of the resulting model. The remaining folds (Fold 2—Fold 9) formed the training set (TS) which was used to train the EN model, and tune parameters, α and λ, to form the best prediction model. This model was then used to predict the labelled HIV statuses of subjects in the VS, in this case, Fold 1. In the second iteration: Fold 2 was designated as the VS, and the remaining folds became a TS. This step was looped 10 times so that each fold was used once as a VS (Figure 4.4). In each iteration, a nested 10-fold CV loop within the training set was used to tune the values of α and λ.

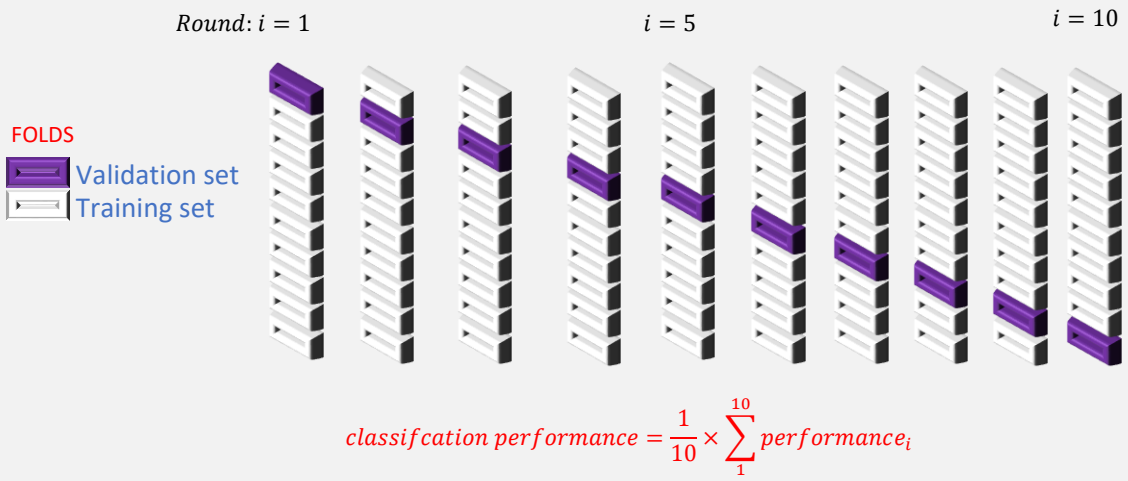


Figure 4.4| 10-fold cross validation scheme—Classification performance metrics are averaged across the folds

For each VS, several classification performance measures were computed: The computation metrics which are defined in equations 4.1 were of *sensitivity*, true positive rate; *specificity*, true negative rate; *accuracy*, the percentage of correct predictions; and lastly, the area under (AUC) the receiver operating characteristic curve (*sensitivity vs 1-specificity*) to show class separability for all possible class decision thresholds, thereby providing an overall performance measure.

$$sensitivity\ (TPR) = \frac{TP}{TP + FN} \qquad specificity\ (TNR) = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TN + TN + FP + FN}$$

- TPR – true positive rate
- TP – true positive
- TN – true negative
- TNR- true negative rate
- FP – false positive
- FN – false negative

Eq. 4.1

Sensitivity and specificity were evaluated at the class decision threshold of 0.5—the optimal cut-off point on the ROC curve closest to true positive rate of 1 and false positive rate of 0. Classification performance measures were averaged across the folds. To obtain a stable estimate of classification performance, the overall CV was repeated 100 times; in each repeat, the data was shuffled and partitioned into different folds. This formed 1000 iterations of training and validation. This way a robust model that generalises well is constructed (Figure 4.5).

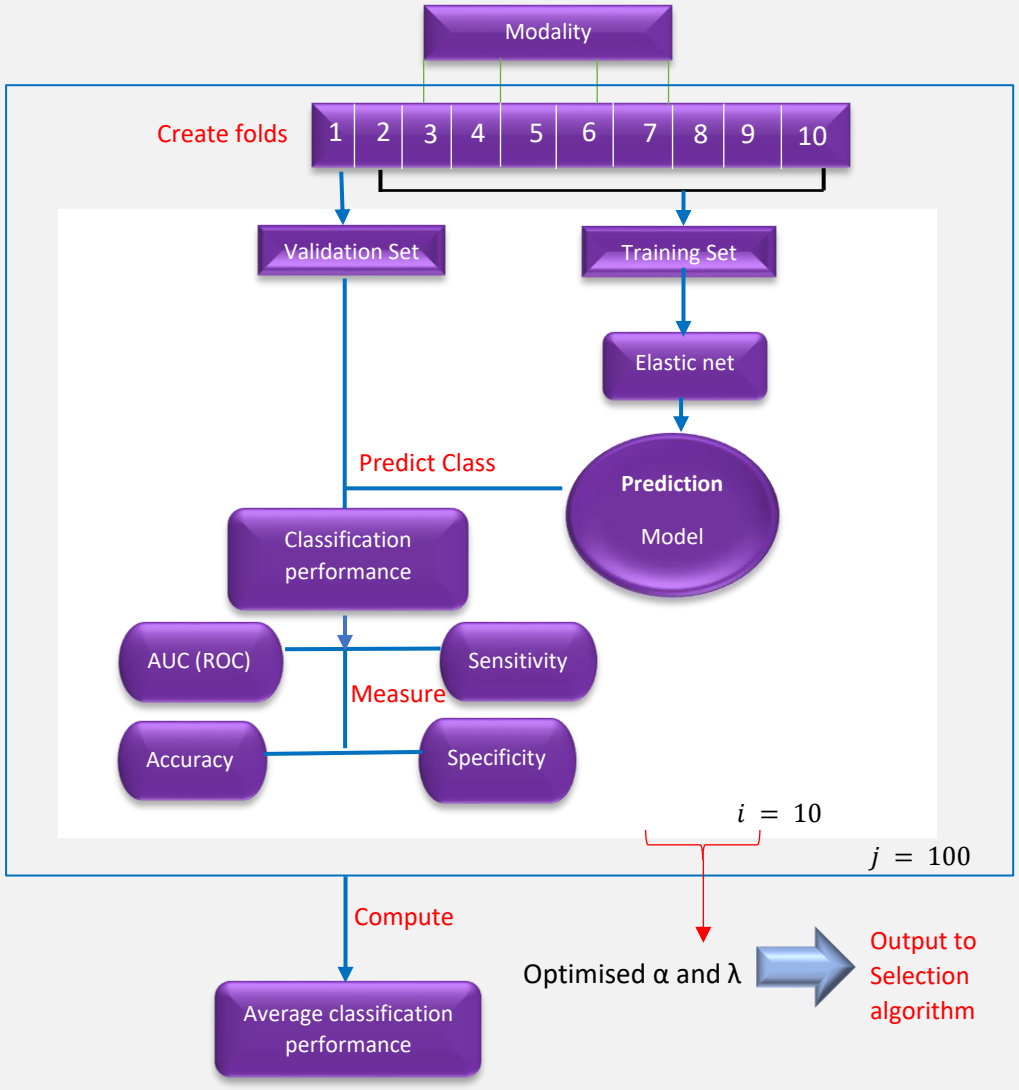


Figure 4.5| A logistic EN regression with 100 iterations of 10-fold cross validation (CV)—Output from CV is optimised penalisation parameters, λ and α. Implemented code in the Appendix {Programming code}.

4.7. Feature selection for elastic-net models

In order to select features that are most useful or relevant in model-building process, to obtain each final model, the EN was retrained on the entire non-partitioned feature space ($VS + TS$) with optimised α and λ parameters from the 10-fold embedded cross-validation procedure which was repeated 100 times as described in the previous section (§4.6).

Using coefficient penalisation or regularisation of EN, the final set of non-zero features was selected (Figure 4.6). It was assumed that selected features from retrained model generated the same classification performance as the embedded cross-validation loops for each repeat cycle. In all models sex, age at scan, and total intracranial volume (*TIV*) were included as confounders with no shrinkage penalty.

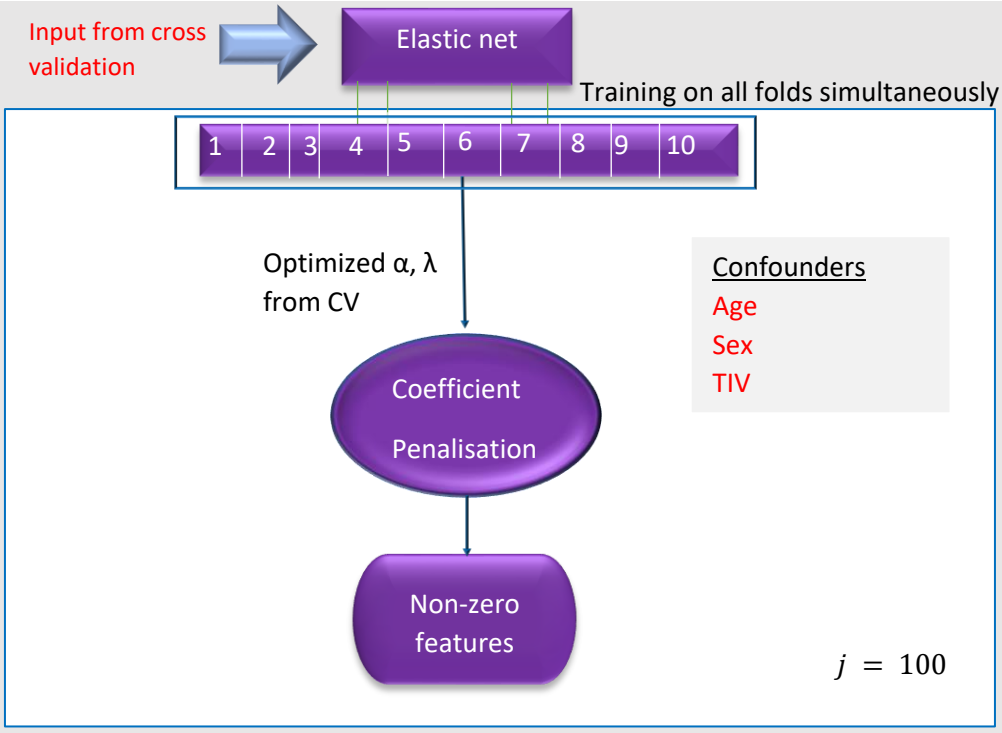


Figure 4.6| Feature selection with CV optimised parameters

The relative importance of a selected feature was evaluated in three ways: the frequency of selection with a non-zero weighting across the 100 repeats, the average value of its weighting in the final model over all iterations including zero-weightings (weighted frequency) and the average of its non-zero weightings (Figure 4.7). Features were determined to be relevant if their frequency of occurrence was greater than 75%. Relevant features were ranked by their weighted frequency.

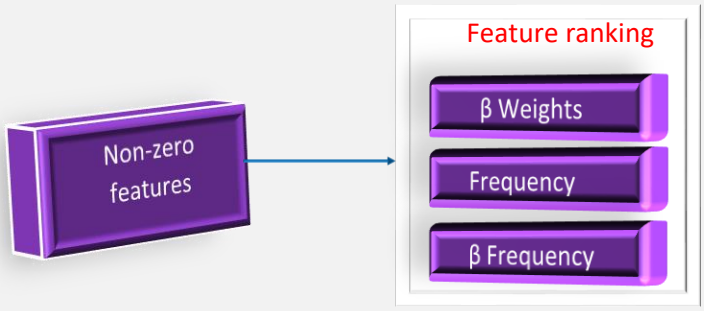


Figure 4.7| Evaluation of selected features during 100 repeat cycle.
Frequency: number of occurrences of each feature per 100-repeat fold shuffling cycle
Average Weight: Average weight in the regularisation model when feature occurs
Weighted frequency: Average weight in the regularisation model including the times the feature does not occur

4.8. Multimodal elastic-net models

Thus far, the implementation of feature selection and classification using EN regularisation has been described on feature sets from a single modality where the predictors are one type of measurements e.g. volume of structures or sMRI volume feature set. An identical process of 100 repeated 10-fold cross validation was followed for multimodal feature sets, the only difference is that candidate predictors (feature space) were then measurements of different types i.e. a combination of predictors from more than one modality, for example one multimodal feature space was formed from WM tract diffusivities and metabolite concentrations in basal ganglia (DTI + ^{1}H-MRS bg).

We explored two techniques for concatenating candidate features to engineer the multimodal feature space before applying EN regularisation for classification and feature selection: they were *nested* and *non-nested* multimodal models.

4.8.1. Non-nested multimodal models

For non-nested EN models, the multimodal feature space is comprised of all candidate features of the combined individual feature sets. This contrasts with nested multimodal feature space where only selected features (a subset) from one modality are kept for combination with a subsequent feature set. The non-nested modelling scheme ensures the best possible combination of variables is selected, thereby preventing overfitting and a biased prediction model.

The non-nested multimodal model-building analyses can be summarised with the following stepwise approach (Schouten et al., 2016) which was undertaken in the analysis:

Step 1	Assess the classification performance of each individual modality §4.6—§4.7)
Step 2	Use the best performing single modality from step 1 as baseline—AUC_1
Step 3	Combine all features from the best performing modality from Step 1 with all the features from each of the remaining feature sets and record AUC e.g. if Step 2 results in DTI as best performing modality, then AUC is computed for DTI + sMRI volumes, DTI + sMRI gyrification, DTI + ^{1}H MRS bg, and so on.
Step 4	Use the combination with the highest AUC from Step 3 as new baseline—AUC_2
Step 5	Combine the best modality features with the remaining sets
Step 6	Repeat until there is no further increase in AUC e.g. $AUC_n < AUC_{n-1}$

4.8.2. Nested multimodal models

In nested models, instead of applying the EN to all the candidate features from multiple modalities, only the best features from the previous step are kept for combination with a subsequent feature set. The nested model building procedure can be demonstrated with an example. EN is cast on the first feature set e.g. sMRI volumes, to produce a subset of relevant features (A), for example:

$$A = \{volume\ 1, volume\ 2, volume\ 3\}$$

Eq. 4.2

Classification performance of this subset is measured. Then, A is concatenated with all candidate features of the next set e.g. DTI measures, and a further EN performed. The result is another subset of relevant features (B):

$$B = \{volume\ 1, volume\ 2, volume\ 3, DTI\ 1, DTI\ 2\}$$

Eq. 4.3

The subset B is engineered in a manner that A is included in the EN model with no shrinkage penalty, so its members are forced to be part of the predictive model. Nested models have an advantage in that the minimum classification performance (AUC, sensitivity, accuracy, and specificity) is that of the initial set (feature set with which one begins the nesting modelling) and will either remain the same or increase with additional features. Building of nested models is also illustrated in Figure 4.8.

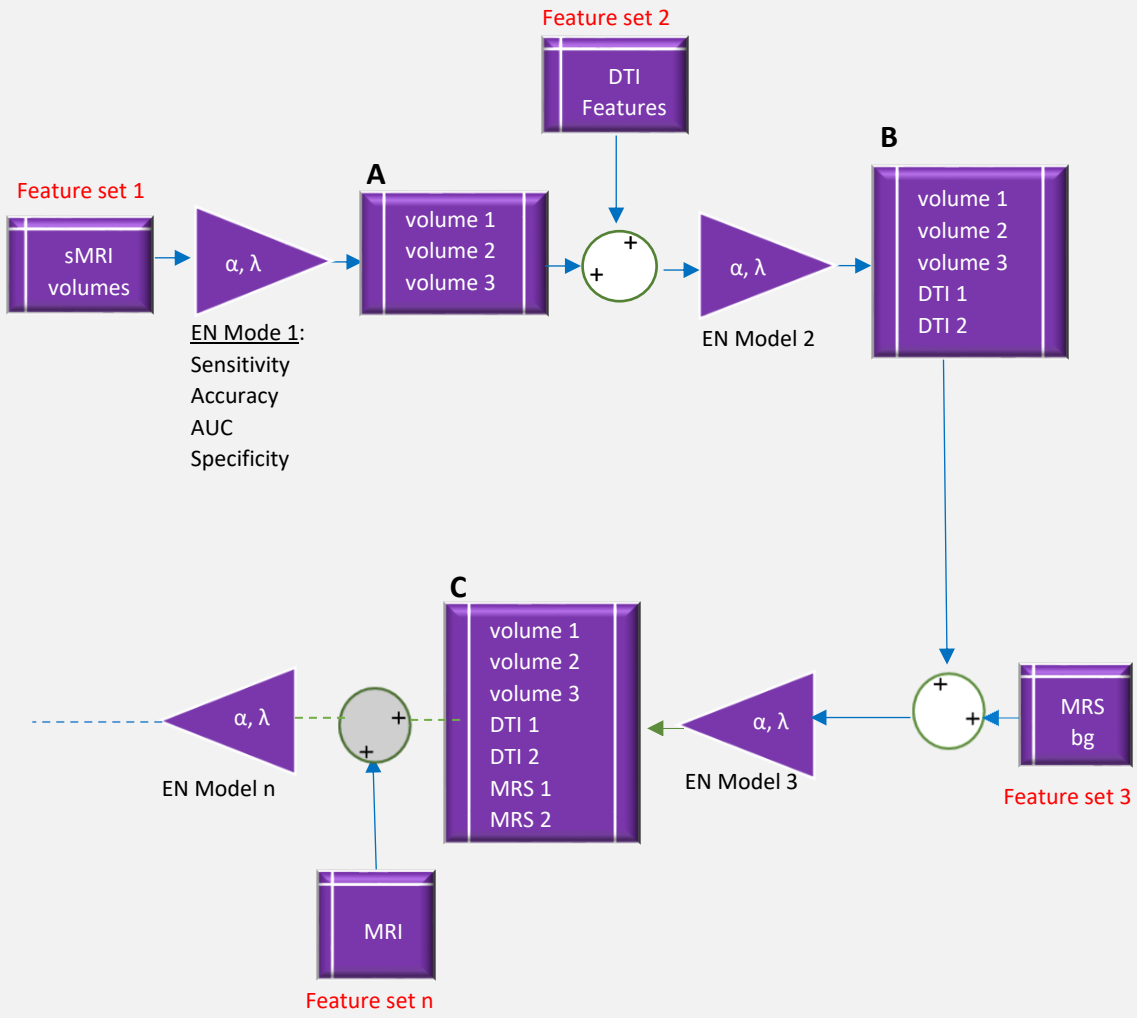


Figure 4.8| Creation of nested models using sMRI volumes as first set—subset of features from sMRI volumes are equivalent to A. The process is a representation of equations 4.2—4.3.

4.9. Model comparisons

Once different EN models are built, in additions to measuring their predictive ability using sensitivity, specificity and the AUC, they may be formally compared in terms of their goodness of fit i.e. to determine whether a model using a larger (multimodal) feature set produces a better fit to the data than one using a smaller subset of features – i.e. a simpler model. Non-nested models are assessed using Akaike Information Criterion (AIC) and likelihood-ratio tests (LRT) are used for their nested counterparts. Both criteria were implemented in R programming language (R Core Team, 2018).

