Neurocognitive Performance Profiling in People Living with HIV

by

Daniela Isabel Gomez

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Psychiatry
University of Alberta

ABSTRACT

Despite the wide availability of effective antiretroviral therapy (ART), neurocognitive problems persist in over 20% of people living with HIV. Neurocognitive problems are highly heterogeneous in HIV, and the underlying etiology of neurocognitive impairment (NCI) is multifactorial. The heterogeneity of NCI in HIV may not be fully captured by the current clinical staging criteria for HIV-associated neurocognitive disorders (HAND). Given the multifactorial nature of NCI, multiple neurocognitive impairment profiles associated with distinct risk-factors may exist. Here, we investigated the neurocognitive performance of 381 HIV-positive participants from the Southern Alberta HIV clinic (SAC) cohort to identify different neurocognitive impairment profiles and their associated risk-factors. Neurocognitive performance was assessed by multi-domain neuropsychological test battery. We implemented Latent Profile Analysis (LPA) to analyze performance and empirically define neurocognitive profiles. Random Forest Analysis (RFA), a machine learning technique, was applied to identify the most important factors associated with these profiles. Three profiles emerged from the LPA: Profile 1 (P1, n=159) achieved the highest performance, while Profile 2 (P2, n=163) had lowered executive functions and verbal memory, and Profile 3 (P3, n=59) was globally impaired. RFA achieved good prediction (area-under-the-curve ≥ 0.80) only for global impairment (P3). Non-North American descent was the dominant predictor of P3, followed by factors coinciding with non-North American descent (female sex and toxoplasma seropositivity). Additional predictors included unemployment, current depressive symptoms, lower nadir CD4, and longstanding HIV. Restricting analyses to North Americans pointed to the additional importance of ART achieving high CSF levels, and older age in prediction of P3. HAND diagnoses were most common in the globally impaired profile (P3=89.8%), followed by the group with reduced higher-order

neurocognitive performance (P2=16.6%). Implementation of LPA and RFA empirically distinguished three distinct neurocognitive performance profiles in this cohort while also highlighting potential risk factors and their relative importance to neurocognitive impairment. These data-driven analytical methods pointed to discernible demographic, HIV- and treatment-related risk-factor constellations in patients born outside and within North America that might influence diagnostic and therapeutic decisions.

Preface

Sections from the literature review in Chapter 2 and the discussion in Chapter 5 were adapted from Gomez, D., Power, C., & Fujiwara, E, (2018), "Neurocognitive Impairment and Associated Genetic Aspects in HIV Infection", Current Topics in Behavioural Neuroscience, doi:10.1007/7854_2018_69, and Gomez, D., Power, C., Gill, M. J., Koenig, N., Vega, R., & Fujiwara, E, (2019), "Empiric neurocognitive performance profile discovery and interpretation in HIV infection", Journal of Neurovirology, vol. 25, issue 1, pages 72-84, doi:10.1007/s13365-018-0685-6.

Chapter 3 and Chapter 4 are reprints of Gomez, D., Power, C., Gill, M. J., Koenig, N., Vega, R., & Fujiwara, E, (2019), "Empiric neurocognitive performance profile discovery and interpretation in HIV infection", Journal of Neurovirology, vol. 25, issue 1, pages 72-84, doi:10.1007/s13365-018-0685-6.

This research project was approved by the University of Calgary, Conjoint Health Research Ethics Board; CHREB13-0615.

Acknowledgements

I must begin by first thanking you, Dr. Esther Fujiwara. Were it not for your guidance, I would have not have achieved all I have accomplished over these past three years. You have been a selfless mentor, generously volunteering your time to me, and patiently teaching me about the world of statistics, a world of numbers I never imaged I would come to love. You cultivated my talent for data analysis, and I will remember fondly the many hours we spent pouring over data, results, and troubleshooting our machine learning algorithm. You have been a dependable mentor; your feedback pushed me to better my work and work on my weaknesses. Were it not for your friendship, I would not have been able to surpass the many professional and personal challenges that came to pass. You have been so incredibly kind, understanding and sympathetic to my struggles during my graduate degree. For all of this, Esther, I am infinitely grateful to you.

Thank you, Dr. Chris Power, for being so welcoming and for believing and trusting in my abilities. You always valued my opinions and helped me bolster my confidence. You have been invested in my growth since the beginning, and I am so thankful for all the times you made yourself available to me for guidance, not just in research, but life in general. You have been a kind, encouraging and compassionate mentor to me. Your passion for research and for all the work you do is inspiring.

Thank you, Dr. Roger Dixon for your insightful feedback, and your encouragements. Thank you for sharing your time and expertise with me. Your meaningful questions and suggestions helped improve all my projects. I also want to thank you and Kirstie McDermott for generously sharing part of your machine learning code with me; without it, this project would only be half as interesting.

Thank you to Dr. M. John Gill, Brenda Beckholt, Noshin Koenig, and all the clinical staff at the Southern Alberta Clinic (SAC). You were all so welcoming and supportive during my time at the clinic. Thank you for taking the time out of your incredibly busy days to train me. I loved working with all of you. Thank you for letting me be a part of your team.

Thank you for all the amazing participants from the SAC clinic; without you volunteering your time and data; none of this would have been possible. You are all inspiring.

Thank you, Tara Checknita, for all of your advice, for guiding me though the process of putting together Research Day for the past few years, and for being such a good friend. You made my short time at the RTF even more enjoyable.

Thank you, Mom, Dad, and German, for all your love and encouragement through this journey. This was not an easy ride, but you all have been such an amazing emotional support system for me.

And lastly, thank you to my beautiful Abuela. You may be no longer with us, but I know you were with me every step of the way.