4.9.1. Likelihood-ratio χ^2 tests

LRT is a hypothesis test that helps in choosing the optimal model, by seeing whether the more complex model fits the data significantly better than the simpler nested model with fewer parameters. This is achieved by comparing their likelihoods ($\mathcal{L}$) — the probability of the data given the parameter estimates, which is maximised near the true values of the parameters θ. A larger LRT indicates that the larger model is an improvement to the smaller model. The test statistic approximately follows a χ^2 distribution. The likelihood ratio is defined as:

$$LRT = -2\,ln\left(\frac{\mathcal{L}_A(\theta)}{\mathcal{L}_B(\theta)}\right) \equiv Deviance_A - Deviance_B$$

- $\mathcal{L}_A$ is the likelihood of model A
- A is the simple model e.g. sMRI volume features
- B is the larger model e.g. sMRI volume + DTI
- θ is the set of model parameters i.e. individual variables of A and B
- And $A \subseteq B$

Eq. 4.4

4.9.2. Akaike Information Criterion

AICs are estimated information losses by models. An AIC value is meaningless in isolation, but it allows models to be compared to other models in terms of their simplicity and goodness of fit (Akaike, 1973). The model with the smallest AIC is the optimum model. AICs for every non-nested model were calculated. The first order AIC is mathematically defined below.

$$AIC = -2\,ln\big(\mathcal{L}_A(\theta)\big) + 2k$$

- θ is the set of model parameters i.e. individual variables of A
- $\mathcal{L}_A(\theta)$ likelihood of A given the data is evaluated at maximum estimate of θ
- k is the number of estimated variables in A

Eq. 4.5

In some cases, the sample size is small to an extent that it biases AIC values to favour models with the largest number of parameters. For such cases, usually when the ratio of *n* to *k* is below 40, the second order AIC (AIC_c) is calculated instead. AIC_c considers *n* by increasing the relative penalty for model complexity. Anderson and Burnham (2002, ch.7) defined it as:

$$AIC_c = -2\,ln\big(\mathcal{L}_A(\theta)\big) + 2nk(n-k-1)$$

- θ is the set of model parameters i.e. individual variables of model A
- $\mathcal{L}_A(\theta)$ likelihood of A given the data is evaluated at maximum estimate of θ
- k is the number of estimated variables/parameters in A
- n is the sample size or number of observations in model A

Eq. 4.6

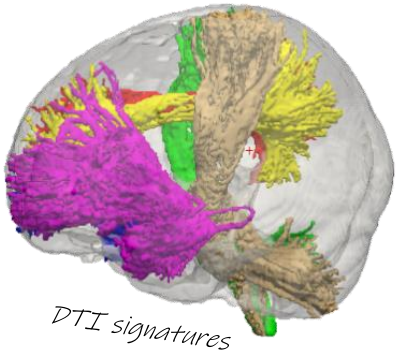
The relative likelihood (the quantity: $e^{-\frac{AIC_{min}-AIC_i}{2}}$) is the probability that models other than the best model will minimise information loss i.e. how likely information will be lost if the i^{th} model is selected instead of the model which best fits the data (AIC_{min}). AIC is closely related to the log-likelihood used in the LRT. Furthermore, it is asymptotically connected to cross-validation—it can be shown that minimising AIC is equivalent to minimising CV error (Peng et al., 2013; Fang, 2011; Stone, 1977). But, AIC works under weaker conditions such as relatively small n, while CV may lead to classification inaccuracy from excess variability of error estimates (as mentioned in §3.4).

4.10. Imputing missing data with Multivariate Imputation using Chained Equations (MICE)

After data quality control, sample sizes of the 7 MRI sets were different. Using a strategy of case-wise deletion of missing data when combining modalities, the overall sample size is that of the smallest dataset. In order to check whether n influenced the classification performance and the behaviour of selected features, the data was imputed i.e. missing or bad values were imputed using Multivariate Imputation using Chained Equations (MICE) as formulated by Azur and colleagues (2011). The procedure can be summarised with the following steps:

Step 1	Perform a simple mean imputation for each missing value as a place holder
Step 2	Set one of the imputed variables back to missing
Step 3	Missing value in Step 2 becomes a dependent variable in a regression model while the other variables are independent.
Step 4	Missing value in Step 2 is then replaced by the prediction from the regression model
Step 5	Steps 2 to 4 are then repeated for every single missing value in the dataset.
Step 6	The number of imputations (repeats of Steps 2 through 4) can be selected, to create a distribution of imputed values. The number was set to 5.

The MICE package in R was used to impute missing values of abovementioned feature sets. The method was set to “classification and regression trees (cart)” with the number of iterations at 5, which produced 5 completely imputed datasets.



5 Findings: classification performance and feature selection

Contents

5.1. Number of observations used for analysis 49
5.2. Single modality elastic net classifiers 50
5.2.1. Classification performance (objective 1) 50
5.2.2. Feature predictive of HIV status in single modalities (objective 2) 50
5.3 Non-nested multimodal elastic net classifiers 55
5.3.1. Classification performance (objective 3) 55
5.3.2. Comparing goodness of fit of single vs multimodal models via AIC (objective 4) 56
5.3.3. Features predictive of HIV status (objective 5) 57
5.4 Nested multimodal elastic net classifiers 60
5.4.1. Classification performance (objective 3) 60
5.4.2. Features predictive of HIV statis (objective 4) 61
5.4.3. Comparison of non-nested vs nested models 62
5.4.4. Features predictive of HIV status (objective 5) 63
5.5 Effect of increasing sample size via imputation of missing data (objective 6) 64
5.5.1 Classification performance 64
5.5.2 Feature selection 64

Chapter 5 Overview

- sMRI volume feature set is the best performing single modality, followed by DTI measures.
- sMRI volumes + DTI + ^{1}H-MRS bg is the best performing multimodal feature set although it is not as likely to minimise information loss as sMRI volumes + DTI

5.1. Number of observations used for analysis

Total number of observations *n* in each modality whose data survived quality control are listed in Table 5.1. These were the sample sizes used for single modality analyses. The number of common subjects in possible multimodal combinations and available for multimodal analysis is listed in Table 5.2.

Table 5.1| A list of participants with useable data in each modality feature set

Measure	Participants (Control/HIV)
sMRI volumes	125 (55/70)
sMRI thickness (RH & LH)	125 (55/70)
sMRI gyrification (RH & LH)	125 (55/70)
^{1}H-MRS mfgm	105 (45/60)
DTI all regions	104 (45/59)
^{1}H-MRS pwm	99 (41/58)
^{1}H-MRS bg	79 (32/47)

Table 5.2| Common participants across modalities

Combination	Common Participants (Control/HIV)
sMRI volumes	125 (55/70)
sMRI volumes + sMRI thickness + sMRI gyrification	125 (55/70)
sMRI all measures + ^{1}H-MRS mfgm	104 (45/59)
sMRI all measures + DTI all regions	102 (45/57)
sMRI all measures + ^{1}H-MRS mfgm + DTI all regions	90 (41/49)
sMRI all measures + ^{1}H-MRS mfgm + DTI all regions + ^{1}H-MRS pwm	79 (35/44)
sMRI all measures + DTI all regions + ^{1}H-MRS bg	69 (29/40)

Structural imaging feature sets had the largest number of observations, and spectroscopy in the basal ganglia had the fewest. As can be seen, multimodal feature sets were necessarily smaller (lower number of subjects *n*) than the largest individual feature sets from structural imaging.

5.2. Single modality elastic net classifiers

5.2.1 Classification performance (objective 1)

Results of training and cross-validated performance for the 7-MRI derived feature sets are shown in Table 5.3. We present the AUC (degree of separability) along with sensitivity, specificity, and accuracy for each individual classifier.

Table 5.3| Classification performance evaluation metrics of the 7-MRI derived feature sets

Modality	AUC	Sensitivity	Specificity	Accuracy
sMRI volumes	0.71	81.18%	66.89%	74.04%
sMRI thickness	0.55	79.65%	51.97%	65.81%
sMRI gyrification	0.54	75.09%	54.00%	64.54%
^{1}H-MRS mfgm	0.58	77.20%	49.65%	63.43%
^{1}H-MRS pwm	0.49	72.27%	31.61%	51.94%
^{1}H-MRS bg	0.55	53.57%	68.52%	61.05%
DTI all regions	0.62	75.31%	69.09%	72.20%

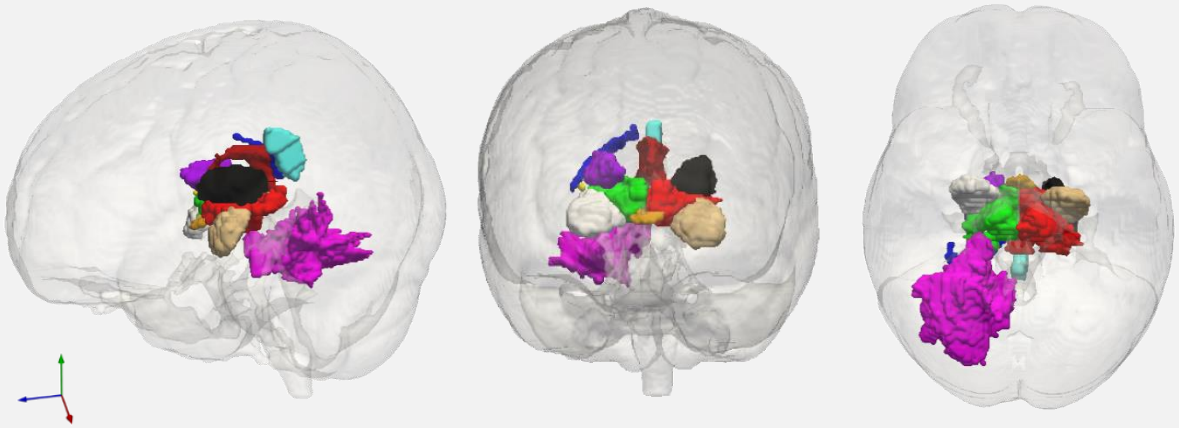
sMRI volumes have better predictive power than other measures as evidenced by higher AUC $(AUC > 0.7)$ — there is a 71% probability that a randomly selected HIV+ child will have a larger statistic than a randomly selected control. Volumes is followed by DTI which has fair discrimination capacity $(AUC > 0.6)$. Spectroscopic measures in the basal ganglia, peritrigonal white matter, and midfrontal gray matter show no predictive ability with AUCs of close to chance $(AUC \approx 0.5)$. Other structural measures, gyrification and cortical thickness, also have an accuracy no better than a coin flip. Sensitivity is generally higher than specificity for all individual EN classifiers.

5.2.2. Features predictive of HIV status in single modalities (objective 2)

As previously noted, feature selection is part of the classification algorithm. We present relevant features that are extracted for each of the 100 repeat cycles. Coefficients are interpreted as the change in dependent variable with one standard deviation increase in

predictor value, with all other predictors held constant. A negative sign indicates a smaller measure in HIV+ children than in healthy controls.

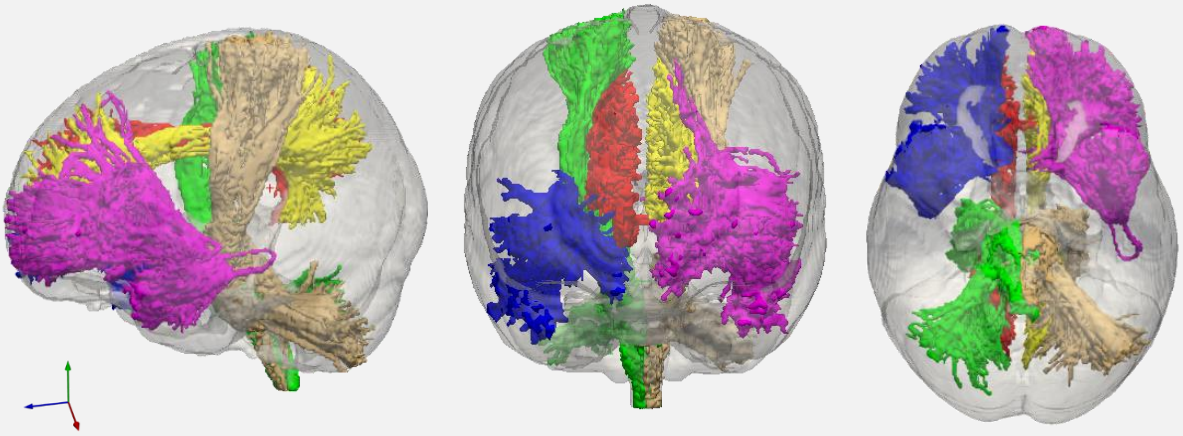
For sMRI volumes, 12 features were found with a frequency > 75%. Within the relevant features, the highest ranked in terms of weighting (weighted frequency) is right vessel and the lowest is right cerebellum white matter. The anatomical location of the features, and their rankings are shown in Figure 5.1. Weightings for 11 of the relevant volume features have a negative association with HIV. Only CSF has a positive weighting in the EN regression model.



Relevant Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Right vessel	-2.37	100	2.37
Optic Chiasm	-0.65	98	0.64
Left Amygdala	-0.26	99	0.25
Left Pallidum	-0.20	100	0.20
Right Pallidum	-0.18	100	0.18
Right choroid plexus	-0.15	95	0.15
Right amygdala	-0.09	93	0.08
Posterior corpus callosum	-0.09	78	0.07
Cerebrospinal fluid (CSF)	0.06	93	0.05
Right ventral diencephalon (DC)	-0.05	84	0.05
Left ventral DC	-0.03	77	0.02
Right cerebellum white matter	-0.01	90	0.01

Figure 5.1| Selected volume structures in situ—Relevant features from sMRI volumes with their mean weight, frequency, and weighted frequency.

For DTI, 13 relevant features were found which are listed in Figure 5.2. Differences in diffusion measures are restricted to three tracts in both hemispheres: uncinate fasciculus, corticospinal tract, and cingulum. All selected features, comprising axial, radial and mean diffusivities in these tracts, are positively associated with HIV.



Relevant Features	$\bar{\beta}$	Frequency	|Weighted frequency|
RD in **left uncinate fasciculus**	5.50	100	5.50
MD in **right uncinate fasciculus**	4.10	100	4.10
MD in **left corticospinal tract**	3.87	100	3.87
MD in **right cingulum**	3.50	95	3.33
RD in right uncinate fasciculus	3.11	100	3.11
RD in left corticospinal tract	3.04	100	3.04
RD in **right corticospinal tract**	3.06	98	2.99
AD in right uncinate fasciculus	3.69	86	3.17
MD in left uncinate fasciculus	2.53	99	2.51
MD in **left cingulum**	2.49	98	2.44
AD in left corticospinal tract	2.28	95	2.16
RD in left cingulum	2.15	93	2.00
MD in right corticospinal tract	1.26	94	1.18

Figure 5.2| Selected DTI tracts in situ—Relevant features from DTI with their mean weight, frequency, and weighted frequency.

The EN classifier using peritrigonal white matter spectroscopy data performs poorly in terms of classification ($AUC = 0.49$). Accordingly, there are no features that meet the threshold frequency for relevance, as all of them are selected approximately half of the time in a 100-repeat cycle. The top 6 most frequently selected features with highest weights are listed in Table 5.4.

Table 5.4|Selected features for ^{1}H-MRS in peritrigonal white matter (pwm) classifier

Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Lipids at 20ppm ratio to creatine	-9.07	52	4.72
N-acetylaspartylglutamate ratio to creatine	-4.48	50	2.24
Lipids at 13ppm ratio to creatine	4.21	50	2.11
Macromolecules at 20ppm	6.24	40	2.49
Total choline ratio to creatine	-3.14	40	1.26
Glycerophosporylcholine ratio to creatine	2.01	40	0.80

--Features have no class discrimination capability--

Results of feature selection for spectroscopy in the midfrontal gray matter are shown in Table 5.5. With an AUC of only 0.58, these features have no class discrimination ability. There are no relevant features selected as the output shows the EN classifier selecting one feature just over half of the time.

Table 5.5| Selected features for ^{1}H-MRS in midfrontal gray matter (mfgm) classifier

Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Total choline ratio to creatine	9.184	55	5.051
N-acetylaspartate	9.115	54	4.922
Lipids at 20ppm ratio to creatine	-4.671	59	2.756
Glutamate ratio to creatine	-2.626	66	1.733
N-acetylaspartylglutamate ratio to creatine	2.960	58	1.717
Myo-inositol ratio to creatine	-1.524	68	1.036

--Features have no class discrimination capability--

Features from basal ganglia spectroscopy similarly have no class discrimination capability. The top-ranked features are listed in Table 5.6.

Table 5.6| Selected features for ^{1}H-MRS in basal ganglia (bg) classifier

Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Lipids at 13ppm ratio to creatine	-8.40	56	4.71
Lipids at 20ppm ratio to creatine	11.88	32	3.80
N-acetylaspartylglutamate ratio to creatine	-10.70	30	3.21
Lipids at 20ppm	9.10	30	2.73
Glycerophosphocholine ratio to creatine	7.02	36	2.53
Creatine	4.87	46	2.24

--Features have no class discrimination capability--

Cortical thickness models also show no class discrimination capacity ($AUC = 0.54$). Although no features are selected on more than 75 iterations, several features are selected at least 60% of the time. The weight, frequency, and weighted frequency for the top 6 ranked features are shown in Table 5.7.

Table 5.7| Selected features for sMRI cortical thickness classifier.

Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Right rostral middle frontal	3.54	68	2.40
Left inferior parietal	3.41	68	2.32
Left lateral occipital	2.99	63	1.88
Left supramarginal	2.72	65	1.77
Right parahippocampal	-2.81	62	1.74
Left banks of superior temporal sulcus	2.73	63	1.72

--Features have no class discrimination capability--

The last single modality EN classifier is sMRI gyrification, which has poor performance ($AUC = 0.54$) and does not have relevant features meeting the feature selection threshold. The top 6 ranked features, which are selected almost half of the time, are shown in Table 5.8.

Table 5.8| Selected features for sMRI gyrification EN classifier

Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Right lateral occipital	9.76	51	4.98
Left cuneus	-5.10	58	2.96
Right supramarginal	-4.59	45	2.07
Right superior parietal	-3.98	47	2.06
Left parahippocampal	-3.89	40	1.56
Right rostral anterior cingulate	-2.75	49	1.35

--Features have no class discrimination capability--

5.3. Non-nested multimodal elastic net classifiers

5.3.1 Classification performance (objective 3)

The best performing individual feature set is sMRI volumes ($AUC = 0.71$). The best combination obtained from concatenating this set with each of the remaining feature sets is sMRI volumes + DTI, with an AUC of 0.78. The measure of separability (AUC) between HIV+ and healthy controls obtained by concatenating this multimodal set with the rest of the individual feature sets is shown in Table 5.9. The model sMRI volumes + DTI + ^{1}H-MRS bg is the best performing multimodal feature set, as there is no further improvement in AUC when further feature sets are added.

Table 5.9| Multimodal classification performance in steps of improving AUC (in red)—The first column gives the feature set combination with which the feature set in each column is concatenated. Only subjects with data for all measurements were included. Since the number of subjects n was different for each modality, each concatenation step was performed on a smaller sample than the previous one.

Concatenation Steps	sMRI volumes	DTI all regions	^{1}H-MRS bg	^{1}H-MRS mfgm	sMRI thickness	sMRI LGI	^{1}H-MRS pwm
1. —	0.71	0.62	0.55	0.58	0.55	0.54	0.49
2. sMRI volumes	—	0.78	0.75	0.68	0.61	0.60	0.62
3. sMRI volumes + DTI	—	—	0.80	0.78	0.72	0.65	0.76
4. sMRI volumes + DTI + ^{1}H-MRS bg	—	—	—	0.77	0.73	0.64	0.76

Classification performance evaluation metrics for the best performing feature sets in each concatenation step are given in Table 5.10. Sensitivity, specificity, and accuracy, as well as AUC improve with each concatenation step.

Table 5.10| Classification performance metrics of non-nested multimodal analyses

Modality	AUC	Sensitivity	Specificity	Accuracy
sMRI volumes	0.71	81.18%	66.89%	74.04%
sMRI volumes + DTI	0.78	82.65%	80.26%	81.45%
sMRI volumes **+** DTI **+** ^{1}H-MRS bg	0.80	84.38%	88.75%	86.56%

5.3.2. Comparing goodness of fit of single vs multimodal models via AIC (objective 4)

Because the ratio of sample size to number of estimated parameters (*n:k*) for individual non-nested models was always less than 40, the second order AIC (AIC_c) was computed and is shown for each model in Table 5.11. For each model the relative likelihood in comparison to the one with the smallest AIC is also shown.

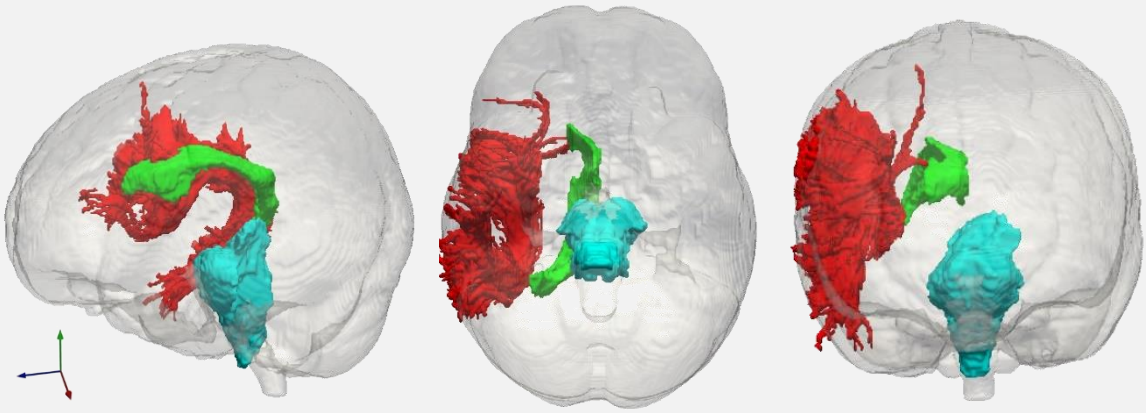
Table 5.11| Comparison of non-nested multimodal classifiers with AIC

Candidate models	k	AIC_c	Relative likelihood $e^{\frac{AIC_c min - AIC}{2}}$
sMRI volumes	17	99.60	0.013
DTI	17	102.87	0.002
^{1}H-MRS bg	30	163.12	2.056×10^{-16}
sMRI volumes + DTI	27	90.88	1.000
sMRI volumes + DTI + ^{1}H-MRS bg	29	102.62	0.002

The sMRI volumes + DTI model has lowest AIC and is thus the best model to minimise information loss. The model using sMRI volumes alone is 0.013 times less likely than the sMRI volumes + DTI model to minimise information loss. Furthermore, sMRI volumes + DTI + ^{1}H-MRS bg model is only 0.002 less likely to minimize information loss than the best model.

5.3.3. Features predictive of HIV status (objective 5)

We present relevant features from the best performing multimodal combinations—namely, sMRI volumes + DTI in step 2 and sMRI Volumes + DTI + ^{1}H-MRS bg in step 3. For the sMRI volumes + DTI combination, 22 relevant features are found. Three features appear in the multimodal combination that are not relevant features in the single modality analyses: volumes of brainstem and right lateral ventricle and AD in the temporal part of the right superior longitudinal fasciculus (illustrated in Figure 5.3). Conversely, volumes of the posterior corpus callosum, right choroid plexus, and right ventral diencephalon which are relevant features in the single modality sMRI volumes EN model appear with a frequency of less than 60% in this multimodal combination. Similarly, RD and MD in the right corticospinal tract are relevant features in the single modality DTI model but are not relevant in the multimodal combination.



--22 out of 122 features selected (only 3 shown)--

Relevant Features	$\bar{\beta}$ x10^{-2}	Frequency	Weighted frequency x10^{-2}
Right vessel	-9.44	100	9.44
RD in left uncinate fasciculus	6.10	100	6.10
Left amygdala	-6.09	100	6.09
Left pallidum	-6.06	99	6.00
Optic chiasm	-5.47	100	5.47
Right amygdala	-5.39	100	5.39
Right pallidum	-5.22	100	5.22
Left ventral DC	-4.65	100	4.65
RD in right corticospinal tract	4.21	100	4.21
MD in right uncinate fasciculus	4.14	100	4.14
MD in left uncinate fasciculus	3.53	100	3.53
RD in right uncinate fasciculus	3.39	100	3.39
AD in right uncinate fasciculus	2.92	100	2.92
Cerebrospinal fluid	2.51	93	2.33
Right cerebellum white matter	2.32	92	2.14
MD in right cingulum	1.89	99	1.87
MD in right corticospinal tract	1.74	100	1.75
Brainstem	-1.66	98	1.62
Right lateral ventricle	-1.18	93	1.10
MD in left cingulum	1.11	97	1.08
RD in right cingulum	1.11	96	1.07
AD in right superior longitudinal fasciculus temporal	1.08	93	1.01

Figure 5.3| Selected features of sMRI volumes + DTI in situ

For sMRI volumes + DTI + ^{1}H-MRS bg, 24 relevant features are found (Table 5.12). Selected features are like the sMRI volumes + DTI combination, with the addition of glutamate to creatine ratio, macromolecules at 20ppm, N-acetylaspartate, guanidinoacetate to creatine ratio and macromolecules + lipids at 20ppm. Volumes of optic chiasm and cerebrospinal fluid and AD in the temporal part of the right superior longitudinal fasciculus are no longer in the model.

Table 5.12|Selected features of sMRI volumes + DTI + ^{1}H-MRS bg—Five spectroscopy in basal ganglia measures are among the relevant features select, albeit having the lowest ranking.

Relevant Features	$\bar{\beta}$ x10^{-2}	Frequency	|Weighted frequency| x10^{-2}
Right vessel	-0.97	100	0.97
RD in left uncinate fasciculus	-0.90	100	0.90
Left amygdala	-0.66	100	0.66
Left pallidum	-0.66	100	0.66
Glutamate ratio to creatine	-0.73	79	0.65
Right amygdala	-0.64	100	0.64
Right pallidum	-0.61	100	0.61
Left ventral DC	-0.67	78	0.59
RD in right corticospinal tract	-0.53	100	0.53
MD in right uncinate fasciculus	-0.62	75	0.53
MD in left uncinate fasciculus	0.51	100	0.52
RD in right uncinate fasciculus	0.50	100	0.50
AD in right uncinate fasciculus	-0.48	100	0.48
Macromolecules at 20 ppm	0.44	100	0.44
Right cerebellum white matter	-0.43	100	0.43
MD in right cingulum	-0.41	100	0.41
MD in right corticospinal tract	-0.38	100	0.38
Brainstem	0.38	90	0.34
Right lateral ventricle	-0.36	77	0.31
MD in left cingulum	0.26	100	0.26
RD in right cingulum	-0.29	79	0.23
N-acetylaspartate	0.18	100	0.18
Guanidinoacetate ratio to creatine	0.17	100	0.17
Macromolecules + lipids at 20ppm	-0.22	78	0.17

--24 out of 150 features selected--

5.4. Nested multimodal elastic net classifiers

5.4.1. Classification performance (objective 3)

Each sequence of nested classifiers starts with the sMRI volumes or DTI feature sets, since their relevant features had already been extracted. Features selected in the previous phase in the sequence are forced to be included in training of the subsequent multimodal EN classifier. In general, AUC improves when adding more predictors, up to a limit beyond which the algorithm does not converge (Table 5.13).

Table 5.13| Classification performance of nested models—Shown in red is the concatenation sequence leading to the highest AUC.

Starting modality (Initial AUC)	+ sMRI volumes	+ DTI all regions	+ ^{1}H-MRS bg	+ ^{1}H-MRS mfgm	+ sMRI thickness	+ sMRI LGI	+ ^{1}H-MRS pwm
sMRI volumes (0.71)	—	0.817	0.823	~*	~*	~*	~*
DTI (0.62)	0.809	—	0.813	~*	~*	~*	~*

~* alpha, lambda does not converge after 10^{8} iterations; parameters can't be tuned/optimised for the multimodal feature sets.

Other classification performance metrics of the best performing nested multimodal combinations are presented in Table 5.14. Measures of sensitivity, specificity and accuracy also improved with the addition of more features.

Table 5.14| Classification performance metrics of best performing nested sequence

Modality	AUC	Sensitivity	Specificity	Accuracy
sMRI volumes	0.71	81%	67%	74%
sMRI volumes + DTI	0.81	85%	79%	82%
sMRI Volumes + DTI + ^{1}H-MRS bg	0.82	86%	80%	83%

5.4.2. Goodness of fit using the LRT (objective 4)

Results of the LRT on the nested multimodal models producing the highest AUCs are presented in this section. We show the formal comparison of the sMRI volumes + DTI model to the nested sMRI volumes model, sMRI volumes + DTI + ^{1}H-MRS bg to sMRI volumes + DTI in Table 5.15.

Table 5.15| Likelihood ratio test of nested models

Candidate models	DF*	Log likelihood	χ^2	Pr (>χ^2)
sMRI volumes	17	-26.80	—	—
sMRI volumes + DTI	29	-51.38	49.15	1.96×10^{-6}
sMRI volumes + DTI	29	-51.38	—	—
sMRI volumes + DTI + ^{1}H-MRS bg	37	-49.13	4.50	0.81

*Degrees of freedom (DF) = number of selected features + confounders + deviance + 1

The sMRI volumes + DTI model minimises the negative log-likelihood, therefore the multimodal classifier is an improvement to the smaller model i.e. sMRI volumes. When adding spectroscopy in the basal ganglia—sMRI volumes + DTI + ^{1}H-MRS bg—the multimodal classifier does not improve on the smaller model (sMRI volumes + DTI) because it does not minimise the negative log-likelihood.

5.4.3. Comparison of non-nested vs nested models

The final check was whether relevant features selected through non-nested EN classification are a better fit to the data than the nested model. We present the AIC_c test results in Table 5.16.

Table 5.16| Comparison of nested and non-nested models with AIC

Candidate models	k*	AIC_c	Relative likelihood $e^{\frac{AIC_c min - AIC}{2}}$
sMRI volumes + DTI (non-nested)	27	90.88	—
sMRI volumes + DTI (nested)	29	102.62	0.003
sMRI volumes + DTI + ^{1}H-MRS bg (non-nested)	29	102.62	—
sMRI volumes + DTI + ^{1}H-MRS bg (nested)	37	164.71	3.201 x10^{-14}

*k = number of selected features + confounders + deviance + 1

The non-nested sMRI volumes + DTI model offers a better fit than its nested counterpart which was only 0.003 times as probable to minimise information loss despite offering a slightly better AUC improvement (78% AUC vs 81% AUC). Similarly, non-nested sMRI volumes + DTI + ^{1}H-MRS bg model fits better to the data than that generated through nested concatenation.

5.4.4. Features predictive of HIV status (objective 5)

A total of 32 relevant features were selected for the nested multimodal model producing the best performance ($AUC = 0.823$). Figure 5.4 illustrates the nested feature selection procedure and relevant features extracted in each phase; the sequence terminates at sMRI volumes + DTI + ^{1}H-MRS bg.

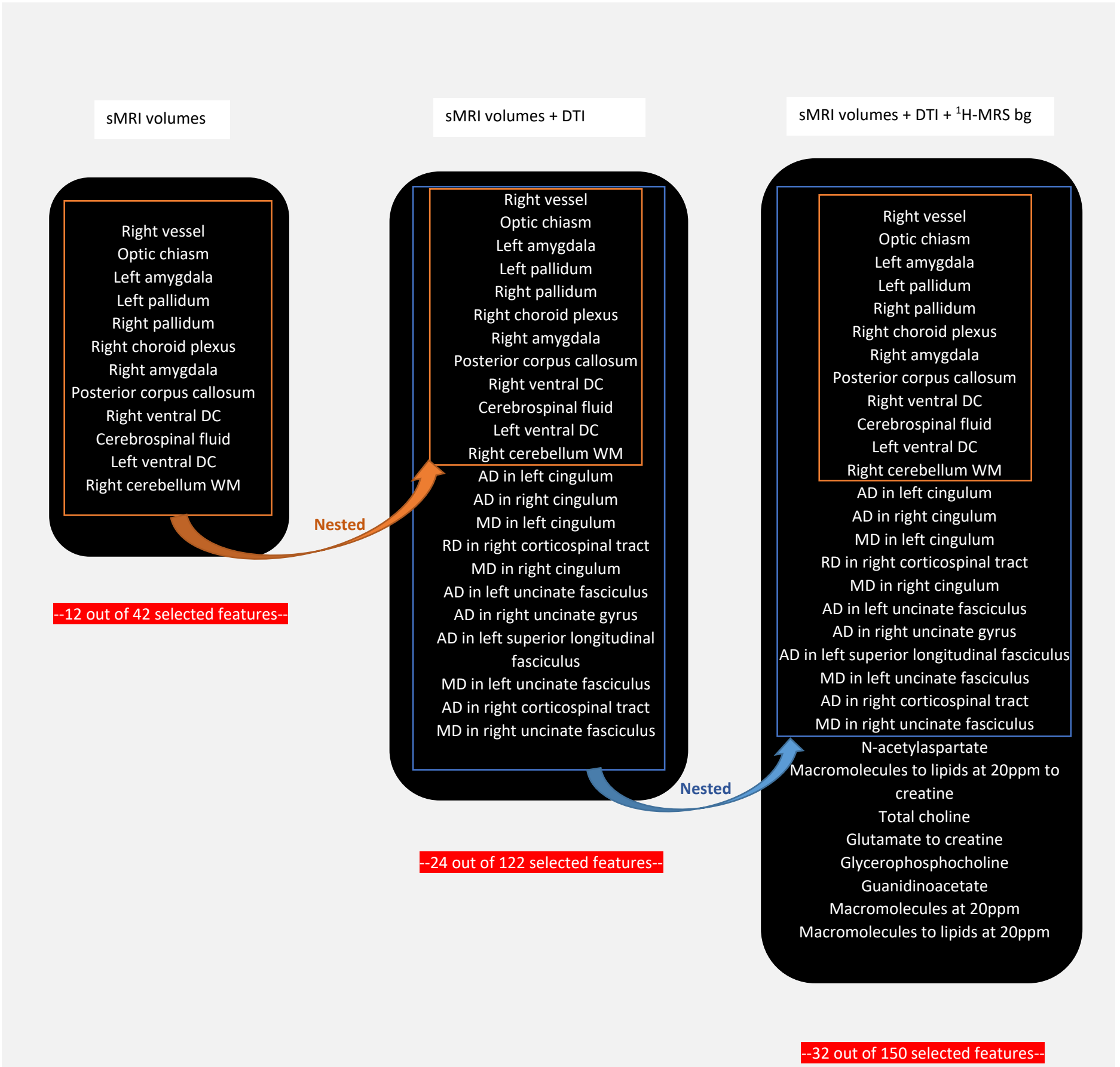


Figure 5.4| Feature selection of nested multimodal classifier—Shown are relevant features in each sequence phase until there is no convergence of penalization parameters.

5.5. Effect of increasing sample size via imputation of missing data

To investigate the feasibility of MICE imputation and the effect of sample size on classification and feature selection, we imputed missing values for the best performing multimodal feature set identified previously — sMRI volumes + DTI. Computation of classification performance metrics and extraction of relevant features for imputed datasets was the same as for non-imputed feature sets (excluding subjects with missing data).

5.5.1. Classification performance

We present classification performance metrics for the best performing dataset (sMRI volumes + DTI) compared to its 5 MICE-imputed datasets in Table 5.17.

Table 5.17| Classification performance of imputed feature set with non-imputed model

Modality	n	AUC	Sensitivity	Specificity	Accuracy
sMRI volumes + DTI	102	0.78	82.65%	80.26%	81.45%
Imputed set 1	127	0.75	78.13%	84.55%	81.34%
Imputed set 2	127	0.81	81.25%	87.50%	84.38%
Imputed set 3	127	0.78	75.00%	87.50%	81.25%
Imputed set 4	127	0.78	79.69%	91.67%	85.68%
Imputed set 5	127	0.68	78.13%	83.33%	80.73%

Imputed datasets (large sample size) have a higher specificity (mean 86.91 ± sd 3.25) but lower sensitivity (mean 77.87 ± sd 2.51) and accuracy (mean 82.67 ± sd 0.02) than models created using fewer observations. The AUC (mean 0.76 ± sd 0.05) of imputed datasets is within the margin of error of that of the model with fewer observations.

5.5.2. Feature selection

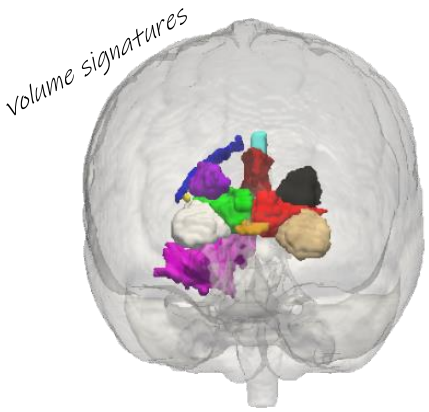
Relevant features were extracted from individual imputed multimodal EN classifiers with the same analysis and thresholding as before. Table 5.18 lists the selected features obtained from the non-nested SMRI volumes + DTI model, showing whether each feature was still relevant in the model when using imputed data. Features with a tick are those found in both imputed and non-imputed datasets.

Eleven similar features are selected regardless of the inclusion of imputed data. However, CSF and AD in right superior longitudinal fasciculus temporal are never selected when imputed data is included.

Table 5.18| Selected features of best performing model that are also selected in imputed sets

Selected features	Imputed set 1	Imputed set 2	Imputed set 3	Imputed set 4	Imputed set 5
Left pallidum	✓	✓	✓	✓	✓
RD in left uncinate fasciculus	✓	✓	✓	✓	✓
Right vessel	✓	✓	✓	✓	✓
Left amygdala	✓	✓	✓	✓	✓
Optic chiasm	✓	✓	✓	✓	✓
Right amygdala	✓	✓	✓	✓	✓
Right pallidum	✓	✓	✓	✓	✓
Left ventral DC	✓	✓	✓	✓	x *
RD in right corticospinal tract	x	x	x	✓	✓
MD in right uncinate fasciculus	✓	✓	✓	✓	✓
MD in left uncinate fasciculus	✓	✓	✓	✓	x **
RD in right uncinate fasciculus	✓	x	x	x	x
AD in right uncinate fasciculus	✓	✓	✓	✓	✓
Right cerebellum white matter	✓	✓	✓	✓	✓
Cerebrospinal fluid	x	x	x	x	x
MD in right cingulum	✓	x	x	✓	x
MD in right corticospinal tract	✓	x	✓	✓	✓
Brain stem	✓	x	x	✓	✓
Right lateral ventricle	x	✓	x	x	x
MD in left cingulum	✓	✓	x	x	x
RD in right cingulum	✓	✓	x	x	x
AD in right superior longitudinal fasciculus temporal	x	x	x	x	x

Non-selected (x) and selected feature (✓)
*Right ventral DC selected instead
**RD in left uncinate fasciculus selected instead



6

Discussion & conclusion: significance of findings, limitations and recommendations

Contents

6.1. Hypothesis 1: a multimodal approach discriminates better between HIV+ children on cART and HIV- controls than single modality analyses 67

6.2. Hypothesis 2: multimodal neuroimaging measurements characteristic of cART-treated perinatally acquired HIV are indentifiable 68

6.3. Selected features: comparison with previous findings 69

6.3.1. Structural volumes alterations in perinatal HIV 70

6.3.2. White matter microstructure alterations in perinatal HIV 74

6.3.3 Features not strongly predictive of HIV status 76

6.4. Limitations and recommendations for future work 78

6.4.1. On generalisable results 78

6.4.2. On confounding 80

6.4.3. On ROI vs voxel-based analysis of DTI 81

6.4.4. On segmentation 81

6.4.5. On spectroscopy 82

6.4.6. On longitudinal analysis 82

6.4.7. On classification 82

6.5. Summary and conclusion 83

Neural development during human childhood includes coordinated and sequential processes that depend on both genetic and environmental factors. As the brain develops, neurometabolite concentrations are maintained to avoid defects in oxidative metabolism, atypical proliferation of glial cells, irregular neuronal and axonal density (Bertholdo et al., 2013; Blüml & Panigrahy, 2013; Považan et al., 2015). The maturation of subcortical gray matter, white matter, total brain volume, cortical thickness, and gyrification complexity supports development of neural circuitry, neuro-generation and fine motor control (Kotov, 2017; Spalleta & Fabrizio, 2018; Thompson et al., 2005; Trefler et al., 2016). Myelination of white matter tracts results in increased fibre coherence, axonal diameter, as well as conduction speed of nerve impulses and functional connectivity (Lentz et al., 2014; Brouwer et al., 2010; Snook et al., 2007 and 2005; Beaulieu et al., 2005; Schneider et al., 2004).