Table of Contents

Section **Page**

1. INTRODUCTION 1

2. BACKGROUND 3

2.1. Epidemiology of HIV 3

2.2. Pathophysiology of HIV and neurological complications 4

2.3. Diagnosis of neurocognitive impairment in HIV 5

2.4. Challenges with diagnosing NCI in HIV 8

2.4.1. Multifactorial nature of NCI 8

2.4.2. Measurement problems 11

2.4.3. Alternative approaches to assessing NCI in HIV 13

2.5. Aims and hypotheses 16

3. MATERIALS AND METHODS 17

3.1. Cohort 17

3.2. Neuropsychological test battery 17

3.3. Demographic and medical variables 21

3.4. Statistical analyses 22

4. RESULTS 27

4.1. Latent profile analysis 27

4.2. Univariate differences between profiles and comparisons with HAND diagnosis 31

4.3. Random forest analyses 35

4.4. North American participants 41

5. DISCUSSION 50

5.2. Confounding influences of demographic variability on NCI 52

5.3. Risk factors for neurocognitive impairment in North American patients 55
5.3.1. Demographic variables 56
5.3.2. HIV- and treatment-related variables 57
5.3.3. Comorbidities 58
5.4. Strengths and limitations 59
5.5. Future directions 61
5.6. Conclusion 62

List of Tables

Table **Page**

1. Neuropsychological test descriptions and mean z-scaled performance (standard deviations) for patients from the Southern Alberta Clinic (SAC) cohort. 20

2. List of variables considered for inclusion in the Random Forest Analyses (x = included). . . 26

3. Synthetic Minority Over-sampling Technique (SMOTE) parameters used to balance the size of P3 (entire cohort; N = 381) 27

4. Latent Profile Analyses on the ten neurocognitive indicators. 28

5. Mean z-scaled performance (standard deviations) in the ten neuropsychological indicators used in Latent Profile Analyses in the entire cohort (N = 381). 29

6. Patient characteristics and differences between LPA profiles (N = 381) 33

7. List of variables considered for inclusion in the Random Forest Analyses between North American (NA) and non-North American (non-NA) participants (x = included). 39

8. Synthetic Minority Over-sampling Technique (SMOTE) parameters used to balance the size of non-North American group (entire sample, N = 381) 40

9. Mean z-scaled performance (standard deviations) in the ten neuropsychological indicators used in Latent Profile Analyses in North American-born patients (n = 300) 43

10. Patient characteristics and differences between latent profiles in North American-born patients (n = 300) 44

11. Synthetic Minority Over-sampling Technique (SMOTE) parameters used to balance the size of P3 (North American only; n = 300) 46

List of Figures

Figures **Page**

1. Decision-tree by Woods, Moore, Weber, and Grant (2009) for the diagnosis of HIV-associated neurocognitive disorders (HAND) according to the Frascati criteria.. 7

2. Pathological mechanisms underlying HIV-associated neurocognitive disorders according to Nightingale et al., (2014) 9

3. Latent Profile Analyses and Random Forest Analyses of 381 HIV-infected participants from the Southern Alberta HIV Clinic. 30

4. Summary of important variables predicting neurocognitive performance in the Southern Alberta Clinic cohort (N= 381) identified by Random Forest Analysis. 37

5. Random Forest Analyses (RFA), showing the importance for specific variables distinguishing North American (NA) versus non-North American-born (non-NA) participants. 40

6. Latent Profile Analyses and Random Forest Analyses of 300 HIV-infected North American participants from the Southern Alberta HIV Clinic. 48

7. Summary of important variables predicting neurocognitive performance in the North American patients from the Southern Alberta Clinic cohort (n = 300) identified by Random Forest Analysis. 49

List of Abbreviations

AAN: American Academy of Neurology

AD: Alzheimer's disease

ADC: AIDS-dementia complex

AIDS: acquired immunodeficiency syndrome

ANI: asymptomatic neurocognitive impairment

ART: antiretroviral therapy

AUC: area under the receiver operating curve

BBB: blood brain barrier

CNS: central nervous system

Cond.: condition

CPE: CNS penetrance effectiveness

D-KEFS: Delis-Kaplan Executive Function System

Dis.: disorder

GPB: grooved pegboard

HAD: HIV-associated dementia

HAND: HIV-associated neurocognitive disorder

HCV: Hepatitis C Virus

HIV: human immunodeficiency virus

HVLT-R: Hopkins Verbal Learning Test- Revised

LPA: Latent Profile Analysis

MCI: mild cognitive impairment

MCMD: HIV-associated minor cognitive/motor disorder

MND: mild neurocognitive disorder

NA: North America

NCI: neurocognitive impairment

PHQ-9: Patient Health Questionnaire-9

QOL: quality of life

RFA: Random Forest Analysis

ROC = receiver operating characteristic

SAC: Southern Alberta Clinic

SDMT: Symbol Digit Modalities Test

SMOTE: Synthetic Minority Over-sampling Technique

TMT: Trail Making Task

Vio.: violence

VL: viral load

WCST: Wisconsin Card Sorting Task

WRAT-4: Wide Range Achievement Test -4

1. INTRODUCTION

The Human Immunodeficiency Virus (HIV) is a systemic virus known for its insidious effects on the immune system. HIV can also affect the brain, leading to neuropathological changes and downstream neurocognitive impairment (NCI) in multiple cognitive domains including attention, memory, language, executive and motor functions in some patients. NCI in patients with HIV is associated with difficulties affecting multiple aspects of everyday life activities such as difficulties with financial management, independence, and employment, all resulting in poor quality of life (Andrade et al., 2013; Doyle et al., 2016). Yet, the pattern of NCI is not uniform and this heterogeneity may not be fully captured by the current clinical staging criteria for HIV-associated neurocognitive disorders (HAND). It has been recently suggested that the definition and measurement of NCI in HIV may benefit from the integration of empirical statistical methods to classify neurocognitive impairment profiles, in addition to (or instead of) following clinical staging guidelines that focus on severity rather than type of NCI (Devlin & Giovannetti, 2017). Furthermore, many factors other than HIV are known to accompany NCI (Iudicello, Morgan, Hussain, Watson, & Heaton, 2019; Nightingale et al., 2014), and current analytical methods are not appropriate to assess all of these potential contributors at once. If there are multiple NCI profiles in HIV, these may also be associated with different underlying risk factors. The following questions were addressed in the current study: 1) Are there distinguishable profiles of neurocognitive impairment in patients living with HIV? 2) What are the factors associated with differentiable neurocognitive impairment patterns in HIV? In order to address these questions, we categorized neurocognitive profiles of 381 HIV-positive participants from the Southern Alberta HIV Clinic (SAC) in Calgary. We assessed their neuropsychological performance using a comprehensive neuropsychological test battery, and analyzed their

performance patters using a Latent Profile Analysis (LPA). A machine learning technique, Random Forest Analysis (RFA), was applied to identify the most important factors associated with these performance profiles. Understanding the patterns of impairment in HIV, and their potential risk-factors, may help clinicians to optimize current cognitive rehabilitation and intervention strategies to help stop, reverse, or prevent neurocognitive impairment in people living with HIV, and ultimately increase the quality of life of patients.

2. BACKGROUND

2.1. Epidemiology of HIV

The Human Immunodeficiency Virus (HIV) is an infectious virus that leads to the depletion of the CD4+ T-lymphocyte cells, causing the progressive weakening of the immune system (World Health Organization [WHO], 2018). The virus is transmitted via direct contact with infected fluids, such as through sexual contact, blood transfusions, needle sharing, and birth (WHO, 2018). Globally, 36.9 million people live with HIV (WHO, 2018) including over 60,000 Canadians (Public Health Agency of Canada, 2018). In Canada, the HIV epidemic is concentrated to specific population groups. The largest group (49.1%) of infected persons identify as gay or bisexual males, and men who have sex with men (Public Health Agency of Canada, 2018). Injection drug users represent 14.6% of people living with HIV in Canada. Indigenous people and people from countries where HIV is endemic (such as sub-Sahara Africa and the Caribbean) represent a disproportionately high proportion of people living with HIV in Canada, at 9.6% and 15% respectively (Public Health Agency of Canada, 2018). The number of people currently living with HIV in Canada is increasing; although new infections contribute to the increase, previously infected individuals who are receiving effective treatment are now living longer with the disease and account for a large proportion of this increase (Public Health Agency of Canada, 2018).

Historically, HIV infection led to the progressive depletion of CD4+ T-cells and resulted in a weakened immune system. Prior to the invention of effective combination antiretroviral therapies (ART), HIV infections would progress to Acquired Immune Deficiency Syndrome (AIDS), a usually fatal stage defined by a nadir CD4+ T-cell count, the lowest CD4 count ever recorded, equal or lower to 200 cells per cubic millilitres of blood (200 cell/mm^3) (CDC, 2015).

However, following the introduction of potent ART, HIV is now considered a manageable, chronic illness and most treated individuals reach near normal life expectancies.

2.2. Pathophysiology of HIV and neurological complications

Although primarily a systemic disease, HIV can enter the central nervous system early in the infection (Davis et al., 1992). It is currently hypothesized that HIV enters the brain via a "Trojan horse" mechanism, where peripherally HIV-infected macrophages and lymphocytes cross the blood-brain barrier (BBB) and release the virus into the CNS. Once within the CNS, free virions infect new macrophages resulting in an exponential increase in viral replication. Viral replication then activates surrounding microglia leading to the secretion of immune and viral neurotoxic molecules that induce inflammation and increase blood-brain barrier's permeability, further facilitating the entry of the virus. Although astrocytes do not support viral replication, HIV can enter astrocytes and induce astrogliosis, further contributing to the neuroinflammation (Kaul, Garden, & Lipton, 2001). HIV also does not directly affect neurons, but the ensuing inflammation and continued viral replication within the CNS and microphages lead to the disruption of neural functioning and, ultimately, neuronal injury.

Imaging studies have shown that both cortical and subcortical brain regions are affected by HIV. Volume loss in subcortical areas (i.e.., caudate, putamen, amygdala, hippocampus, and thalamus) and white matter regions (i.e., the corpus callosum, orbitofrontal cortex, parietal and frontal cortices) and cortical thinning occur in people living with HIV (Ances, Ortega, Vaida, Heaps, & Paul, 2012; Chang et al., 2011; Chiang et al., 2007; Jernigan et al., 2005; Thompson et al., 2005). Changes in cortex and subcortical structures have been often associated to low nadir CD4 counts as well as lower cognitive performance (Chang et al., 2011; Sanford et al., 2017;

Thompson et al., 2005). Therefore, the CNS disturbances arising from HIV viral replication are believed to underlie some of the behavioural and neurocognitive impairment (NCI) observed in some people living with HIV.

2.3. Diagnosis of neurocognitive impairment in HIV

The expression and diagnosis of NCI associated with HIV infection has undergone a dramatic change after the introduction of modern ART. In the pre-ART era, marked motor deficits and dementia-like signs and symptoms were frequently observed in a subset of patients with AIDS (Navia, Jordan, & Price, 1986). These signs and symptoms were then known as AIDS dementia complex (ADC), HIV encephalopathy, or HIV dementia. ADC was characterized by a progressive loss of attention and concentration, marked motor problems (motor slowing and loss of balance), and behavioural changes, including apathy, agitation, and even mania. This disabling condition was a severe and progressive syndrome, usually resulting in death within the year (Navia et al., 1986).

In 1991, the first consensus set of criteria was introduced by the American Academy of Neurology (AAN), five years after ADC/HIV encephalopathy/HIV dementia syndrome was first described. The AAN criteria differentiated between severe HIV-associated dementia (HAD; which replaced ADC nomenclature) and a milder, less impaired form of impairment termed HIV-associated minor cognitive/motor disorder (MCMD).