This makes the developmental phase between 5 and 10 years an especially vulnerable time when children are susceptible to developing structural, cognitive and functional abnormalities due to perinatally acquired HIV or other neurological diseases. It is therefore important to investigate neurological markers of cART-treated paediatric HIV to identify changes that may hinder normal development. While previous studies have looked at neuroimaging modalities separately (Jankiewicz et al., 2017; Nwosu et al., 2018; Robertson et al., 2018), no study to date has combined neuroimaging modalities and/or used a multivariate pattern recognition approach to identify features that can be used for classification of treated or untreated paediatric HIV infection.

6.1. Hypothesis 1: a multimodal approach discriminates better between HIV+ children on cART and HIV- controls than single modality analyses

In line with our hypothesis, multimodal neuroimaging improves classification performance in cART-treated paediatric HIV compared to single modalities. We began with classification accuracies of 61%, 72%, and 74% for metabolites in the basal ganglia, diffusion measures, and volume structures respectively, to a multimodal combination accuracy of 87%. This improvement is in line with investigations implementing similar embedded feature selection techniques where classification accuracy of schizophrenia, dementia, Parkinson's disease,

depression and ADHD increased by 2—7% with a multimodal approach (Bowman et al., 2016; Colby et al., 2012; Kambeitz et al., 2017).

The best performing multimodal classifier for distinguishing HIV+ children from controls consists of subcortical sMRI volumes and DTI features resulting in 83% sensitivity, 80% specificity, 81% accuracy and a 78% AUC. This is an improvement on classification performance of single modalities of both volumes (81% sensitivity, 67% specificity, 71% AUC) and DTI (75% sensitivity, 69% specificity, 62% AUC). Similar improvements in performance have been observed when sMRI and DTI are combined in other pathologies like Alzheimer's disease (Schouten et al., 2016), schizophrenia (Sui et al., 2015), mild cognitive impairment (Ritter et al., 2015), and Parkinson's disease (Bowman et al., 2016). Within individual neuroimaging modalities, embedded feature selection methods (LASSO, SVM and chained regularization), have previously been found to discriminate between HIV+ and HIV- adults with accuracies of greater than 80% using sMRI (Adeli et al., 2018) and DTI (Tang et al., 2017) data. However, AUCs for these studies was on average better than ours, for example, $AUC > 0.9$ in Tang et al. (2017), because the authors implemented more powerful embedded techniques or stability selection with their classification and feature selection algorithms.

6.2. Hypothesis 2: multimodal neuroimaging measurements characteristic of cART-treated perinatally acquired HIV are identifiable

We find a neuroimaging signature of cART-treated paediatric HIV comprising 22 stable features from the multimodal combination of subcortical volumes and DTI— these include smaller bilateral globus pallidus and amygdala as well as increased radial and mean diffusivities (RD and MD) in the uncinate fasciculus and right corticospinal tract.

Although most of the selected predictors in the multimodal combination are simply the combined predictors selected in the 2 single modality EN models, the multimodal EN model replaces 5 less relevant features from these models with 3 better ones when provided with the entire multimodal feature set. The posterior corpus callosum, right choroid plexus, right diencephalon, RD and axial diffusivity (AD) in left corticospinal tract are selected in the single modality analyses, but are replaced by brainstem, right lateral ventricle, and AD in right superior temporal longitudinal fasciculus in the multimodal analysis without any information

loss. This implies that when sMRI and DTI are combined, the latter 3 features provide better, or at least equivalent, explanatory power than the former 5.

The ability to replace a less useful feature from one modality with a feature from another is a strength of the non-nested model building process. Although the nested sMRI volumes+DTI+^{1}H-MRS bg model has a marginally higher AUC (0.82) than the corresponding non-nested model (0.80), and the nested sMRI volumes + DTI model has a higher AUC (0.81) than the non-nested model (0.78), the non-nested model in each case has fewer parameters and consequently a lower AIC than its nested counterpart, suggesting that these models would generalise better to new data.

6.3. Selected features: comparison with previous findings

Building EN models using data from individual modalities, we find that structural volumes (81% sensitivity, 67% specificity) and DTI metrics (75% sensitivity, 69% specificity) produce better discrimination between HIV+ children and controls than other measures considered (Table 5.3). In agreement with previous studies (Ackermann et al., 2019, 2016; Blokhuis et al., 2019, 2017; Wade et al., 2019; Yadav et al., 2017; Lewis-de los Angeles et al., 2016), the hallmark of cART-treated paediatric HIV appears to be irregularities in subcortical structure and white matter integrity. A high concentration of the virus in central periventricular white matter and basal ganglia has been attributed to neuro-invasion via perivascular pathways since early investigations (e.g. Belman et al., 1986; Dickson et al., 1989).

The 22 features (Figure 5.3) identified using multimodal EN models in this dissertation confirm the impact HIV has on subcortical volumes and diffusion in white matter even in the presence of cART. We infer that there are more consistent effects of cART-treated HIV on subcortical structure and WM microstructure at age 7 than on cortical thickness, gyrification, and neurometabolite concentrations, even though several previous studies in the CHER cohort did find HIV-related alterations in these measures (Nwosu et al., 2018; Robertson et al., 2018; Mbugua et al., 2016). With EN regularisation, a feature that is not on its own highly predictive of HIV can be discovered when combined with another feature or a group thereof. It may explain measures in our models that have not been found to be altered in previous investigations in the same cohort.

6.3.1. Structural volumes alterations in perinatal HIV

Using EN regularisation, we can consider a larger set of candidate features and discover previously unknown neuroimaging measures that are characteristic of paediatric HIV, without being constrained by multiple comparison correction as in traditional statistical analyses. Many of the FreeSurfer structural volumes that appear with high frequency in EN models were not examined in previous single modality analyses (Nwosu et al., 2018), as there was no strong *a priori* hypothesis about their sensitivity to effects of perinatal HIV infection. These include the globus pallidus, amygdala, cerebellum white matter and CSF.

Although Nwosu et al. (2018) found reduced putamen volumes in the CHER cohort at age 7, the putamen is not a relevant feature in our single modality or multimodal EN models in the same population. This suggests that smaller putamen may not be a robust finding across perinatally HIV+ children and may reflect a change that is specific to a HIV+ subgroup. In fact, Blokhuis and colleagues (2017) reported trend-level putamen *increases* in children aged between 8 and 18 years and Randall et al. (2017) similarly reported increases in putamen size at age 5 in the CHER cohort. Despite these findings, our results suggest that HIV may affect globus pallidus more than other components of the basal ganglia.

The globus pallidus regulates voluntary movements that occur on the subconscious level (Gillies et al., 2017). It is the important structure of the basal ganglia and stabilities excitatory state of cerebellum to allow controlled smooth movement. HIV infection of the CNS impacts mostly spiny neurons in the putamen and medium-sized neurons in the globus pallidus, as well as large pyramidal neurons in the neocortex (Kovalevich & Langford, 2012), and up to 33% of HIV+ children show bilateral and symmetric calcifications involving predominantly the globus pallidus and putamen (Donzuso et al., 2019; De Girolami et al., 2004).

In a cohort of Thai children with HIV, there was a shape deformation of the right pallidal surface and baseline CD4 count was negatively associated with left globus pallidus volume (Wade et al., 2019). Shape deformation has also been associated with peak viral load in adolescents (Lewis-de los Angeles et al., 2016). Although our models suggest reduced volume of the left and right globus pallidus in HIV, larger globus pallidus volumes have also been

observed in HIV+ children aged less than 12 years, but not in older children (Paul et al., 2018). The effect of HIV on the globus pallidus may therefore vary with factors such as age and past HIV disease severity.

Our EN models also reflect an association of HIV infection and reduced volumes of the right and left amygdala, structures involved in emotion-regulation processes. The left amygdala is involved specifically with sustained emotional evaluation, while right amygdala plays a crucial role in dynamic emotional stimulus detection (Baeken et al., 2014). Amygdala volume alterations are not commonly reported in HIV infection but may be a consequence of HIV-associated stressors, rather than a direct result of HIV itself. However, a study of HIV orphans found smaller hippocampal volumes, but no difference from control children in amygdala (Zuo et al., 2019), suggesting no effect of early life adversity on amygdala volume. Although a reduction in amygdala volume in our cohort may result from HIV-associated stress factors, the hippocampus was not a relevant feature in our models. Moreover, in the PREDICT cohort in Thailand, Wade et al. (2019) found associations of CD4 count with local surface area of the bilateral posterior and left anterior amygdala, suggesting an association with disease severity. In addition, larger volumes of the amygdala have been observed in HIV+ children under 12 years (Paul et al., 2018).

In pre-cART studies (e.g. George et al., 2009), the subcortical frontal WM and the cerebellum was found to calcify in paediatric HIV+ patients. The right cerebellum controls much of the movement involving the right side of the body, and as abovementioned, its excitatory motion is balanced by the globus pallidus. Our finding of reduced cerebellar WM volume in combination with globus pallidus volume may indicate motor dysfunction. In early studies, mild cerebellar dysfunction was observed in some HIV+ children, with lesions in the right cerebellar WM (Liptai et al. 2007; Nuttall et al. 2004). The subjects studied usually had progressive multifocal leukoencephalopathy which is infrequent in children compared to adults and occurred before access to cART became widespread.

Ventricular enlargement and/or sulcal widening are among the more common neuroimaging findings in children with HIV on cART (van Arnhem et al., 2013). Collective expansion of the ventricular system is a feature of cART-treated HIV (Thompson et al., 2006). As such, our finding of increased CSF in HIV+ children is not unexpected. The FreeSurfer definition of CSF refers to the collective ventricular system including both lateral ventricles, third ventricle, and

fourth ventricle (Filipek et al., 1994). The HIV virus enters the brain primarily via the CSF or fluid-filled spaces such as the lateral ventricles. This may lead to HIV-infected mononuclear cells easily penetrating the ventricles, leading to neuronal atrophy in nearby regions (Ances et al., 2012). Pre-cART, HIV replication in CSF was associated with neurocognitive impairments in children (Pratt et al., 1996; Sei et al., 1996), and changes to the brain water content and nutritional state of HIV patients were responsible for enlarged CSF spaces. However, since the introduction of cART, viral load in plasma and CSF is often below the detection limit (Steinbrink et al., 2013).

Surprisingly, despite increased total CSF, our EN models suggest reduced volume of the right lateral ventricle in HIV infection. In HIV+ adults, lateral ventricles are frequently enlarged (Thompson et al., 2015, 2005; Wang et al., 2010). Previous studies have shown that HIV+ children have lateral ventricles that are asymmetrical (Thompson & Jahanshad, 2015; Bruck et al., 2001), have T2 lengthening (Prado et al., 2011), and dilate with total brain volume loss (Mariam & Assefa, 2012). In addition to the neuroimaging features described above, some unusual features that are fairly specific to FreeSurfer segmentation were found to be relevant predictors in the models. These are the brainstem, right vessel of the basal ganglia, optic chiasm and left ventral diencephalon.

Brainstem morphometry is not often examined. The brainstem is of prime importance in conveyance of motor pathways such as corticospinal tract and is a sensory pathway for the auditory nerve. There are many studies that reported on hearing impairment in HIV+ children on cART (Torre et al., 2012; Hrapcak et al., 2016; Chao et al., 2012; Laughton et al., 2012) but the nature of most hearing deficits are largely unconfirmed, and commonly associated with incidences of otitis media or abnormalities in auditory brainstem responses. The first relay of the primary auditory pathway occurs in the cochlear nuclei in the brainstem, where the basic decoding of duration, intensity, and frequency signal occurs. We speculate that the brainstem may be playing a role in slight hearing loss in HIV+ children than previously thought.

The left and right vessel are small labels that FreeSurfer includes for vessels entering the basal ganglia that would otherwise be labelled as lesions (Fischl et al., 2002; Filipek et al., 1994). Perivascular or Virchow-Robin spaces form around these blood vessels (lenticulostriate arteries) and appear with the same intensity as CSF (hypointense) on T1-weighted image (Kwee, 2007). These vessels are not highly correlated predictors, which can be inferred from

the asymmetrical selection of the right vessel as the only relevant feature. Reduced volume of the right vessel in HIV+ children may reflect altered blood flow to the basal ganglia, which has previously been detected in cART-treated perinatally HIV-infected children with mild psychomotor symptoms (Blokhuis et al., 2017). But, the right vessel and globus pallidus are consistently highly ranked predictors displaying a strong association, which suggests that an altered blood flow only in the right vessel results in HIV disproportionately affecting the globus pallidus than other parts of the basal ganglia or vice versa.

The optic chiasm is a small visual pathway—part of the brain where optic nerves partially cross. Located at the bottom of the brain immediately inferior to the hypothalamus, and essential for binocular vision and depth perception (Herrera et al., 2019). The literature on ocular manifestations of HIV infection in children is very limited because of their low prevalence (Padhani et al., 2000; Dennehy et al., 1989). By contrast, in the pre-cART era between 40 and 90% of HIV+ adults had ocular involvement (Kestenyl et al. 1990; Freeman et al. 1989). Furthermore, differences found in optic chiasm are usually benign, and leave major aspects of visual function intact (Ather et al., 2019; Puzniak et al., 2019).

FreeSurfer ventral diencephalon label contains the hypothalamus, its main component, along with basal forebrain, sub-lenticular extended amygdala, and large portion of the ventral tegmentum (Makris et al., 2013, 2008). Volume of the hypothalamus is difficult to measure *in vivo* as it is a group of nuclei not a discrete structure (Makris et al., 2013) and these are not easily distinguishable in standard MRI images. As such, other investigators have used this FreeSurfer-defined ROI as the best representation of the neuroendocrine or homeostatic system in other pathologies such as paediatric neurotrauma (Bigler et al., 2019), major depression disorder (Kim et al., 2019), and first-episode schizophrenia (Emsley et al., 2015).

Although volume alterations of the hypothalamus may contribute to endocrine dysfunction, we did not find previous neuroimaging studies directly measuring volume of hypothalamus or other parts of the ventral diencephalon in HIV+ patients of any age. However, endocrine dysfunction has been extensively reported in HIV+ children (Pandey et al., 2018; Loomba-Albrecht et al., 2014; Chantry et al., 2007; Rondanelli et al., 2002). The authors speculate that HIV may have an impact on the hypothalamus—the neural control centre for the endocrine system. These studies focused on growth hormone deficiency and/or the hypothalamic-

pituitary-adrenal axis, which is pivotal in maintenance of basal and stress-related homeostasis in HIV+ children.

6.3.2. White matter microstructure alterations in perinatal HIV

In terms of WM microstructure, our findings include increased RD and/or MD but no FA alterations, implicating various white matter tracts including the corticospinal, uncinate fasciculus, superior longitudinal fasciculus and cingulum, most of which were not implicated in previous analysis at 7 years (Jankiewicz et al. 2017).

We find increased RD and MD in the right corticospinal tract, suggestive of reduced myelination (Lebel & Beaulieu, 2011) and less-organised structure (Feldman et al., 2010) respectively. The corticospinal tract constitutes the descending motor pathway and comprises fibres that connect motor cortex with spinal cord and a primary pathway for voluntary motor function (Mori, 2007). Notably, FA in the CST is not a relevant feature in our analysis at age 7. However, in cART-treated children aged 6 to 15 years younger age has been associated with decreased FA in the right corticospinal tract and lower haemoglobin with decreased FA in the left corticospinal tract (Hoare et al., 2015).

Previous studies in the CHER cohort found the CST to be affected at age 5 (Ackermann et al., 2016), but not at age 7 (Jankiewicz et al., 2017). At age 5, Ackerman and colleagues (2016) found two clusters in right corticospinal tract with lower FA which was attributable predominantly to increased RD. Increased MD in the same tract was due to both increases in RD and AD. In a follow-up study at age 7, Jankiewicz et al. (2017) found no continued WM damage in this tract, a resolution attributable possibly to cART. The inconsistency between these results and ours may be explained by noting that in this dissertation we employ a region-of-interest (ROI) based method, where we average DTI indices of FA, RD, MD, and AD, along the entire tract (ROI), instead of voxel-based group analyses as employed by Jankiewicz et al. (2017) and Ackerman et al. (2016). This may also explain why MD in the inferior fronto-occipital fasciculus, inferior longitudinal fasciculus, superior corona radiata and anterior thalamic radiation did not feature in our models, despite regional increases being observed along these tracts in HIV+ children (Jankiewicz et al. 2017).

Similarly, in our ROI-analysis, we find increased RD and MD in bilateral uncinate fasciculus as a relevant feature of cART-treated HIV, indicating respectively decreasing myelination and poorly organised structure, although no voxel-wise group differences were found in identical clusters at age 7 years (Jankiewicz et al., 2017). We also find increased AD in right uncinate fasciculus, which suggests fibre coherence as axons strengthen along the entire tract. The uncinate fasciculus consists of fibres that connect the anterior temporal lobe, including the amygdala and hippocampal formation, to the orbital cortex (Mori, 2007). Hoare and colleagues (2015) also reported decreased FA—a measure of neuronal damage— in left uncinate fasciculus, which was associated with lower haemoglobin. Moreover, MD differences were found in clusters along bilateral uncinate fasciculus at age 5 years in the CHER cohort.

We find increased AD only in the right temporal part of the superior longitudinal fasciculus. This was the lowest ranked feature, which implies that differences between HIV+ children and controls are not consistently present. The superior longitudinal fasciculus is located at the supero-lateral side of the putamen and forms a large arc; it sends branches to the temporal lobe where the auditory cortex is located (Mori, 2007; Oishi et al., 2008). In typically developing adolescents, better working memory performance, attention and reading-related skills are associated with increased FA along this tract (Frye et al., 2010; Vestergaard et al., 2011).

Previous studies have found increased MD in superior longitudinal fasciculus in cART-treated children aged 8 to 12 years (Hoare et al., 2012) and 13 to 17 years (Li et al., 2015), while Uban et al. (2015) found increased MD in the inferior longitudinal fasciculus between 16 and 20 years of age. Others have found decreased AD in left superior longitudinal fasciculus in cART-treated children aged 6 to 15 years, which was associated with poorer fronto-striatal cognition (Hoare et al., 2015). The authors also reported that higher total protein and viral loads were associated with increased MD in this tract.

Our finding of increased MD in bilateral cingulum, and RD increase in the right cingulum only suggest demyelination of fibres in the right side of the cingulum but an overall less organisation and less dense axonal packing of the tract. The cingulum is the WM bundle that runs within the cingulate gyrus along its entire length (Mori, 2007). It is challenging to discern the implication of WM microstructural changes in this tract since the cingulate gyrus is

traditionally divided into 4 areas, each associated with a function (Vogt, 2005). However, when grouped with other parts of the limbic system such as the amygdala, it is associated with pain and emotion processing. MD and RD differences along this tract have been reported in HIV+ adults (Zhu et al., 2013), however we were unable to find studies that report on the cingulum in cART-treated children.