With the introduction of ART in 1996, the course of HIV infection evolved from a largely fatal diagnosis to a chronic but manageable disorder (Collier et al., 1996; D'Aquila et al., 1996). Cases of HAD sharply declined, and milder forms of neurocognitive impairment became more evident as ART was increasingly effective in suppressing viral burden and improving immune status (Heaton et al., 2011). ADC/HAD-like phenotypes were associated with

subcortical motor symptoms and neuropathological changes such as HIV encephalitis including microglial nodules, multinucleated giants cell, and the presence of viral antigens, which became uncommon due to the effectiveness of ART (Heaton et al., 2010; Saylor et al., 2016). Given this dramatic change in the expression of neurocognitive impairment in HIV patients, the current consensus diagnostic guidelines, the "Frascati" criteria, were developed in 2007 and introduced HAND as an umbrella term to classify the range of symptoms as follows: asymptomatic neurocognitive impairment, mild neurocognitive disorders, and HAD (Antinori et al., 2007; **Figure 1**). To arrive at a HAND diagnosis, the Frascati criteria require multi-domain neuropsychological testing, an assessment of patients' functional impairment in daily life and the exclusion of comorbidities that confound the interpretation of the neuropsychological results (e.g., head injury, CNS co-infections, developmental delay, psychosis etc.). Asymptomatic neurocognitive impairment (ANI) is characterized by abnormal cognitive performance in at least two cognitive domains without functional impairment on a patient's day-to-day activities. Mild neurocognitive disorder (MND) requires cognitive deficits in two or more domains and at least mild functional impairment in activities of daily living and/or work. HAD, the most severe form of HAND, requires severe cognitive impairment with marked functional deficits.

The prevalence of HAND and the proportions of patients developing any form of HAND vary widely across cohorts. For example, diagnoses of HAND reported in cohort studies range from 26 to 74% (Bonnet, Amieva, Marquant, Bernard, Bruyand, Dauchy, Mercie, Greib, Richert, Neau, Catheline, Dehail, Dabis, Morlat, Dartigues, Chene, & ANRS CO3 Aquitaine Study Group, 2013; Cysique et al., 2014; Dufouil, Richert, Thiebaut, Bruyand, Amieva, Dauchy, Dartigues, Neau, Morlat, Dehail, Dabis, Bonnet, Chene, & ANRS CO3 Aquitaine Study Group, 2015; Garvey, Surendrakumar, & Winston, 2011; Sacktor et al., 2016; Simioni et al., 2010;

Vassallo et al., 2014; Winston et al., 2013; Wright et al., 2015). According to Saylor et al. (2016), most HAND patients receiving current ART have the milder forms, while only 2–8% of patients develop HAD.

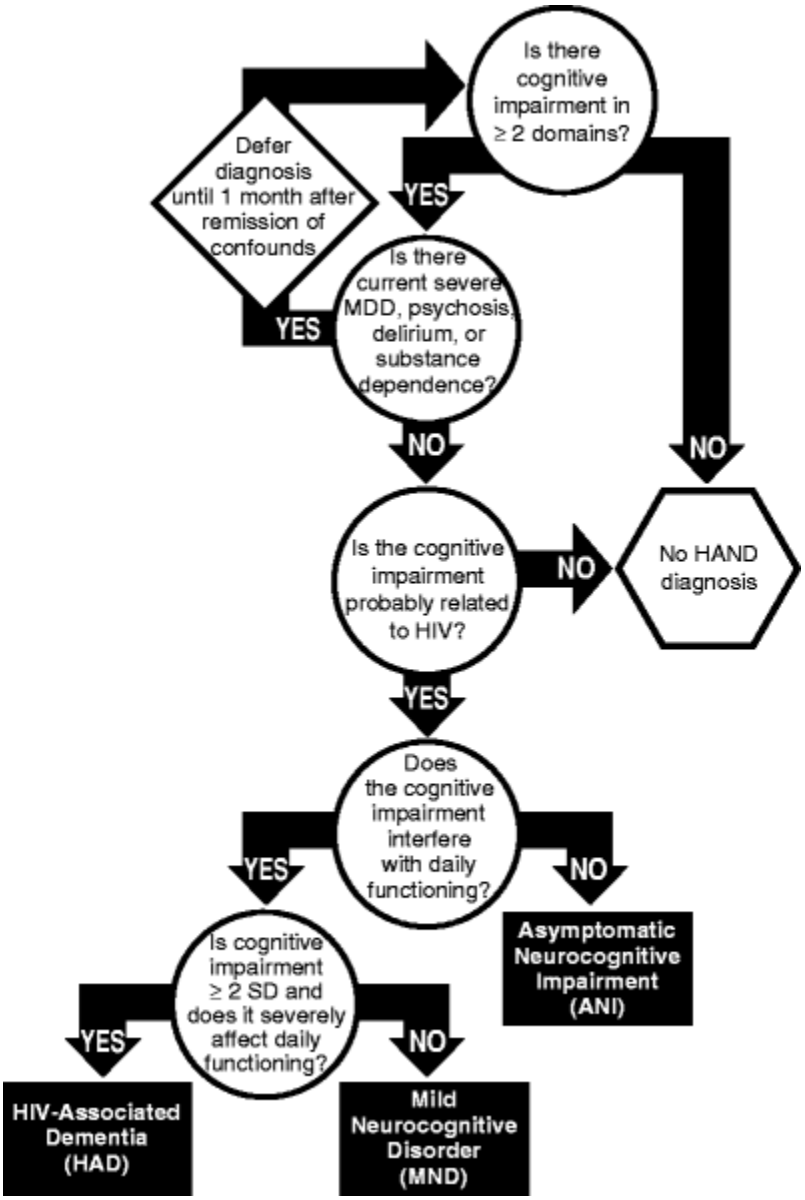

Figure 1. Decision-tree by Woods et al. (2009) for the diagnosis of HIV-associated neurocognitive disorders (HAND) according to the Frascati criteria.

2.4. Challenges with diagnosing NCI in HIV

Despite the update to the diagnostic guidelines to better capture the current presentation of NCI in HIV, there remain diagnostic ambiguities in the current consensus criteria that may account for such a wide range in the rates of impairment in current cohorts. Here, I describe how the multifactorial nature of NCI in HIV, and measurement problems make the use of the current diagnostic criteria challenging, and why neurocognitive profiling may be aid us in better assessing cognition in people living with HIV.

2.4.1. Multifactorial nature of NCI

Although the current diagnostic criteria presume that NCI observed in people living with HIV are due solely to HIV infection, several other factors may also predispose an individual to develop NCI (**Figure 2**; Nightingale et al., 2014). Neurocognitive deficits in people living with HIV have been consistently associated with severe immunosuppression. In the pre-ART era, current CD4+ counts, viral load, disease severity (AIDS), and duration of HIV diagnoses were highly predictive of NCI (Heaton et al., 2011). In the post-ART era, current immunosuppression and disease severity are no longer consistently associated with NCI due to the efficacy of current ART in stopping disease progression. However, low nadir CD4, or the lowest ever recorded CD4 count, remain highly predictive of NCI in HIV. This might be due to 'legacy' irreversible brain damage due to inflammation within the CNS before treatment (Ellis et al., 2011). Studies have also found that certain ART are neurotoxic (K. Robertson, Liner, & Meeker, 2012; Vivithanaporn, Asahchop, Acharjee, Baker, & Power, 2016), thus ART treatment itself may contribute to the development of NCI.

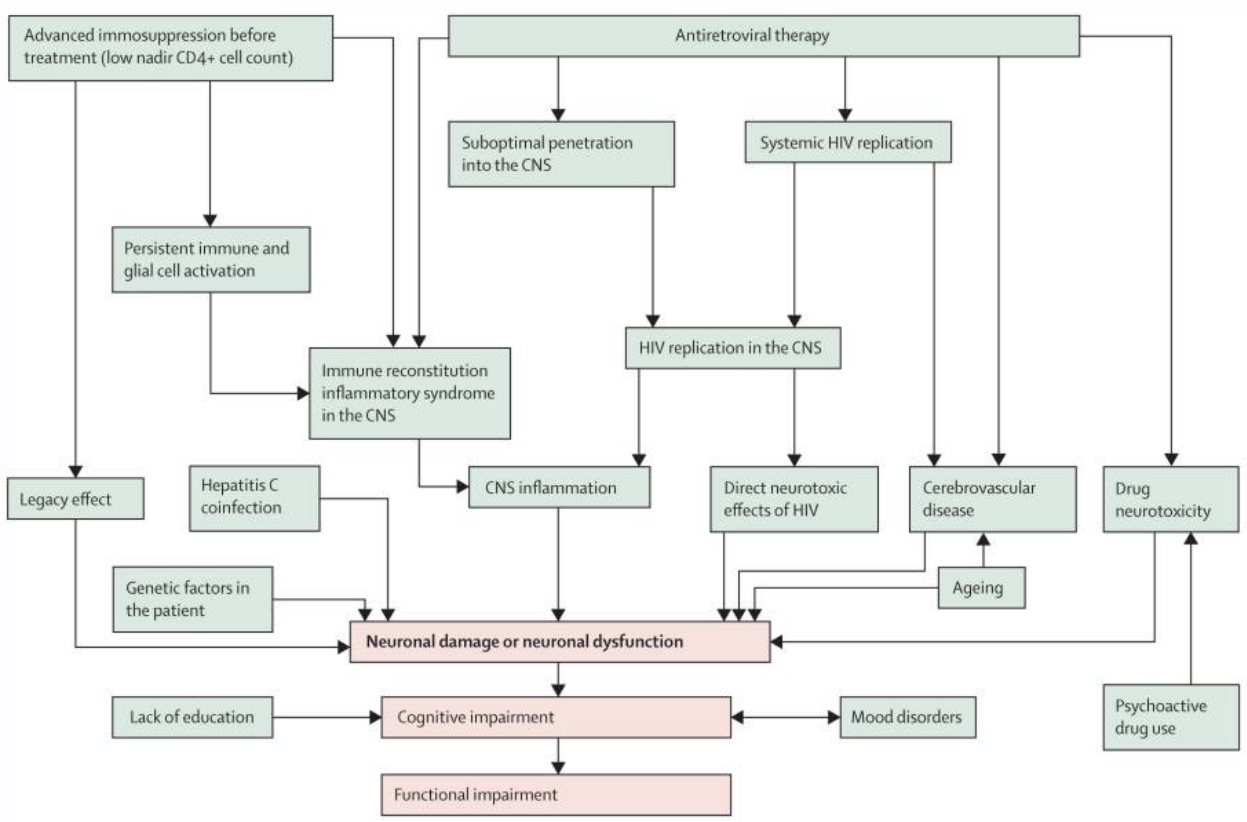

Figure 2. Pathological mechanisms underlying HIV-associated neurocognitive disorders according to Nightingale et al. (2014).

Demographic factors such as age and sex may also play a role. Female sex has been associated with an increased risk for neurocognitive deficits, particularly in HIV-positive women with depression, low education, and multiple life stressors (Maki & Martin-Thormeyer, 2009; Maki et al., 2015). Older age is independently associated with neurocognitive decline, however HIV infection may compound the cognitive changes in older people with HIV (K. R. Robertson et al., 2007). Individuals with HIV exhibit neurocognitive impairments comparable to deficits seen in healthy people who are older by a decade or more (Cohen, Seider, & Navia, 2015; Sheppard et al., 2017). People living with HIV tend to develop age-related comorbidities earlier,

including cardiovascular and metabolic diseases, which has been suggested to indicate an acceleration of biological ageing processes by HIV (Guaraldi et al., 2011).