6.3.3. Features not strongly predictive of HIV status

While here we find no prominent markers in cortical thickness and gyrification, interestingly, Nwosu et al. (2018) found differences in the same study population cross-sectionally at age 7. cART-treated children had thicker lateral occipital cortex than controls in the left hemisphere as well as lower gyrification in large bilateral medial parietal and right temporal regions. Ignoring the low classification performance of our cortical thickness EN models, an increase in thickness of the left lateral occipital region is the third ranked feature, after thicker left inferior parietal cortex and right rostral middle frontal cortex (Table 5.7). Likewise, without considering the performance of cortical folding models, the highest ranked features are decreased gyrification of left cuneus as well as right superior parietal and rostral anterior cingulate regions (Table 5.8). Left cuneus, lateral occipital, and superior temporal regions overlap with those found by Nwosu et al. (2018). However, classification performance cannot be disregarded as it provides information about whether these features can discriminate accurately between HIV+ children and uninfected controls, and more importantly whether they can be generalised to another study population. Our findings suggest it is unlikely ($AUC \approx 0.5$) for these markers to be generalisable. We suggest that previous findings in cortical thickness and folding may have been driven by a subgroup of HIV+ children with unusually low gyrification and thickness measures, which does not have an impact on overall cross validated EN models.

A notable difference between this study and Nwosu et al. (2018) is our use of mean gyrification/cortical thickness in atlas-based regions of interest, rather than voxel by voxel comparison of these measures. If regions affected by HIV are substantially smaller than the atlas regions, differences might not be detectable using our technique. Studies in other populations have found reduced gyrification in HIV (Hoare et al., 2019; Lewis-de los Angeles et al., 2016). In a longitudinal study, Arnhem and colleagues (2013) found that ventricular

enlargement and/or sulcal widening resolves and attributed the changes to cART. From our findings, we also speculate that cART may minimise alterations in cortical thickness and folding in HIV+ children.

Spectroscopy measures in the basal ganglia, midfrontal gray matter, and peritrigonal white matter are poor predictors of HIV at age 7 according to our individual modality EN analyses. Correspondingly, in the same study population, Robertson and colleagues (2018) using linear regression models found no differences in metabolite concentrations in the basal ganglia at age 7. In our models, the highest ranked features appear less than half of the time supporting the poor class separability of the classifier and suggesting that there are no generalisable spectroscopy features that are characteristic of HIV. The top-ranked marker with 56% selection rate is ratio of lipids to creatine at 1.3 ppm which is an arbitrary spectral feature used in algorithm of LCModel that Robertson et al. (ibid.) did not consider in their analyses.

Preliminary spectroscopy results in the midfrontal gray matter using linear regression models found lower glutamate levels and higher choline in HIV+ children than in controls at age 7. The EN models using MFGM spectroscopy data showed poor performance and absolute N-acetylaspartate (NAA) and choline were not amongst the most frequently selected features. However, the ratio of choline to creatine was one of the top-ranked features, showing an elevation in HIV, which corresponds with previous findings. Other high ranked features include elevated NAA and ratio of NAA to creatine (Cr) in HIV+ children, lowered ratio of glutamate to Cr, which again is consistent with previous findings, as well as lower lipids at 20ppm/Cr which was not considered in previous analyses. However, since the generalisability of the model is poor, these effects may not be general to all HIV+ children.

***

6.4. Limitations and recommendations for future work

Neuroimaging research is frequently limited by number of observations, since in general relatively small cohorts of subjects are scanned. Our study sample size was initially 127, however this number was further reduced after quality control criteria had been applied. In addition, there were more HIV+ children than uninfected controls (Table 5.1), so prediction models are theoretically biased toward higher sensitivity than specificity. Cross validation and random shuffling of subjects during repeat cycles was performed to minimise bias, however our findings suggest that the models still tended to identify HIV+ children better than controls; the exception was spectroscopy in the basal ganglia model. Moreover, this slight bias created an unsatisfactory match of sensitivity and specificity to AUC measures.

The larger sample size does not directly translate to higher classification performance since DTI models showed better performance than cortical thickness and gyrification models, and multimodal classifiers have better performance that those from any single modality. We attempted to investigate this limitation by replacing bad or missing data with multiple imputed values. First, MICE results are inconsistent, specificity increases while sensitivity decreases when simulated data is included, so the AUC is equivalent to that obtained without using imputation (Table 5.17). Second, half of the features are selected, regardless of inclusion of imputed data, only 2 are no longer selected when imputed data is used. Although other features are dropped at least once (Table 5.18), it is difficult to make conclusions about whether observed differences in the final models result from the altered sample size or the imputation itself. The MICE algorithm is not very accurate when many missing values are imputed (note that imputation of all DTI features was performed for 25 subjects). Although this is often not possible, we would recommend that future studies use a constant and balanced number of subjects for all modalities.

6.4.1. On generalisable results

Traditional statistics used in previous studies evaluates variable relevance with p-values and correlation or regression with the dependent variable. Here, however, we use EN regularization for embedded feature selection and classification and evaluate the resulting models with cross validation, and repeated resampling.

Traditional statistical methods lack generalisability—the ability to extend findings from a subset population to the population at large. However, the EN regression explicitly maximises the generalisability of the prediction model. By using cross-validation, EN is optimised to perform well on data that was never seen during training. The regularisation and embedded feature selection ensure that the model complexity as well as information loss is minimised. Another approach to feature selection would be to use variables from traditional statistical analysis that have strong associations with HIV in a classification algorithm – this corresponds to filter-based feature selection. However, as summarised in Chapter 3, filters are not optimised for the classifier used, and do not consider interactions between variables. Although we did not explicitly compare filter-based methods to the embedded feature selection of the EN, we expect that classification performance would be poorer because of dependencies between neuroimaging features and increased model complexity.

For model evaluation, we can compare cross-validated performance or examine fit of the model to the data using an information criterion, such as the AIC. It has been shown that for large datasets minimizing AIC is equivalent to minimizing the CV error (Peng et al., 2013; Stone, 1977). We used both methods to evaluate our models in terms of predictive ability and fit to the data. It should, however, be noted that our model parameters were optimised to minimize cross-validation error rather than the AIC. Interestingly, although the model with the lowest CV error was the nested sMRI volumes + DTI + ^{1}H-MRS bg model (AUC 0.82), the penalization of additional parameters in the nested model resulted in the non-nested model having lower AIC. Similarly, the simpler non-nested model (sMRI volumes + DTI only) had a lower AIC than the more complex one (sMRI volumes + DTI + ^{1}H-MRS bg). Since our data set was relatively small, our training and tests were also necessarily small, which makes cross-validation results less reliable (Varoquaux, 2018).

It would be interesting to test whether the classification performance obtained in this study would be similar if the models were used to predict HIV status in children from other HIV cohorts. If our data is similar to that in other HIV+ cohorts, we would expect our results to hold using the model that minimizes CV error (maximum AUC). However, if the distribution of our data is different to the actual data distribution of other cohorts, the models with fewer parameters (and lower AIC) might perform better.

6.4.2. On confounding

Two categories of empirical evidence for confounding exist: internal and external. Statistical evidence from the data at hand is internal evidence, and external evidence refers to evidence from prior research. Strong external evidence overrides internal evidence for multivariable models that focus on generalisability (Brookhart et al., 2010; McNamee, 2003). Generalisability of the predictive models will also be affected by external confounding variables such as nutritional intake and socioeconomic factors that may be difficult to control for. Confounding in this dissertation is limited to measured variables. Other potential confounders are unknown, for example, medical compliance. Based on previous findings (Fish et al., 2017; Lyall et al., 2015; Seider et al., 2016; Sutcliffe, 2010; van Dijk et al., 2011) we chose to control for age at scan, sex, and total intracranial volume by including them as non-penalised variables in the models. The remaining coefficients are thus adjusted for these variables.

In predictive models, there are two ways of dealing with confounders that are known to affect imaging measures: they can be included as predictors in the model i.e. estimated jointly with neuroimaging data (Linn et al., 2016) or the effects of possible confounders may be regressed out of the neuroimaging data before the model is constructed (e.g. Adeli et al., 2018), that is to say estimated marginally with data. If non-imaging variables are associated with the image in a disease-dependent way, estimating disease patterns jointly with non-imaging variables could reveal the relative importance of neuroimaging versus demographic and clinical features in the classification (Linn et al. 2016). However, this can be a problem when the distribution of confounders in the training sample does not match the distribution in the population of interest (Linn et al., 2016; Rao et al., 2017).

Although here we included confounders in our logistic regression models, this is possibly not optimal for small sample sizes in which there may be a bias in the training sample: a relationship between a confounder and the outcome variable that is not representative of the wider population. In our data, there is no difference in age distribution between HIV+ and control children. However, the are 6% more females in the HIV+ group. For comparison, it would be worth assessing how estimating cART-treated HIV patterns marginally across non-imaging variables such as age, sex, and birth weight with elastic net regularisation influences

classification and feature selection. Linn et al. (2016) found that it does not have an impact on the maximum margin hyperplane of support vector machines (SVMs) in high-dimensional problems due to the curse of dimensionality.

6.4.3. On ROI vs voxel-based analysis of DTI

As introduced in Chapter 2, these two methods measure different aspects of WM development in childhood. ROI-based method has a few drawbacks including user-bias in defining meaningful ROIs (Ressel et al., 2018; Snook et al., 2007). In addition, averaged tensor indices in ROI analysis include voxels that may not have any signal, and the averaging may miss true differences if the signal variations are not uniform over the entire ROI. This may explain why the elastic net selects an increase in RD and MD but not the corresponding increase in FA.

Voxel-based group analysis may perform better as an alternative. However, voxel-based methods were avoided in this dissertation because they would exacerbate the dimensionality problem that feature selection is trying to solve (Chapter 3). The more voxels forming the feature space, the more likely the prediction models will overfit to noise in the sample, thereby making the DTI signatures less generalisable. We recommend future studies to use either smaller ROIs or first perform an embedded technique in DTI, for example (Tang et al., 2017), then only include the selected subset voxels in a multimodal analysis.

6.4.4. On segmentation

Investigation of volumes was limited to structures automatically segmented by FreeSurfer. We found automatically segmented structure that are not usually considered in HIV studies to be relevant in our predictive models —namely, brainstem, right vessel, CSF, and ventral diencephalon. These structures should be examined in further neuroimaging studies or correlation studies with neurocognitive tests. Vessels entering the basal ganglia, ventricular system, flow of blood and CSF should be investigated with vascular imaging using MRI or PET; rs-fMRI and auditory brainstem responses to test for hearing impairment and contribution of the brainstem in hearing loss, and ocular examination to examine the implications of reduction in optic chiasm. Furthermore, future studies could explore manual segmentation of

smaller structures such as the hypothalamus and/or release of hormones from the neuroendocrine system.

6.4.5. On spectroscopy

Insofar as any atom with nuclear non-zero spin (^{13}C, ^{21}P, ^{23}Na) can be targeted in theory, spectroscopic measures investigated in this dissertation are in terms of proton (^{1}H nuclei) signals in brain tissue. Furthermore, data is acquired via single voxel spectroscopy (SVS) instead of chemical shift imaging (CSI), which offers a trade-off between multivoxel acquisition and lowered spectral resolution. Future studies should investigate the concentration of metabolites in pwm including lipids and macromolecules. In studies of heterogeneous structures such as the basal ganglia, it might be more informative to obtain higher resolution metabolite maps using CSI, in order to better separate substructures such as globus pallidum and thalamus.

6.4.6. On longitudinal analysis

In our study, data was analysed cross-sectionally at age 7, which prevented us from exploring longitudinal changes due to cART-treated HIV. Since there is data for these children at ages 5, 7, 9, and 11 years, we recommend a future longitudinal multimodal neuroimaging study. Lifelong exposure to HIV and cART in this population warrants a longitudinal assessment of these signatures, their changes over time, relationship to neuroinflammation and vascular dysfunction, and whether they can be useful in predicting neurocognitive outcomes.

6.4.7. On classification

Elastic net regularisation has limitations: First, it is computationally expensive because there are 2 parameters that need to be tuned in multiple CV loops. Second, feature selection becomes unstable when more features are selected. Future studies should implement stability selection with classification by subsampling with the selection algorithm to control for errors in false positives (Meinshausen & Bühlmann, 2010). Embedded feature selection with EN regularised regression is a simple form of supervised learning. It has been shown to

be reduceable to SVM (Zhou et al., 2015), but there are more powerful techniques that could be explored for this application.

The closest method to the elastic net is chained regularisation that Adeli and colleagues (2018) implemented to find brain patterns specific to HIV infection. The authors did a formal comparison and found that it outperforms sparse SVM, EN + SVM, and sequential feature selection + SVM. Otherwise, artificial neural networks have input, output and hidden layers of training and testing to produce good prediction models (Wang et al., 2019; James et al., 2018; Morris et al., 2018), which would be interesting to explore in future work.

6.5. Summary and conclusion

The multimodal signature of cART-treated perinatally-acquired HIV in 7-year old children is smaller right and left globus pallidus, bilateral amygdala, right vessel of the basal ganglia, optic chiasm, brainstem, right lateral ventricle, left ventral diencephalon, and right cerebellum white matter; enlargement of the ventricular system; increased RD and MD in uncinate fasciculus, MD and RD in left corticospinal tract, and MD and RD in cingulum, and AD in right superior temporal longitudinal fasciculus. This multimodal signature has a classification performance of 83% sensitivity, 80% specificity, 81% accuracy, and 78% AUC, which is better than signature of volumes and DTI separately. Measures of cortical thickness, gyrification, and spectroscopy in basal ganglia, peritrigonal white matter, or midfrontal gray matter, are not useful in the classification of cART-treated paediatric HIV.

While the ability to predict HIV-status based on neuroimaging is not clinically advantageous particularly in resource-limited settings, demonstration of its feasibility could help identify neuroimaging signatures that are common to treated HIV infection and offer a separate yet complementary dimension to the clinical diagnosis. Building on this knowledge, future research might make it possible to predict cognitive or clinical outcomes, and to better tailor cART treatment to the individual.

WORKS CITED

Ackermann, C., Andronikou, S., Saleh, M. G., Laughton, B., Alhamud, A. A., Van Der Kouwe, A., … Meintjes, E. M. (2016). Early antiretroviral therapy in HIV-infected children is associated with diffuse white matter structural abnormality and corpus callosum sparing. In *American Journal of Neuroradiology*. https://doi.org/10.3174/ajnr.A4921

Ackermann, Christelle, van Toorn, R., & Andronikou, S. (2019). Human immunodeficiency virus-related cerebral white matter disease in children. *Pediatric Radiology*. https://doi.org/10.1007/s00247-018-4310-x

Adeli, E., Kwon, D., Zhao, Q., Pfefferbaum, A., Zahr, N. M., Sullivan, E. V., & Pohl, K. M. (2018). Chained regularization for identifying brain patterns specific to HIV infection. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2018.08.022

Akaike, H. (1973). Maximum likelihood identification of gaussian autoregressive moving average models. *Biometrika*. https://doi.org/10.1093/biomet/60.2.255

Alhamud, A., Tisdall, M. D., Hess, A. T., Hasan, K. M., Meintjes, E. M., & Van Der Kouwe, A. J. W. (2012). Volumetric navigators for real-time motion correction in diffusion tensor imaging. *Magnetic Resonance in Medicine*. https://doi.org/10.1002/mrm.23314

Ances, B. M., Ortega, M., Vaida, F., Heaps, J., & Paul, R. (2012). Independent effects of HIV, aging, and HAART on brain volumetric measures. *Journal of Acquired Immune Deficiency Syndromes*. https://doi.org/10.1097/QAI.0b013e318249db17

Anderson, K. P., & Burnham, D. A. (2002). *Model Selection and Multi-Model Inference : A Practical Information-Theoretic Approach (2nd Edition)*. *Ecological Modelling*. https://doi.org/10.1016/j.ecolmodel.2003.11.004

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry - The methods. *NeuroImage*. https://doi.org/10.1006/nimg.2000.0582

Ather, S., Proudlock, F. A., Welton, T., Morgan, P. S., Sheth, V., Gottlob, I., & Dineen, R. A. (2019). Aberrant visual pathway development in albinism: From retina to cortex. *Human Brain Mapping*. https://doi.org/10.1002/hbm.24411

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*. https://doi.org/10.1002/mpr.329

Baeken, C., Marinazzo, D., Van Schuerbeek, P., Wu, G. R., De Mey, J., Luypaert, R., & De Raedt, R. (2014). Left and right amygdala - Mediofrontal cortical functional connectivity is differentially modulated by harm avoidance. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0095740

Ballabh, P., Braun, A., & Nedergaard, M. (2004a). The blood-brain barrier: An overview: Structure, regulation, and clinical implications. *Neurobiology of Disease*. https://doi.org/10.1016/j.nbd.2003.12.016

Ballabh, P., Braun, A., & Nedergaard, M. (2004b). The blood-brain barrier: An overview: Structure, regulation, and clinical implications. *Neurobiology of Disease*. https://doi.org/10.1016/j.nbd.2003.12.016

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., … Reiss, A. L. (2005). White matter development during childhood and adolescence: A cross-sectional diffusion tensor imaging study. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhi062

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*. https://doi.org/10.1016/S0006-3495(94)80775-1

Beaulieu, C., Plewes, C., Paulson, L. A., Roy, D., Snook, L., Concha, L., & Phillips, L. (2005). Imaging brain connectivity in children with diverse reading ability. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2004.12.053

Belman, A. L., Lantos, G., Horoupian, D., Novick, B. E., Ultmann, M. H., Dickson, D. W., & Rubinstein, A. (1986). AIDS: Calcification of the basal ganglia in infants and children. *Neurology*. https://doi.org/10.1212/wnl.36.9.1192

Bertholdo, D., Watcharakorn, A., & Castillo, M. (2013). Brain Proton Magnetic Resonance Spectroscopy: Introduction and Overview. *Neuroimaging Clinics of North America*. https://doi.org/10.1016/j.nic.2012.10.002

Bigler, E. D., Abildskov, T. J., Eggleston, B., Taylor, B. A., Tate, D. F., Petrie, J. A., … Wilde, E. A. (2019). Structural neuroimaging in mild traumatic brain injury: A chronic effects of neurotrauma consortium study. *International Journal of Methods in Psychiatric Research*. https://doi.org/10.1002/mpr.1781

Blakemore, S. J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2011.11.080

Blokhuis, C., Peeters, C. F. W., Cohen, S., Scherpbier, H. J., Kuijpers, T. W., Reiss, P., … Pajkrt,

D. (2019). Systemic and intrathecal immune activation in association with cerebral and cognitive outcomes in paediatric HIV. *Scientific Reports*. https://doi.org/10.1038/s41598-019-44198-z

Blokhuis, Charlotte, Mutsaerts, H. J. M. M., Cohen, S., Scherpbier, H. J., Caan, M. W. A., Majoie, C. B. L. M., … Pajkrt, D. (2017). Higher subcortical and white matter cerebral blood flow in perinatally HIV-infected children. *Medicine (United States)*. https://doi.org/10.1097/MD.0000000000005891

Blüml, S., & Panigrahy, A. (2013). *MR spectroscopy of pediatric brain disorders*. *MR Spectroscopy of Pediatric Brain Disorders*. https://doi.org/10.1007/978-1-4419-5864-8

Bottomley, P. A. (1984). NMR in medicine. *Computerized Radiology*. https://doi.org/10.1016/0730-4862(84)90065-9

Bowman, F. D. B., Drake, D. F., & Huddleston, D. E. (2016). Multimodal imaging signatures of Parkinson's disease. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2016.00131