Comorbidities themselves have also been identified as risk-factors of NCI in HIV, including cardiovascular comorbidities, diabetes, lipodystrophy and dyslipidemia (Heaton et al., 2010; Wright et al., 2015; Wright et al., 2010). Substance abuse and hepatitis C infection are also independently associated with NCI (Fletcher & McKeating, 2012; Gould, 2010), and HIV may compound their effects (Abutaleb, Kattakuzhy, Kottilil, O'Connor, & Wilson, 2018). Mental health conditions, such as depression have been found to covary with mildly lowered cognitive functions in HIV-infection (Bryant et al., 2015; Fellows, Byrd, & Morgello, 2013; Fialho, Pereira, Mendonca, & Ouakinin, 2013).

Genetic factors have also been identified as risk-factors for NCI in people living with HIV. I recently reviewed the possible contributions of host genetic diversity, epigenetic changes, and a genetic susceptibility to ART toxicity to the development NCI in HIV (Gomez, Power, and Fujiwara, 2018; see Appendix for summary tables). The *CCR5-Δ-32* allele codes for the C-C chemokine receptor type 5, which is the most common co-receptor for viral entry, and has been consistently associated with decreased disease progression in heterozygote individuals (Gonzalez et al., 1999; Ioannidis et al., 2001; Zimmerman et al., 1997). Since disease progression is associated with neurocognitive decline, *CCR5-Δ-32* allele may decrease an individuals' susceptibility for NCI. Factors associated with ageing/accelerated ageing also show promise. For example, HIV-positive older carriers of the ApoE ε4 allele (age ≥50 years) were at higher risk of developing NCI than HIV-positive non-carriers and HIV-uninfected ApoE ε4 carriers (Panos et al. 2013; Valcour et al. 2004; Mukerji et al. 2016; Wendelken et al. 2016). Emerging evidence also suggests that epigenetic aging (based on DNA methylation) in brain and blood tissue was is

faster in individuals living with HIV, which in turn has also negatively associated with neurocognitive performance (Horvath & Levine, 2015; Horvath et al., 2018; Rickabaugh et al., 2015). Despite these interesting findings, most genetic biomarker results await replication.

This brief summary highlighted the multiple factors implicated in the development of NCI in HIV, which complicates our ability to determine causality. The multifactorial nature of HIV-associated neurocognitive impairment may account for the highly variable rates of impairment across HIV-infected cohorts (Bonnet, Amieva, Marquant, Bernard, Bruyand, Dauchy, Mercie, Greib, Richert, Neau, Catheline, Dehail, Dabis, Morlat, Dartigues, Chene, & ANRS CO3 Aquitaine Cohort, 2013; Cysique, Maruff, & Brew, 2004; Heaton et al., 2008).

2.4.2. Measurement problems

According to the Frascati criteria (Antinori et al., 2007), neuropsychological assessment to determine HAND requires testing of at least five cognitive domains, assessed with at least two test measures within each domain. Ideally, the neuropsychological battery would assess performance in at least five of the following domains: language functions, attention/working memory, executive functions, memory, information processing speed, and motor functions. In practice however, the number of tested domains and the number of measures comprised within domains vary widely in the literature. For example, in some cohorts only one individual test measure has been declared as a 'domain', the same measure is counted into more than one domain, and as little as three domains were assessed to arrive at a diagnosis (Cysique et al., 2014; Dufouil, Richert, Thiebaut, Bruyand, Amieva, Dauchy, Dartigues, Neau, Morlat, Dehail, Dabis, Bonnet, Chene, & ANRS CO3 Aquitaine Study Group, 2015; Sacktor et al., 2016; Winston et al., 2013; Wright et al., 2015).

Furthermore, a HAND diagnosis also requires at least two domains to be "impaired",

however there are multiple scoring techniques that have been applied to arrive at this categorical label. In published cohorts, neurocognitive impairment has been established by using either 'domain averaging', such as the Gisslen criteria, (Gisslen, Price, & Nilsson, 2011); global deficit scores (Carey, Woods, Gonzalez, et al., 2004); or clinical ratings (Woods et al., 2004).These methods define 'cognitive impairment' differently. For example, domain averaging refers to raw scores being converted into z-scores based on normative reference cohorts, e.g., from test manuals. These z-scores are then averaged per domain. According to the most conservative domain averaging method, the Gisslen criteria, a domain is deemed impaired when the average z-score of a domain is $z \leq -1.5$ below normative ranges. A z-score of ≤ -1 is also sufficient to classify a domain as "impaired" according to Frascati. Clinical ratings convert raw scores into T-scores, which are then binned into clinical rating scores that range from 1 (above average, T-score ≥ 55, z-score ≥ 0.55) to 9 (definite impairment, T-score ≤ 19, z-score ≤ -3) with a cut off score of 5 (T-score: 35-39, z-scores between -1.5 and -1.0) representing impairment. Instead of domain averaging, a domain is deemed impaired when any *one* of the tests within a domain attains a score of 5 or above. Therefore, more weight is given to impairment in clinical ratings (Blackstone et al., 2012). In Global Deficits Scores, T-scores are converted into deficit scores ranging from zero (above average performance, T-score ≥ 40) to 5 (severe impairment, T-score ≤ 19). Domain deficit scores are then calculated by averaging all tests within a domain, and a global score is calculated by averaging all domain scores. A Global Deficit Score of ≥ 0.5 is deemed impaired. When compared to each other these criteria have low agreement (e.g., 18% concordance between HAND diagnoses across three criteria; Blackstone et al., 2012; Su et al., 2015; Tierney et al., 2017).Thus, there is a lack of consensus over which standard confers the best balance between specificity and sensitivity to NCI in HIV.

Lastly, assessment of functional impairment in daily life is needed for HAND staging but is challenging. The most commonly used self-report scales are subjective, were originally developed for other neurological disorders, and hence are not HIV-specific (Clifford & Ances, 2013). Even when more quantitative measures are used (e.g., the Columbia Medication Management Test; Heaton et al., 2004), both self-reported and performance-based measures are related to patients' educational and sociocultural background and do not predict progression of HAND and/or neurocognitive performance. These inconsistencies across diagnostic criteria pose a challenge for the classification and study of neurocognitive functions in HIV cohorts.

2.4.3. Alternative approaches to assessing NCI in HIV

As highlighted above, there are multiple problems with the current diagnostic criteria for NCI in HIV. The Frascati criteria presume that HIV is the main cause of neurocognitive deficits observed in patients; however, as outlined above there are multiple possible contributors to NCI in HIV apart from the disease processes themselves. Although the criteria suggest excluding confounding and consider comorbid conditions, in practice determining the extent to which factors other than HIV influence NCI is difficult. As discussed above, there are also measurement ambiguities inherent in the Frascati criteria. Lastly, the current criteria focus on a patients' overall degree of impairment without taking into account the *pattern* of impairment. The categorical nature of the HAND diagnosis fails to capture the heterogeneity of the NCI observed in patients. Given the multifactorial nature of the impairment in HIV, it is possible that there exist multiple types, or profiles, of cognitive impairment in HIV that are currently not being captured by the categorical label of HAND.

Therefore, the definition and measurement of NCI may benefit from the integration of empirical statistical methods to classify neurocognitive impairment profiles (Devlin & Giovannetti, 2017). These tools could be used to arrive at neurocognitive profiles reflective of NCI in HIV to help validate and potentially adjust the clinical staging. Only a handful of studies have attempted to identify cognitive subtypes within HIV-infected cohorts, but the results have been varied. Most studies have implemented cluster analysis, whereas a newer study conducted by Molsberry et al. (2018) applied a mixture-modeling technique known as Latent Profile Analysis, a technique that is arguably superior to cluster analysis. Results from these studies are outlined below.

Several studies using cluster-analytical techniques have examined neurocognitive profiles in HIV/AIDS (Devlin & Giovannetti, 2017). Untreated or monotherapy-treated individuals were included in some studies (Lojek & Bornstein, 2005; van Gorp et al., 1993) and patients receiving modern ART in others (Dawes et al., 2008; Fazeli et al., 2014). One study (Dawes et al., 2008) used relative neurocognitive strengths/weaknesses rather than the absolute performance used by others (Fazeli et al., 2014; Lojek & Bornstein, 2005; van Gorp et al., 1993). Apart from demographic characteristics (age, premorbid verbal abilities, and ethnicity), neurocognitive profiles were distinguished by symptomatic HIV-infection in untreated or monotherapy-treated cohorts (Lojek & Bornstein, 2005; van Gorp et al., 1993), which is uncommon in developed countries today. A more recent study (n = 78) identified two profiles (Fazeli et al., 2014), with neurocognitive impairment in one profile, linked to co-morbid risk factors. However, cluster analyses, especially in smaller cohorts, can result in idiosyncratic and unstable solutions (Aldenderfer & Blashfield, 1984).

Mixture-modeling techniques, like Latent Profile Analysis (Nylund, Asparouhov, & Muthen, 2007), are an alternative approach that offers rigorous statistical methods to determine the true number of profiles. LPA has been used to differentiate cognitive profiles in other neurological disease populations, including multiple sclerosis (Frndak et al., 2016), Alzheimer disease and mild cognitive impairment (Kohler et al., 2013; McGuinness, Barrett, McIlvenna, Passmore, & Shorter, 2015), Parkinson disease (Flensborg Damholdt, Shevlin, Borghammer, Larsen, & Ostergaard, 2012), and has recently been applied to investigate neurocognitive profiles in an HIV/AIDS cohort. Molsberry et al. (2018) uncovered three distinct neurocognitive profiles through LPA in a large sample of HIV-infected and uninfected men from the Multicenter AIDS Cohort Study. Profile 1 (20% of sample) consisted of participants with marked cognitive deficits, particularly in the motor domain. Profile 2 (35%) had normative performance in most cognitive domains but performed below average in learning and memory. Lastly, Profile 3 (45%) identified individuals with above average learning and memory abilities. There were few significant univariate differences among the profiles, mostly differentiating the globally impaired Profile 1 from the others: Compared to Profiles 2 and 3, individuals in Profile 1 were older, had lower IQ, showed higher levels of depression, and many were of non-White ethnicity. Predictive, multivariate models of profile membership were not tested. Notably, each profile contained roughly ~50% of HIV-infected and uninfected individuals, with similar rates of infection across profiles, similar rates of patients with AIDS, and similar levels of immune functioning/viral replication such that specificity of LPA-based cognitive profiles to HIV-infection remains to be tested.

As previously mentioned, neurocognitive impairment in contemporary HIV-cohorts is likely multifactorial, and requires statistical tools that accommodate many potentially related

susceptibility factors. Random Forest Analysis (Strobl, Malley, & Tutz, 2009) is a machine learning technique that represents such a tool that has yet to be applied to analyses of neurocognitive impairment in HIV-infection. RFA is frequently used on large biomedical data-sets (Touw et al., 2013), including in HIV/AIDS (Segal, Barbour, & Grant, 2004; Wade et al., 2015; Xu, Huang, Xu, & Zhang, 2007). RFA has also been used to identify neuroimaging correlates of cognitive/clinical subtypes of Alzheimer's disease (Gray et al., 2013) and multiple sclerosis (Mesaros et al., 2012), as well as identifying important demographic and lifestyle factors preserving memory in aging (McDermott, McFall, Andrews, Anstey, & Dixon, 2016). One advantage of using this statistical tool is that it can deal with highly dimensional data and can tolerate moderately correlated predictors. Details of RFA will be described in section 3.4 Statistical analysis.

2.5. Aims and hypotheses

In the current study, we aimed to address the following questions:

1) What is the nature and number of neurocognitive impairment profiles in ART-treated HIV-infected patients?