Brookhart, M. A., Stürmer, T., Glynn, R. J., Rassen, J., & Schneeweiss, S. (2010). Confounding control in healthcare database research: Challenges and potential approaches. *Medical Care*. https://doi.org/10.1097/MLR.0b013e3181dbebe3

Brouwer, R. M., Mandl, R. C. W., Peper, J. S., van Baal, G. C. M., Kahn, R. S., Boomsma, D. I., & Hulshoff Pol, H. E. (2010). Heritability of DTI and MTR in nine-year-old children. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2010.03.017

Bruck, I., Tahan, T. T., Rodrigues Da Cruz, C., Martins, L. T. F., Antoniuk, S. A., Rodrigues, M., … De Bruyn, L. R. (2001). Developmental milestones of vertically HIV infected and seroreverters children: Follow up of 83 children. *Arquivos de Neuro-Psiquiatria*. https://doi.org/10.1590/s0004-282x2001000500007

Buckner, C. M., Luers, A. J., Calderon, T. M., Eugenin, E. A., & Berman, J. W. (2006). Neuroimmunity and the blood-brain barrier: Molecular regulation of leukocyte transmigration and viral entry into the nervous system with a focus on neuroAIDS. *Journal of Neuroimmune Pharmacology*. https://doi.org/10.1007/s11481-006-9017-3

Bunea, F., She, Y., Ombao, H., Gongvatana, A., Devlin, K., & Cohen, R. (2011). Penalized least squares regression methods and applications to neuroimaging. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2010.12.028

Cabral, C., Kambeitz-Ilankovic, L., Kambeitz, J., Calhoun, V. D., Dwyer, D. B., Von Saldern,

S., … Koutsouleris, N. (2016). Classifying Schizophrenia Using Multimodal Multivariate Pattern Recognition Analysis: Evaluating the Impact of Individual Clinical Profiles on the Neurodiagnostic Performance. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbw053

Caniglia, E. C., Cain, L. E., Justice, A., Tate, J., Logan, R., Sabin, C., … Hernan, M. A. (2014). Antiretroviral penetration into the CNS and incidence of AIDS-defining neurologic conditions. *Neurology*, *83*(2), 134–141. https://doi.org/10.1212/WNL.0000000000000564

Chantry, C. J., Frederick, M. M., Meyer, W. A., Handelsman, E., Rich, K., Paul, M. E., … Moye, J. (2007). Endocrine abnormalities and impaired growth in human immunodeficiency virus-infected children. *Pediatric Infectious Disease Journal*. https://doi.org/10.1097/01.inf.0000247131.76584.af

Chao, C. K., Czechowicz, J. A., Messner, A. H., Alarcoń, J., Roca, L. K., Larragán Rodriguez, M. M., … Zunt, J. R. (2012). High prevalence of hearing impairment in HIV-infected Peruvian children. *Otolaryngology - Head and Neck Surgery*. https://doi.org/10.1177/0194599811429271

Clifford, D. B., & Ances, B. M. (2013). HIV-associated neurocognitive disorder. *The Lancet Infectious Diseases*. https://doi.org/10.1016/S1473-3099(13)70269-X

Colby, J. B., Rudie, J. D., Brown, J. A., Douglas, P. K., Cohen, M. S., & Shehzad, Z. (2012). Insights into multimodal imaging classification of ADHD. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/fnsys.2012.00059

Cotton, M. F., Violari, A., Otwombe, K., Panchia, R., Dobbels, E., Rabie, H., … Babiker, A. G. (2013). Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: Results from the children with HIV early antiretroviral (CHER) randomised trial. *The Lancet*. https://doi.org/10.1016/S0140-6736(13)61409-9

Davatzikos, C. (2019). Machine learning in neuroimaging: Progress and challenges. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2018.10.003

De Girolami, U., Sharer, L. R., Gabuzda, D., & Sotrel, A. (2004). Neuropathology of AIDS. In *AIDS and Other Manifestations of HIV Infection*. https://doi.org/10.1016/B978-012764051-8/50030-5

Dennehy, P. J., Warman, R., Flynn, J. T., Scott, G. B., & Mastrucci, M. T. (1989). Ocular Manifestations in Pediatric Patients with Acquired Immunodeficiency Syndrome.

Archives of Ophthalmology. https://doi.org/10.1001/archopht.1989.01070020040025

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dickson, D. W., Belman, A. L., Park, Y. D., Wiley, C., Horoupian, D. S., Llena, J., … Cho, S. (1989). Central nervous system pathology in pediatric AIDS: An autopsy study. *APMIS, Supplement*.

Donzuso, G., Mostile, G., Nicoletti, A., & Zappia, M. (2019). Basal ganglia calcifications (Fahr﷽s syndrome): related conditions and clinical features. *Neurological Sciences*, 1–13.

Drost, D. J., Riddle, W. R., & Clarke, G. D. (2002). Proton magnetic resonance spectroscopy in the brain: Report of AAPM MR Task Group #9. *Medical Physics*. https://doi.org/10.1118/1.1501822

Ellis, R. J., Rosario, D., Clifford, D. B., McArthur, J. C., Simpson, D., Alexander, T., … Grant, I. (2010). Continued high prevalence and adverse clinical impact of human immunodeficiency virus-associated sensory neuropathy in the era of combination antiretroviral therapy: The CHARTER study. *Archives of Neurology*. https://doi.org/10.1001/archneurol.2010.76

Emsley, R., Asmal, L., Chiliza, B., du Plessis, S., Carr, J., Kidd, M., … Kahn, R. S. (2015). Changes in brain regions associated with food-intake regulation, body mass and metabolic profiles during acute antipsychotic treatment in first-episode schizophrenia. *Psychiatry Research - Neuroimaging*. https://doi.org/10.1016/j.pscychresns.2015.06.014

Fang, Y. (2011). Asymptotic Equivalence between Cross-Validations and Akaike Information Criteria in Mixed-Effects Models. *Journal of Data Science*.

Feldman, H. M., Yeatman, J. D., Lee, E. S., Barde, L. H. F., & Gaman-Bean, S. (2010). Diffusion tensor imaging: A review for pediatric researchers and clinicians. *Journal of Developmental and Behavioral Pediatrics*. https://doi.org/10.1097/DBP.0b013e3181dcaa8b

Filipek, P. A., Richelme, C., Kennedy, D. N., & Caviness, V. S. (1994). The young adult human brain: An MRI-based morphometric analysis. *Cerebral Cortex*.

https://doi.org/10.1093/cercor/4.4.344

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*. https://doi.org/10.1016/S0896-6273(02)00569-X

Fish, A. M., Cachia, A., Fischer, C., Mankiw, C., Reardon, P. K., Clasen, L. S., … Raznahan, A. (2017). Influences of Brain Size, Sex, and Sex Chromosome Complement on the Architecture of Human Cortical Folding. *Cerebral Cortex (New York, N.Y. : 1991)*. https://doi.org/10.1093/cercor/bhw323

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v033.i01

Frye, R. E., Hasan, K., Malmberg, B., Desouza, L., Swank, P., SMITH, K., & Landry, S. (2010). Superior longitudinal fasciculus and cognitive dysfunction in adolescents born preterm and at term. *Developmental Medicine and Child Neurology*. https://doi.org/10.1111/j.1469-8749.2010.03633.x

Garvie, P. A., Zeldow, B., Malee, K., Nichols, S. L., Smith, R. A., Wilkins, M. L., & Williams, P. L. (2014). Discordance of cognitive and academic achievement outcomes in youth with perinatal HIV exposure. *Pediatric Infectious Disease Journal*. https://doi.org/10.1097/INF.0000000000000314

George, R., Andronikou, S., Du Plessis, J., Du Plessis, A. M., Van Toorn, R., & Maydell, A. (2009). Central nervous system manifestations of HIV infection in children. In *Pediatric Radiology*. https://doi.org/10.1007/s00247-009-1170-4

Gillies, M. J., Hyam, J. A., Weiss, A. R., Antoniades, C. A., Bogacz, R., Fitzgerald, J. J., … Green, A. L. (2017). The Cognitive Role of the Globus Pallidus interna; Insights from Disease States. *Experimental Brain Research*. https://doi.org/10.1007/s00221-017-4905-8

Glick, N. (1978). Additive estimators for probabilities of correct classification. *Pattern Recognition*. https://doi.org/10.1016/0031-3203(78)90029-8

Gorgolewski, K., Esteban, O., Schaefer, G., Wandell, B., & Poldrack, R. (2017). OpenNeuro﷽a free online platform for sharing and analysis of neuroimaging data. *Organization for Human Brain Mapping. Vancouver, Canada*, *1677*.

Grus, J. (2016). Chapter 10. Working with Data. *Data Science from Scratch : First Principles with Python*.

Guyon, I., & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research (JMLR)*. https://doi.org/10.1016/j.aca.2011.07.027

Haase, A., Frahm, J., Hanicke, W., & Matthaei, D. (1985). 1H NMR chemical shift selective (CHESS) imaging. *Physics in Medicine and Biology*. https://doi.org/10.1088/0031-9155/30/4/008

Hahn, E. L. (1950). Spin echoes. *Physical Review*. https://doi.org/10.1103/PhysRev.80.580

Herrera, E., Agudo-Barriuso, M., & Murcia-Belmonte, V. (2019). Cranial Pair II: The Optic Nerves. *The Anatomical Record*, *302*(3), 428–445. https://doi.org/10.1002/ar.23922

Hoare, J., Fouche, J. P., Phillips, N., Joska, J. A., Donald, K. A., Thomas, K., & Stein, D. J. (2015). Clinical associations of white matter damage in cART-treated HIV-positive children in South Africa. *Journal of NeuroVirology*. https://doi.org/10.1007/s13365-014-0311-1

Hoare, J., Heany, S. J., Fouche, J. P., Phillips, N., Joska, J. A., Myer, L., … Stein, D. J. (2019). Initiation of antiretroviral therapy after the critical neuronal developmental period of the second postnatal year affects white matter microstructure in adolescents living with HIV. *Journal of NeuroVirology*. https://doi.org/10.1007/s13365-018-0712-7

Holmes, M. J., Robertson, F. C., Little, F., Randall, S. R., Cotton, M. F., Van Der Kouwe, A. J. W., … Meintjes, E. M. (2017). Longitudinal increases of brain metabolite levels in 5-10 year old children. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0180973

Hrapcak, S., Kuper, H., Bartlett, P., Devendra, A., Makawa, A., Kim, M., … Ahmed, S. (2016). Hearing loss in HIV-infected children in Lilongwe, Malawi. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0161421

Izquierdo-Useros, N., Naranjo-Gómez, M., Erkizia, I., Puertas, M. C., Borràs, F. E., Blanco, J., & Martinez-Picado, J. (2010). HIV and mature dendritic cells: Trojan exosomes riding the Trojan horse? *PLoS Pathogens*. https://doi.org/10.1371/journal.ppat.1000740

Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., … Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*. https://doi.org/10.1002/jmri.21049

Jacqueline, H., Jenny, W. T., Jean-Paul, F., Bruce, S., Robert, P., Kevin, T., … John, J. (2012). A diffusion tensor imaging and neuropsychological study of prospective memory impairment in South African HIV positive individuals. *Metabolic Brain Disease*. https://doi.org/10.1007/s11011-012-9311-0

James, T. O., Gulumbe, S. U., & Danbaba, A. (2018). Resilient Back-Propagation Algorithm in the Prediction of Mother to Child Transmission of HIV. *OALib*. https://doi.org/10.4236/oalib.1104538

Jankiewicz, M., Holmes, M. J., Taylor, P. A., Cotton, M. F., Laughton, B., van der Kouwe, A. J. W., & Meintjes, E. M. (2017). White Matter Abnormalities in Children with HIV Infection and Exposure. *Frontiers in Neuroanatomy*, *11*. https://doi.org/10.3389/fnana.2017.00088

Janssen, R. J., Mourão-Miranda, J., & Schnack, H. G. (2018). Making Individual Prognoses in Psychiatry Using Neuroimaging and Machine Learning. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2018.04.004

Jie, B., Zhang, D., Gao, W., Wang, Q., Wee, C. Y., & Shen, D. (2014). Integration of network topological and connectivity properties for neuroimaging classification. *IEEE Transactions on Biomedical Engineering*. https://doi.org/10.1109/TBME.2013.2284195

Johnson, K. A., Fox, N. C., Sperling, R. A., & Klunk, W. E. (2012). Brain imaging in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*. https://doi.org/10.1101/cshperspect.a006213

Jollans, L., Boyle, R., Artiges, E., Banaschewski, T., Desrivières, S., Grigis, A., … Whelan, R. (2019). Quantifying performance of machine learning methods for neuroimaging data. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2019.05.082

Kallianpur, K. J., Gerschenson, M., Mitchell, B. I., LiButti, D. E., Umaki, T. M., Ndhlovu, L. C., … Shikuma, C. M. (2016). Oxidative mitochondrial DNA damage in peripheral blood mononuclear cells is associated with reduced volumes of hippocampus and subcortical gray matter in chronically HIV-infected patients. *Mitochondrion*. https://doi.org/10.1016/j.mito.2016.02.006

Kambeitz, J., Cabral, C., Sacchet, M. D., Gotlib, I. H., Zahn, R., Serpa, M. H., … Koutsouleris, N. (2017). Detecting Neuroimaging Biomarkers for Depression: A Meta-analysis of Multivariate Pattern Recognition Studies. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2016.10.028

Kauffman, W. M., Sivit, C. J., Fitz, C. R., Rakusan, T. A., Herzog, K., & Chandra, R. S. (1992). CT and MR evaluation of intracranial involvement in pediatric HIV infection: A clinical-imaging correlation. *American Journal of Neuroradiology*.

Kieck, J. R., & Andronikou, S. (2004). Usefulness of neuro-imaging for the diagnosis of HIV

encephalopathy in children. *South African Medical Journal*. https://doi.org/10.7196/SAMJ.2655

Kim, D., Kang, P., Kim, J., Kim, C. Y., Lee, J. H., Suh, S., & Lee, M. S. (2019). Machine learning classification of first-onset drug-naive MDD using structural MRI. *IEEE Access*. https://doi.org/10.1109/ACCESS.2019.2949128

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *Appears in the International Joint Conference on Articial Intelligence (IJCAI)*. https://doi.org/10.1067/mod.2000.109031

Kotov, S. (2017). MRI Morphometry of the Brain and Neurological Diseases. *New Insights into Morphometry Studies*, 47.

Kovalevich, J., & Langford, D. (2012). Neuronal toxicity in HIV CNS disease. *Future Virology*. https://doi.org/10.2217/fvl.12.57

Kreis, R., Ernst, T., & Ross, B. D. (1993). Absolute Quantitation of Water and Metabolites in the Human Brain. II. Metabolite Concentrations. *Journal of Magnetic Resonance, Series B*. https://doi.org/10.1006/jmrb.1993.1056

Kurth, F., Luders, E., & Gaser, C. (2015). Voxel-Based Morphometry. In *Brain Mapping: An Encyclopedic Reference*. https://doi.org/10.1016/B978-0-12-397025-1.00304-3

Kwee, R. M., & Kwee, T. C. (2007). Virchow-Robin spaces at MR imaging. *Radiographics*. https://doi.org/10.1148/rg.274065722

Laughton, B., Cornell, M., Grove, D., Kidd, M., Springer, P. E., Dobbels, E., … Cotton, M. F. (2012). Early antiretroviral therapy improves neurodevelopmental outcomes in infants. *AIDS*. https://doi.org/10.1097/QAD.0b013e328355d0ce

Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5302-10.2011

Lentz, M. R., Peterson, K. L., Ibrahim, W. G., Lee, D. E., Sarlls, J., Lizak, M. J., … Hammoud, D. A. (2014). Diffusion tensor and volumetric magnetic resonance measures as biomarkers of brain damage in a small animal model of HIV. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0105752

Letendre, S., Marquie-Beck, J., Capparelli, E., Best, B., Clifford, D., Collier, A. C., … Ellis, R. J. (2008). Validation of the CNS penetration-effectiveness rank for quantifying antiretroviral penetration into the central nervous system. *Archives of Neurology*.

https://doi.org/10.1001/archneurol.2007.31

Lewis-de los Angeles, C. P., Alpert, K. I., Williams, P. L., Malee, K., Huo, Y., Csernansky, J. G., … Wang, L. (2016). Deformed Subcortical Structures Are Related to Past HIV Disease Severity in Youth With Perinatally Acquired HIV Infection. *Journal of the Pediatric Infectious Diseases Society*. https://doi.org/10.1093/jpids/piw051

Li, J., Wu, G., Wen, Z., Zhang, J., Lei, H., Gui, X., & Lin, F. (2015). White matter development is potentially influenced in adolescents with vertically transmitted HIV infections: A tract-based spatial statistics study. *American Journal of Neuroradiology*. https://doi.org/10.3174/ajnr.A4417

Linn, K. A., Gaonkar, B., Doshi, J., Davatzikos, C., & Shinohara, R. T. (2016). Addressing Confounding in Predictive Models with an Application to Neuroimaging. *International Journal of Biostatistics*. https://doi.org/10.1515/ijb-2015-0030

Loomba-Albrecht, L. A., Bregman, T., & Chantry, C. J. (2014). Endocrinopathies in children infected with human immunodeficiency virus. *Endocrinology and Metabolism Clinics of North America*. https://doi.org/10.1016/j.ecl.2014.06.001

Lyall, A. E., Shi, F., Geng, X., Woolson, S., Li, G., Wang, L., … Gilmore, J. H. (2015). Dynamic Development of Regional Cortical Thickness and Surface Area in Early Childhood. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhu027

Madhi, S. A., Adrian, P., Cotton, M. F., McIntyre, J. A., Jean-Philippe, P., Meadows, S., … Violari, A. (2010). Effect of HIV Infection Status and Anti-Retroviral Treatment on Quantitative and Qualitative Antibody Responses to Pneumococcal Conjugate Vaccine in Infants. *The Journal of Infectious Diseases*. https://doi.org/10.1086/653704

Makris, N., Angelone, L., Tulloch, S., Sorg, S., Kaiser, J., Kennedy, D., & Bonmassar, G. (2008). MRI-based anatomical model of the human head for specific absorption rate mapping. *Medical and Biological Engineering and Computing*. https://doi.org/10.1007/s11517-008-0414-z

Makris, N., Swaab, D. F., van der Kouwe, A., Abbs, B., Boriel, D., Handa, R. J., … Goldstein, J. M. (2013). Volumetric parcellation methodology of the human hypothalamus in neuroimaging: Normative data and sex differences. *NeuroImage*, *69*, 1–10. https://doi.org/10.1016/j.neuroimage.2012.12.008

Mariam, A. G., & Assefa, G. (2012). Clinical and neuroimaging profile of HIV-I encephalopathy in infancy and childhood in a Sub-saharan african country. *Ethiopian*

Medical Journal.

Mazziotta, J. C., Woods, R., Iacoboni, M., Sicotte, N., Yaden, K., Tran, M., … Toga, A. W. (2009). The myth of the normal, average human brain-The ICBM experience: (1) Subject screening and eligibility. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2008.07.062

Mbugua, K. K., Holmes, M. J., Cotton, M. F., Ratai, E. M., Little, F., Hess, A. T., … Meintjes, E. M. (2016). HIV-associated CD4+/CD8+ depletion in infancy is associated with neurometabolic reductions in the basal ganglia at age 5 years despite early antiretroviral therapy. *AIDS*. https://doi.org/10.1097/QAD.0000000000001082

McNamee, R. (2003). Confounding and confounders. *Occupational and Environmental Medicine*, *60*(3), 227–234. https://doi.org/10.1136/oem.60.3.227

Meinshausen, N., & Bühlmann, P. (2010). Stability selection. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. https://doi.org/10.1111/j.1467-9868.2010.00740.x

Mitchell, W. (2001). Neurological and developmental effects of HIV and AIDS in children and adolescents. *Mental Retardation and Developmental Disabilities Research Reviews*. https://doi.org/10.1002/mrdd.1029

Moore, E. S., Ward, R. E., Wetherill, L. F., Rogers, J. L., Autti-Rämö, I., Fagerlund, Å., … Mattson, S. N. (2007). Unique facial features distinguish fetal alcohol syndrome patients and controls in diverse ethnic populations. *Alcoholism: Clinical and Experimental Research*. https://doi.org/10.1111/j.1530-0277.2007.00472.x

Mori, S. (2007). Human White Matter Atlas. *American Journal of Psychiatry*. https://doi.org/10.1176/appi.ajp.164.7.1005

Morris, P., Dasilva, Y., Clark, E., Hahn, W. E., & Barenholtz, E. (2018). Convolutional Neural Networks for Predicting Molecular Binding Affinity to HIV-1 Proteins. In *ACM-BCB 2018 - Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. https://doi.org/10.1145/3233547.3233596

Mwangi, B., Tian, T. S., & Soares, J. C. (2014). A Review of Feature Reduction Techniques in Neuroimaging. *Neuroinformatics*, *12*(2), 229–244. https://doi.org/10.1007/s12021-013-9204-3

Nichols, S. L., Brummel, S. S., Smith, R. A., Garvie, P. A., Hunter, S. J., Malee, K. M., … Willen,

E. (2015). Executive Functioning in Children and Adolescents With Perinatal HIV Infection. *Pediatric Infectious Disease Journal*. https://doi.org/10.1097/INF.0000000000000809

Nichols, S. L., Chernoff, M. C., Malee, K. M., Sirois, P. A., Woods, S. P., Williams, P. L., … Kammerer, B. (2016). Executive functioning in children and adolescents with perinatal HIV infection and perinatal HIV exposure. *Journal of the Pediatric Infectious Diseases Society*. https://doi.org/10.1093/jpids/piw049

Nwosu, E. C., Robertson, F. C., Holmes, M. J., Cotton, M. F., Dobbels, E., Little, F., … Meintjes, E. M. (2018). Altered brain morphometry in 7-year old HIV-infected children on early ART. *Metabolic Brain Disease*. https://doi.org/10.1007/s11011-017-0162-6

Ogutu, J. O., Schulz-Streeck, T., & Piepho, H.-P. (2012). Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. *BMC Proceedings*. https://doi.org/10.1186/1753-6561-6-S2-S10

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., … Mori, S. (2008). Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2008.07.009

Padhani, D. H., Manji, K. P., & Mtanda, A. T. (2000). Ocular manifestations in children with HIV infection in Dar es Salaam, Tanzania. *Journal of Tropical Pediatrics*. https://doi.org/10.1093/tropej/46.3.145

Pandey, A., John, B. M., & Gupta, R. (2018). Endocrinological Abnormalities and Growth Impairment in Human Immunodeficiency Virus-Infected Children. *Journal of Pediatric Infectious Diseases*. https://doi.org/10.1055/s-0037-1617414

Paul, R., Prasitsuebsai, W., Jahanshad, N., Puthanakit, T., Thompson, P., Aurpibul, L., … Ananworanich, J. (2018). Structural Neuroimaging and Neuropsychologic Signatures in Children With Vertically Acquired HIV. *The Pediatric Infectious Disease Journal*, *37*(7), 662–668. https://doi.org/10.1097/INF.0000000000001852

Paulus, M. P., Kuplicki, R., & Yeh, H. W. (2019). Machine Learning and Brain Imaging: Opportunities and Challenges. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2019.07.007

Pavlakis, S. G., Lu, D., Frank, Y., Bakshi, S., Pahwa, S., Barnett, T. A., … Hyman, R. A. (1995). Magnetic resonance spectroscopy in childhood AIDS encephalopathy. *Pediatric*

Neurology. https://doi.org/10.1016/0887-8994(95)00048-K

Peng, H., Yan, H., & Zhang, W. (2013). The connection between cross-validation and Akaike information criterion in a semiparametric family. *Journal of Nonparametric Statistics*. https://doi.org/10.1080/10485252.2013.767338

Považan, M., Hangel, G., Strasser, B., Gruber, S., Chmelik, M., Trattnig, S., & Bogner, W. (2015). Mapping of brain macromolecules and their use for spectral processing of 1H-MRSI data with an ultra-short acquisition delay at 7T. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2015.07.042

Prado, P. T. C., Escorsi-Rosset, S., Cervi, M. C., & Santos, A. C. (2011). Image evaluation of HIV encephalopathy: A multimodal approach using quantitative MR techniques. *Neuroradiology*. https://doi.org/10.1007/s00234-011-0869-8

Pratt, R. D., Nichols, S., McKinney, N., Kwok, S., Dankner, W. M., & Spector, S. A. (1996). Virologic markers of human immunodeficiency virus type 1 in cerebrospinal fluid of infected children. *Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/174.2.288

Provencher, S. W. (2001). Automatic quantitation of localized in vivo1H spectra with LCModel. *NMR in Biomedicine*. https://doi.org/10.1002/nbm.698

Puzniak, R. J., Ahmadi, K., Kaufmann, J., Gouws, A., Morland, A. B., Pestilli, F., & Hoffmann, M. B. (2019). Quantifying nerve decussation abnormalities in the optic chiasm. *NeuroImage: Clinical*. https://doi.org/10.1016/j.nicl.2019.102055

R Core Team. (2018). A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*.

Randall, S. R., Warton, C. M. R., Holmes, M. J., Cotton, M. F., Laughton, B., van der Kouwe, A. J. W., & Meintjes, E. M. (2017). Larger subcortical gray matter structures and smaller corpora callosa at age 5 years in HIV infected children on early ART. *Frontiers in Neuroanatomy*. https://doi.org/10.3389/fnana.2017.00095

Rao, A., Monteiro, J. M., & Mourao-Miranda, J. (2017). Predictive modelling using neuroimaging data in the presence of confounds. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2017.01.066

Rees, G. (2004). Statistical parametric mapping. *Practical Neurology*. https://doi.org/10.1111/j.1474-7766.2004.00266.x

Ressel, V., van Hedel, H. J. A., Scheer, I., & O'Gorman Tuura, R. (2018). Comparison of DTI

analysis methods for clinical research: influence of pre-processing and tract selection methods. *European Radiology Experimental*. https://doi.org/10.1186/s41747-018-0066-1

Ritter, K., Schumacher, J., Weygandt, M., Buchert, R., Allefeld, C., & Haynes, J.-D. (2015). Multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. https://doi.org/10.1016/j.dadm.2015.01.006

Robertson, F. C., Holmes, M. J., Cotton, M. F., Dobbels, E., Little, F., Laughton, B., … Meintjes, E. M. (2018). Perinatal HIV Infection or Exposure Is Associated With Low N-Acetylaspartate and Glutamate in Basal Ganglia at Age 9 but Not 7 Years. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2018.00145

Rondanelli, M., Caselli, D., Aricò, M., Maccabruni, A., Magnani, B., Bacchella, L., … Minoli, L. (2002). Insulin-like growth factor I (IGF-I) and IGF-binding protein 3 response to growth hormone is impaired in HIV-infected children. *AIDS Research and Human Retroviruses*. https://doi.org/10.1089/088922202753519106

Samboju, V., Philippi, C. L., Chan, P., Cobigo, Y., Fletcher, J. L. K., Robb, M., … Valcour, V. (2018). Structural and functional brain imaging in acute HIV. *NeuroImage: Clinical*. https://doi.org/10.1016/j.nicl.2018.07.024

Saylor, D., Dickens, A. M., Sacktor, N., Haughey, N., Slusher, B., Pletnikov, M., … McArthur, J. C. (2016). HIV-associated neurocognitive disorder - Pathogenesis and prospects for treatment. *Nature Reviews Neurology*. https://doi.org/10.1038/nrneurol.2016.27

Schmithorst, V. J., Wilke, M., Dardzinski, B. J., & Holland, S. K. (2002). Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: A cross-sectional diffusion-tensor MR imaging study. *Radiology*. https://doi.org/10.1148/radiol.2221010626

Schneider, J. F. L., Il'yasov, K. A., Hennig, J., & Martin, E. (2004). Fast quantitative diffusion-tensor imaging of cerebral white matter from the neonatal period to adolescence. *Neuroradiology*. https://doi.org/10.1007/s00234-003-1154-2

Schouten, T. M., Loitfelder, M., de Vos, F., Seiler, S., van der Grond, J., Lechner, A., … Rombouts, S. A. R. B. (2016). Combining anatomical, diffusion, and resting state functional magnetic resonance imaging for individual classification of mild and moderate Alzheimer's disease. *NeuroImage: Clinical*.

https://doi.org/10.1016/j.nicl.2016.01.002

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Büchel, C., … Struve, M. (2010). The IMAGEN study: Reinforcement-related behaviour in normal brain function and psychopathology. *Molecular Psychiatry*. https://doi.org/10.1038/mp.2010.4

Sei, S., Stewart, S. K., Parley, M., Mueller, B. U., Lane, J. R., Robb, M. L., … Pizzo, P. A. (1996). Evaluation of human immunodeficiency virus (HIV) type 1 RNA levels in cerebrospinal fluid and viral resistance to zidovudine in children with HIV encephalopathy. *Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/174.6.1200

Seider, T. R., Gongvatana, A., Woods, A. J., Chen, H., Porges, E. C., Cummings, T., … Cohen, R. A. (2016). Age exacerbates HIV-associated white matter abnormalities. *Journal of NeuroVirology*. https://doi.org/10.1007/s13365-015-0386-3

Shah, A., Gangwani, M. R., Chaudhari, N. S., Glazyrin, A., Bhat, H. K., & Kumar, A. (2016). Neurotoxicity in the Post-HAART Era: Caution for the Antiretroviral Therapeutics. *Neurotoxicity Research*. https://doi.org/10.1007/s12640-016-9646-0

SHARER, L. R., DOWLING, P. C., MICHAELS, J., COOK, S. D., MENONNA, J., BLUMBERGH, B. M., & EPSTEIN, L. G. (1990). Spinal cord disease in children with HIV-1 infection: a combined molecular biological and neuropathological study. *Neuropathology and Applied Neurobiology*. https://doi.org/10.1111/j.1365-2990.1990.tb01266.x

Snook, L., Paulson, L. A., Roy, D., Phillips, L., & Beaulieu, C. (2005). Diffusion tensor imaging of neurodevelopment in children and young adults. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2005.03.016

Snook, L., Plewes, C., & Beaulieu, C. (2007). Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2006.07.021

Sorenson, J. (1990). MRI Physics for Physicians , by Alfred L. Horowitz . *Medical Physics*. https://doi.org/10.1118/1.596576

Spalleta, gianfranco, & Fabrizio, P. (2018). Brain morphometry. In *Neuromethods*. https://doi.org/10.1007/978-1-4939-7647-8_21

Stats SA. (2018). Statistical Release P0441. *Stats SA*.

Steinbrink, F., Evers, S., Buerke, B., Young, P., Arendt, G., Koutsilieri, E., … Husstedt, I. W. (2013). Cognitive impairment in HIV infection is associated with MRI and CSF pattern of neurodegeneration. *European Journal of Neurology*.

https://doi.org/10.1111/ene.12006

Stejskal, E. O., & Tanner, J. E. (1965). Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*. https://doi.org/10.1063/1.1695690

Stone, M. (1977). An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal of the Royal Statistical Society: Series B (Methodological)*. https://doi.org/10.1111/j.2517-6161.1977.tb01603.x

Sui, J., Pearlson, G. D., Du, Y., Yu, Q., Jones, T. R., Chen, J., … Calhoun, V. D. (2015). In Search of Multimodal Neuroimaging Biomarkers of Cognitive Deficits in Schizophrenia. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2015.02.017

Sutcliffe, C. G. (2010). *The care and treatment of HIV-infected children in Sub-Saharan Africa. Dissertation Abstracts International: Section B: The Sciences and Engineering*.

Tang, Z., Liu, Z., Li, R., Yang, X., Cui, X., Wang, S., … Tian, J. (2017). Identifying the white matter impairments among ART-naïve HIV patients: a multivariate pattern analysis of DTI data. *European Radiology*. https://doi.org/10.1007/s00330-017-4820-1

Thompson, P. M., Dutton, R. A., Hayashi, K. M., Lu, A., Lee, S. E., Lee, J. Y., … Becker, J. T. (2006). 3D mapping of ventricular and corpus callosum abnormalities in HIV/AIDS. *Neuroimage*, *31*(1), 12–23.

Thompson, P. M., Dutton, R. A., Hayashi, K. M., Toga, A. W., Lopez, O. L., Aizenstein, H. J., & Becker, J. T. (2005). Thinning of the cerebral cortex visualized in HIV/AIDS reflects CD4+ T lymphocyte decline. *Proceedings of the National Academy of Sciences*, *102*(43), 15647–15652.

Thompson, P. M., & Jahanshad, N. (2015). Novel Neuroimaging Methods to Understand How HIV Affects the Brain. *Current HIV/AIDS Reports*. https://doi.org/10.1007/s11904-015-0268-6

Toborek, M., Lee, Y. W., Flora, G., Pu, H., András, I. E., Wylegala, E., … Nath, A. (2005). Mechanisms of the blood-brain barrier disruption in HIV-1 infection. *Cellular and Molecular Neurobiology*. https://doi.org/10.1007/s10571-004-1383-x

Tohka, J., Moradi, E., Huttunen, H., & Alzheimer'S Disease Neuroimaging Initiative. (2016). Comparison of Feature Selection Techniques in Machine Learning for Anatomical Brain MRI in Dementia. *Neuroinformatics*. https://doi.org/10.1007/s12021-015-9292-3

Tournier, J. Donald. (2019). Diffusion MRI in the brain – Theory and concepts. *Progress in*

Nuclear Magnetic Resonance Spectroscopy. https://doi.org/10.1016/j.pnmrs.2019.03.001

Tournier, Jacques Donald, Mori, S., & Leemans, A. (2011). Diffusion tensor imaging and beyond. *Magnetic Resonance in Medicine*. https://doi.org/10.1002/mrm.22924

Trefler, A., Sadeghi, N., Thomas, A. G., Pierpaoli, C., Baker, C. I., & Thomas, C. (2016). Impact of time-of-day on brain morphometric measures derived from T1-weighted magnetic resonance imaging. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.02.034

Uban, K. A., Herting, M. M., Williams, P. L., Ajmera, T., Gautam, P., Huo, Y., … Sowell, E. R. (2015). White matter microstructure among youth with perinatally acquired HIV is associated with disease severity. *AIDS*. https://doi.org/10.1097/QAD.0000000000000648

UNAIDS. (2017). UNAIDS fact sheet - Latest statistics on the status of the AIDS epidemic. *Ending the Aids Epidermics*. https://doi.org/2017

UNICEF. (2018). Global and Regional Trends - UNICEF DATA.

van Arnhem, L. A., Bunders, M. J., Scherpbier, H. J., Majoie, C. B. L. M., Reneman, L., Frinking, O., … Pajkrt, D. (2013). Neurologic Abnormalities in HIV-1 Infected Children in the Era of Combination Antiretroviral Therapy. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0064398

van der Kouwe, A. J. W., Benner, T., Salat, D. H., & Fischl, B. (2008). Brain morphometry with multiecho MPRAGE. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2007.12.025

van Dijk, J. H., Sutcliffe, C. G., Munsanje, B., Sinywimaanzi, P., Hamangaba, F., Thuma, P. E., & Moss, W. J. (2011). HIV-infected children in rural Zambia achieve good immunologic and virologic outcomes two years after initiating antiretroviral therapy. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0019006

Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2017.06.061

Verani, A., Gras, G., & Pancino, G. (2005). Macrophages and HIV-1: Dangerous liaisons. *Molecular Immunology*. https://doi.org/10.1016/j.molimm.2004.06.020

Vestergaard, M., Skakmadsen, K., Baaré, W. F. C., Skimminge, A., Ejersbo, L. R., Ramsøy, T. Z., … Jernigan, T. L. (2011). White matter microstructure in superior longitudinal fasciculus associated with spatial working memory performance in children. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.2010.21592

Violari, A., Cotton, M. F., Gibb, D. M., Babiker, A. G., Steyn, J., Madhi, S. A., … McIntyre, J. A. (2008). Early antiretroviral therapy and mortality among HIV-infected infants. *The New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa0800971

Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn1704

Wade, B. S. C., Valcour, V. G., Puthanakit, T., Saremi, A., Gutman, B. A., Nir, T. M., … Jahanshad, N. (2019). Mapping abnormal subcortical neurodevelopment in a cohort of Thai children with HIV. *NeuroImage: Clinical*. https://doi.org/10.1016/j.nicl.2019.101810

Wang, G., Wei, W., Jiang, J., Ning, C., Chen, H., Huang, J., … Ye, L. (2019). Application of a long short-term memory neural network: A burgeoning method of deep learning in forecasting HIV incidence in Guangxi, China. *Epidemiology and Infection*. https://doi.org/10.1017/S095026881900075X

Wang, Y., Zhang, J., Gutman, B., Chan, T. F., Becker, J. T., Aizenstein, H. J., … Thompson, P. M. (2010). Multivariate tensor-based morphometry on surfaces: Application to mapping ventricular abnormalities in HIV/AIDS. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2009.10.086

Whelan, R., & Garavan, H. (2014). When optimism hurts: Inflated predictions in psychiatric neuroimaging. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2013.05.014

White, T. (2002). Brain Volumes and Surface Morphology in Monozygotic Twins. *Cerebral Cortex*. https://doi.org/10.1093/cercor/12.5.486

Wilson, T. W., Heinrichs-Graham, E., Becker, K. M., Aloi, J., Robertson, K. R., Sandkovsky, U., … Swindells, S. (2015). Multimodal neuroimaging evidence of alterations in cortical structure and function in HIV-infected older adults. *Human Brain Mapping*. https://doi.org/10.1002/hbm.22674

Woo, C. W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*. https://doi.org/10.1038/nn.4478

World Health Organisation, W. (2017). WHO | HIV and youth. *Who*.

Wu, M. J., Mwangi, B., Bauer, I. E., Passos, I. C., Sanches, M., Zunta-Soares, G. B., … Soares, J. C. (2017). Identification and individualized prediction of clinical phenotypes in bipolar disorders using neurocognitive data, neuroimaging scans and machine learning.

NeuroImage. https://doi.org/10.1016/j.neuroimage.2016.02.016

Yadav, S. K., Gupta, R. K., Garg, R. K., Venkatesh, V., Gupta, P. K., Singh, A. K., … Haris, M. (2017). Altered structural brain changes and neurocognitive performance in pediatric HIV. *NeuroImage: Clinical*. https://doi.org/10.1016/j.nicl.2017.01.032

Ye, J., Farnum, M., Yang, E., Verbeeck, R., Lobanov, V., Raghavan, N., … Narayan, V. A. (2012). Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. *BMC Neurology*. https://doi.org/10.1186/1471-2377-12-46

Zhang, D., Wang, Y., Zhou, L., Yuan, H., & Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2011.01.008

Zhang, Y., & Wu, L. (2012). An MR brain images classifier via principal component analysis and kernel support vector machine. *Progress in Electromagnetics Research*. https://doi.org/10.2528/PIER12061410

Zhou, Q., Chen, W., Song, S., Gardner, J. R., Weinberger, K. Q., & Chen, Y. (2015). A reduction of the Elastic Net to Support Vector machines with an application to GPU computing. In *Proceedings of the National Conference on Artificial Intelligence*.

Zhou, Y., Li, R., Wang, X., Miao, H., Wei, Y., Ali, R., … Li, H. (2017). Motor-related brain abnormalities in HIV-infected patients: a multimodal MRI study. *Neuroradiology*. https://doi.org/10.1007/s00234-017-1912-1

Zhu, T., Zhong, J., Hu, R., Tivarus, M., Ekholm, S., Harezlak, J., … Schifitto, G. (2013). Patterns of white matter injury in HIV infection after partial immune reconstitution: A DTI tract-based spatial statistics study. *Journal of NeuroVirology*. https://doi.org/10.1007/s13365-012-0135-9

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Zucchella, C., Federico, A., Martini, A., Tinazzi, M., Bartolo, M., & Tamburin, S. (2018). Neuropsychological testing. *Practical Neurology*. https://doi.org/10.1136/practneurol-2017-001743

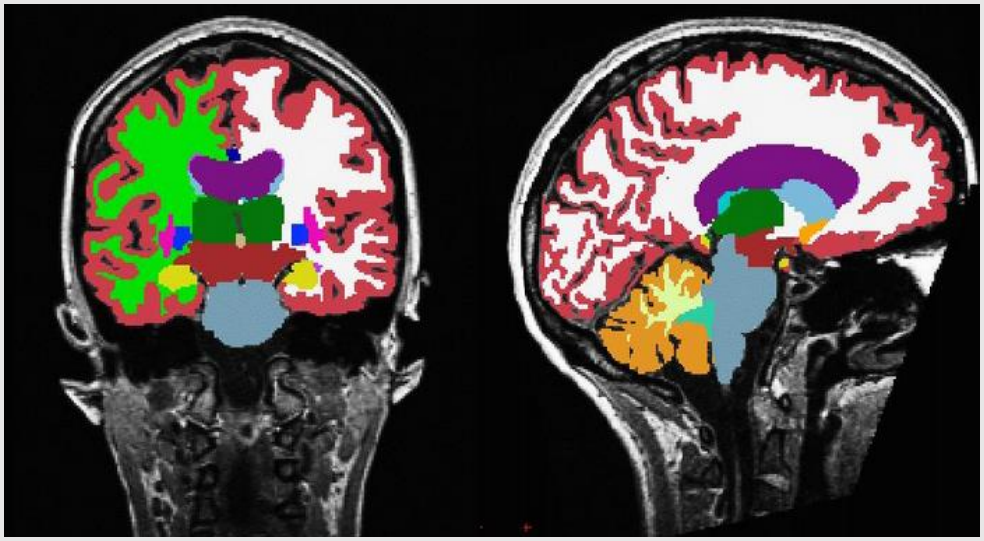
Zuo, P., Wang, Y., Liu, J., Hu, S., Zhao, G., Huang, L., & Lin, D. (2019). Effects of early adversity on the brain: Larger-volume anterior cingulate cortex in AIDS orphans. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0210489

APPENDIX

List of candidate features in each set

Structural volumes (42 candidate features):

Voxels in the normalised brain volume were assigned one of the following labels by FreeSurfer:



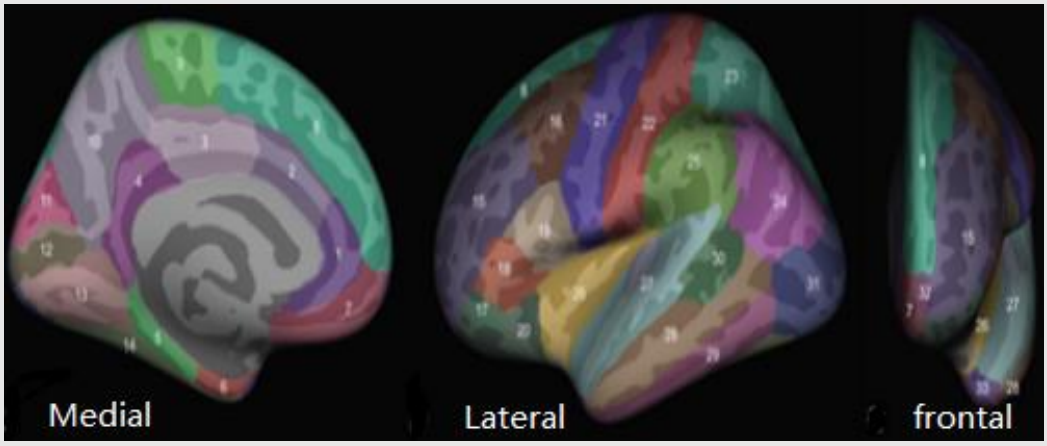
- Lateral ventricle (L/R)
- Inferior lateral ventricle (L/R)
- Cerebral cortex (L/R)
- Thalamus proper (L/R)
- Caudate (L/R)
- Putamen (L/R)
- Globus pallidus (L/R)
- Third ventricle (L/R)
- Brainstem
- Hippocampus (L/R)
- Amygdala (L/R)
- Cerebrospinal fluid (CSF)
- Accumbens area (L/R)
- Cerebellum white matter (L/R)
- Vessel (L/R)
- Choroid plexus (L/R)
- Fourth ventricle (L/R)
- Optic chiasm
- Corpus callosum (P/MP/C/MA/A)
- Ventral diencephalon (L/R)

L – left, R – right, P – posterior, MP – mid-posterior, C – central, MA – mid-anterior

A - anterior

Cortical thickness and LGI (68 candidate features each)

Cortical thickness and folding measures were from 68 brain regions (34 structures in left and right hemispheres). ROIs from Desikan brain template for each hemisphere are listed below:

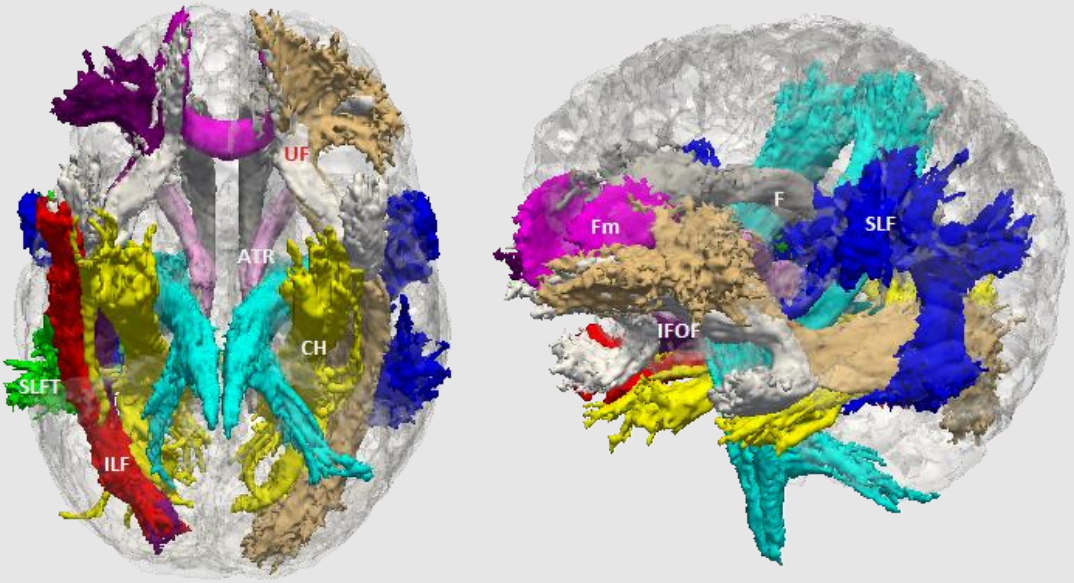


1. Rostral anterior cingulate gyri
2. Caudal anterior cingulate gyri
3. Posterior cingulate gyri
4. Isthmus of cingulate gyri
5. Parahippocampal gyri
6. Entorhinal cortex
7. Medial orbitofrontal cortex
8. Superior frontal gyri
9. Paracentral lobule
10. Precuneus
11. Cuneus
12. Pericalcarine cortex
13. Lingual gyri
14. Fusiform gyri
15. Rostral middle frontal gyri
16. Caudal middle frontal gyri
17. Pars orbitalis of inferior frontal gyri
18. Pars triangularis of inferior frontal gyri
19. Pars opercularis of inferior frontal gyri
20. Lateral orbitofrontal cortex
21. Precentral gyri
22. Postcentral gyri
23. Superior parietal cortex
24. Inferior parietal cortex
25. Supramarginal gyri
26. Insular cortex
27. Superior temporal gyri
28. Middle temporal gyri
29. Inferior temporal gyrus
30. Banks of the superior temporal sulcus
31. Lateral occipital cortex
32. Frontal pole
33. Temporal pole
34. Transverse temporal cortex

Adapted from Desikan et al. (2006)

DTI (80 candidate features):

FA, AD, RD, and MD were extracted from ROIs defined by the JHU atlas.



1. SLF – superior longitudinal fasciculus (L/R)
2. CG – cingulum (L/R)
3. ATR – anterior thalamic radiation (L/R)
4. UF – uncinate fasciculus (L/R)
5. SLFT – temporal part of superior longitudinal fasciculus (L/R)
6. ILF – inferior longitudinal fasciculus (L/R)
7. IFOF – inferior fronto-occipital fasciculus (L/R)
8. Fm - Forceps minor
9. F - Forceps major
10. CST - Corticospinal tract (L/R)
11. CH - Cingulum hippothalamus (L/R)

L – left, R - right

Adapted from Mori (2007)

^{1}H-MRS in bg, pwm, and mfgm (28 candidate features each)

Estimation of neurometabolites by LCModel were obtained in the basal ganglia, peritrigonal white matter, and midfrontal gray matter. The list of metabolites from each region is given below.

Metabolites

1. Cr - creatine
2. Glu – glutamate
3. mI – myo-inositol
4. NAA – N-acetylaspartate
5. Gua – guanidinoacetate
6. NAANAAG – ratio of N-acetylaspartylglutamate to N-acetylaspartate
7. GPC - Glycerophosporylcholine
8. GPCPCh – total choline
9. GluGln – ratio of glutamate to glutamine
10. MM20 – macromolecules at 20ppm
11. Lip13a – lipids at 13ppm
12. Lip13b – lipids at 13ppm
13. Lip20 – lipids at 20ppm
14. MM20Lip20 – macromolecules and lipids at 20ppm

Ratio of metabolites to creatine

15. CrCrPCr – ratio of creatine to creatine
16. GluCrPCr – ratio of glutamate to creatine
17. MICrPCr – ratio of myo-inositol to creatine
18. NAACrPCr – ratio of N-acetylaspartate to creatine
19. GuaCrPCr – ratio of guanidinoacetate to creatine
20. NAANAAGCrPCr ratio of NAANAAG to creatine
21. GPCCrPCr ratio of GPC to creatine
22. GPCPChCrPCr ratio of total choline to creatine
23. GluGlnCrPCr ratio of GluGln to creatine
24. MM20CrPCr – ratio of macromolecules at 20ppm to creatine
25. Lip13aCrPCr – ratio of lipids at 13ppm to creatine
26. Lip13bCrPCr – ratio of lipids at 13ppm to creatine
27. Lip20CrPCr – ratio of lipids at 20ppm to creatine
28. MM20Lip20 CrPCr– ratio of macromolecules and lipids at 20ppm to creatine

LCModel, Provencher (2001)

Programming code

Used packages:

```
library(tidyverse)      # clean, process, model, and visualise data
library(haven)          # read and write various data formats
library(Hmisc)          # advanced table making and utility operations
library(caret)          # data splitting, pre-processing, and tuning
library(AUC)            # computation of area under the ROC
library(pROC)           # roc functions
library(ROCR)           # visualising performance of scoring classifiers
library(mice)           # multivariate imputation by chained equations
library(glmnetUtils)    # optimising e-net parameters simultaneously
library(ModelMetrics)   # evaluating models: LRT, AIC
library(AICcmodavg)     # advanced AIC and BIC models
```

Classification and feature selection were implemented in R using the following sample script:

```
#Load feature set
feature_set <- read.csv(file = "C:/Users/KHBISA001/Documents
                        /OneDrive - University of Cape Town/Clean_Datasets/data.txt",
                        header = TRUE, row.names = NULL, encoding = "UTF-8", sep = ",",
                        dec = ".", quote = "\"", comment.char = "")
#----------------------------------START-------------------------------------------------#
#-------Measuring Classification Performance Using k fold cross validation------------#
#-------------------------------------------------------------------------------------#
set.seed(1)
# Initialise loop variables:
mean_acc <- 0                   # Mean accuracy
mean_auc <- 0                   # Mean auc
mean_sens <- 0                  # Mean sensitivity
mean_spec <- 0                  # Mean specificity

#----------------------100 repeat cycle loop-----------------------------------------#
for (j in 1:100) {
    # List of observations
    feature_set <- feature_set$ID
    # Create 10 folds for the observations
    observ_folds <- createFolds(feature_set, k = 10, list = T, returnTrain = F)
    # Initialise nested loop output variables
    opt_alpha <- 0                                          # Optimised alpha
    opt_lambda <- 0                                         # Optimised lambda
    cv_acc <- 0                                             # Mean accuracy
    cv_auc <- 0                                             # Mean auc
    cv_sens <- 0                                            # Mean sensitivity
    cv_spec <- 0                                            # Mean specificity
    non_zero_coef                                           # Non-zero features
#----------------------10 fold cross-validation loop---------------------------------#
    for (i in 1:10) {
        test_fold_indices <- feature_folds[[j]]             # Testing fold indices
        test_set <- feature_set[test_fold_indices,]         # Testing/validation set
        train_set <- feature_set[-test_fold_indices,]       # Training set indices
        # standardise (z-score) training set
        x_train <- scale(select(train_set, - status, - ID))
        #standardise (z-score)
        x_test <- scale(select(test_set, - status, - ID))
        # Responses of training and validation sets binarised: control (0), HIV (1)
        y_train <- train_set$status
        y_test <- test_set$status
```

```
        # Create penalisation vector
        pen_loop <- rep(1, length(x_train))                # penalise all variables...
        pen_loop[length(x_train)-2:length(x_train)] <- 0   # except age, gender, TIV
        # Train model
        train_model <- cva.glmnet(as.matrix(x_train), y_all, family = c("binomial"),
                                  offset = NULL, nlambda = 200, type.measure = 'auc',
                                  nfolds = 10, penalty.factor = pen_loop, maxit = 200000)
        # Optimum alpha
        opt_alpa[i] <- train_model$alpha
        # Create a prediction model
        predict_model <- predict(train_model, as.matrix(x_test),
                            alpha = train_model$alpha, type = 'response')
        # Create prediction object
        pred_object <- prediction(predict_model, y_test)
        # Area under the receiver output characteristic curve ROC(AUC)
        cv_auc[i] <- auc(testing_set$status, predict_model)
        # Accuracy of nested loop
        a <- performance(pred_object, 'acc')
        cv_acc[l] <- max(unlist(a@y.values))
        # Sensitivity and specificity for nested loop
        roc_curve <- performance(pred_object, 'sens', 'spec')
        # Finding optimum cut-off point for sensitivity and specificity
        opt.cut = function(perf, pred) {
            cut.ind = mapply(FUN = function(x, y, p) {
                d = (x - 0) ^ 2 + (y - 1) ^ 2
                ind = which(d == min(d))
                c(sensitivity = y[[ind]], specificity = 1 - x[[ind]],
            cutoff = p[[ind]])
            }, perf@x.values, perf@y.values, pred@cutoffs)
        }
        cv_spec <- opt.cut(roc_curve, pred_object)[1]                  # specificity
        cv_spec <- opt.cut(roc_curve, pred_object)[2]                  # sensitivity
    }
    # Train model on unpartitioned dataset
    x_train_fs <- as.matrix(scale(select(feature_set, -status, -ID)))
    train_model <- cv.glmnet(x_train, y_all, family = c("binomial"), alpha = opt_alpha,
                             nlambda = 200,type.measure = 'auc', penalty.factor = pen_loop,
                             maxit = 200000)
    # Significant coefficents using optimised lambda
    non_zero_coef[j]<- coef(train_model_all, s = train_model_all$lambda.min)
    # Number of significant features in each repeat
    non_zero_features_all[j] <- sum(as.matrix(non_zero_coef) != 0, na.rm = 0) - 1
    # average cross validation (loop) erros, sensitivity, and specificity
    mean_acc[j] <- mean(cv_acc)
    mean_auc[j] <- mean(cv_auc)
    mean_sens[j] <- mean(cv_sens)
    mean_spec[j] <- mean(cv_spec)
}
```

Model evaluation with second order AIC

```
aic_models <- list()
aic_models[[1]] <- glm(status ~ Age7yrscan                    # volumes
                       + Gender
                       + EstimatedTotalIntraCranialVol
                       + Rightvessel
                       + OpticChiasm
                       + LeftAmygdala
                       + LeftPallidum
                       + RightPallidum
                       + Rightchoroidplexus
                       + RightAmygdala
                       + CC_Posterior
                       + RightVentralDC
                       + CSF
                       + LeftVentralDC
                       + RightCerebellumWhiteMatter, family = binomial,
                       data = volume_dti_bg)
aic_models[[2]] <- glm(status ~ Age7yrscan                  # diffusivities
                       + Gender
                       + RD_ucinate_fasc_l
                       + MD_ucinate_fasc_r
                       + MD_cortsp_tract_l
                       + MD_cing_gyr_r
                       + RD_ucinate_fasc_r
                       + RD_cortsp_tract_l
                       + RD_cortsp_tract_r
                       + AD_ucinate_fasc_r
                       + MD_ucinate_fasc_l
                       + MD_cing_gyr_l
                       + AD_cortsp_tract_l
                       + RD_cing_gyr_l
                       + MD_cortsp_tract_r, family = binomial,
                       data = volume_dti_bg)
aic_models[[3]] <- glm(status ~ Age7yrscan                  # combined
                       + Gender
                       + Race
                       + EstimatedTotalIntraCranialVol
                       + Rightvessel
                       + OpticChiasm
                       + LeftAmygdala
                       + LeftPallidum
                       + RightPallidum
                       + RightAmygdala
                       + CSF
                       + LeftVentralDC
                       + RightCerebellumWhiteMatter
                       + BrainStem
                       + RightLateralVentricle
                       + RD_dti_ucinate_fasc_l + MD_dti_ucinate_fasc_r
                       + MD_dti_ucinate_fasc_l + RD_dti_ucinate_fasc_r
                       + MD_dti_cing_gyr_r + MD_dti_cing_gyr_l
                       + MD_dti_cortsp_tract_l + RD_dti_cing_gyr_r
                       + AD_dti_ucinate_fasc_r + AD_dti_cortsp_tract_l
                       + AD_dti_sup_long_fasc_temp_r,
                       family = binomial, data = volume_dti_bg)
aictab(aic_models, modnames = c("sMRI volumes", "DTI", "Volumes + DTI"))
```

ETHICS

This study complies with the Human Research Ethics Committee of the University of Cape Town

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E53-46 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6626
Email: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

11 March 2019

HREC REF: 123/2019

Dr Frances Robertson
Biomedical Engineering
Anatomy Building, Room 7.22

Dear Dr Robertson

PROJECT TITLE: MULTIMODAL NEUROIMAGING BIOSIGNATURES OF HIV: DISTINGUISHING PERINATALLY INFECTED CHILDREN ON EARLY ART USING AN ELASTIC-NET REGULARIZATION (SUB-STUDY LINKED TO 448/2011) MSc Candidate - Mr I Khobo

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

Approval is granted for one year until the 30 March 2020.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please quote the HREC REF in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

The HREC acknowledge that the student, Isaac Khobo will also be involved in this study.

Yours sincerely

Signature Removed

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

HREC 123/2019

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DoH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI), and Declaration of Helsinki (2013) guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.