2) Which risk-factors are the most important determinants of membership in these profiles?

Based on the previous literature, the number and type of profiles were unclear, making this an exploratory study. However, we expected at least two different profiles of neurocognitive *impairment* (e.g., predominant motor impairment vs. higher-order, non-motor impairment). Among the prominent predictors of neurocognitively impaired profiles, we expected to observe nadir CD4 counts and variables associated with older age (age-related comorbidities like metabolic and cardiovascular conditions).

3. MATERIALS AND METHODS

3.1. Cohort

The study was approved by the University of Calgary, Conjoint Health Research Ethics Board; CHREB13-0615. Written informed consent was obtained from all participants. All participants were HIV-seropositive, in active treatment at the Southern Alberta HIV Clinic (SAC) in Calgary, Alberta, Canada. This is a multidisciplinary clinic providing free medical care laboratory testing and ART. It is staffed by multiple health care professionals, with services provided under universal healthcare (Asahchop et al., 2016; McCombe, Vivithanaporn, Gill, & Power, 2013). All participants had normal or corrected-to-normal vision and hearing, and fluency in English. Exclusions were head trauma with loss of consciousness exceeding five minutes, acute intoxication during neuropsychological testing, less than nine years of education, and severe/dominant neurological or psychiatric disorders (e.g., psychosis, severe stroke, opportunistic infections). Past or present substance use was no exclusion criterion unless sufficiently severe/protracted to preclude interpretation of current neurocognitive deficits. The absence of any predominant factors confounding the interpretation of neurocognitive performance was verified by repeated clinical assessment, chart review, cerebrospinal fluid assessments and cranial MRI to clarify any aetiologies suggestive of non-HIV-related CNS disease. From 445 participants, 64 were conservatively excluded due to confounding comorbidities (more details regarding the cohort will be discussed in the Results section).

3.2. Neuropsychological test battery

Neuropsychological testing was conducted at SAC (45-90 minutes) by a psychometrist, comprising tests of attention/processing speed, memory, motor, language, and executive

functions (**Table 1**). Ten scores from the following tests were included: Symbol Digit Modalities Test (SDMT; Smith, 1973); Number-Sequencing (TMT-2) and Number-Letter-Sequencing (TMT-4) from the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001); Grooved Pegboard (GPB; Trites, 1977); Hopkins Verbal Learning Test (HVLT-R; Brandt & Benedict, 2001); Verbal Fluency Tests (Phonemic Fluency: letters F-A-S; Semantic Fluency: animals) from the D-KEFS; and the Wisconsin Card Sorting Test (WCST-64; Kongs, Thompson, Iverson, & Heaton, 2000). Normative age- and education-adjustments from test manuals were applied to all tests (TMT-2/-4: Fine, Delis, & Holdnack, 2011), with additional adjustments for sex in the HVLT-R, D-KEFS Categorical Fluency, and GPB tests (Schretlen, Testa, & Pearlson, 2010). Raw scores were z-scaled based on normative adjustments. Performance in the Word Reading subtest of the Wide Range Achievement Test (WRAT-4; Wilkinson & Robertson, 2006) provided a measure for premorbid reading ability, and was included in a cognitive reserve score (Patel et al., 2013). Based on performance in this battery we also ascertained HAND diagnoses using the "Frascati" criteria (Antinori et al., 2007). HAND required performance of at least 1 standard deviation below normative means in two or more of the five neurocognitive domains (**Table 1**). The presence of functional impairment in daily life was ascertained through patient report involving multiple steps: each patient was asked at the time of intake by a registered nurse whether they were experiencing difficulties with memory, related cognitive abilities or activities, and if these deficits impacted their lives on a day-to-day basis. Answers to these inquiries were recorded in the general medical chart. At the time of neuropsychological testing, each patient was asked again by a trained psychometrist whether they experienced recent difficulties with memory, concentration, decision making, motor functions and behavioral troubles (agitation, anger, apathy), and how it affected their lives. Each

patient seen by neurology was then asked if they experienced difficulties with short- or long-term memory, concentration, decision making, anger/irritability, apathy, difficulty with speed of thinking, moving or tremor as well as how these concerns might be evident in everyday life, probing patients to provide specific examples. Furthermore, the presence of functional impairment was confirmed by informant-report through accompanying family members/caregivers when available. Final HAND diagnoses were determined by consensus decision of the clinical team.

Table 1. Neuropsychological test descriptions and mean z-scaled performance (standard deviations) for patients from the Southern Alberta Clinic (SAC) cohort.

Neuropsychological Tests	SAC cohort (n = 381)	Normative Adjustment
Attention		
SDMT: correct responses	-.06 (1.05)	Age, education
TMT-2: completion time	-.38 (0.98)	Age, education
Executive Functions		
TMT-4: completion time	-.50 (1.04)	Age, education
WCST-64: perseverative errors	-.56 (0.88)	Age, education
Memory		
HVLT-R: immediate recall	-.05 (1.16)	Age, education, sex
HVLT-R: delayed recall (25 min.)	-.25 (1.09)	Age, education, sex
Language		
D-KEFS (FAS): correct words	-.56 (0.95)	Age, education, sex
D-KEFS (Animals): correct words	.04 (1.09)	Age, education, sex
Motor Functions		
GPB (dom. hand): completion time	-.06 (1.07)	Age, education, sex
GPB (non-dom. hand): completion time	.01 (1.13)	Age, education, sex

Abbreviations: D-KEFS: Delis-Kaplan Executive Function System; dom.: dominant; GPB: Grooved Pegboard; HVLT-R: Hopkins Verbal Memory Test – Revised; SDMT: Symbol Digit Modalities Test; TMT: Trail Making Test; WCST: Wisconsin Card Sorting Task, 64 card version.

3.3. Demographic and medical variables

Several continuous and dichotomous variables (0 = no; 1 = yes) with potential neurocognitive performance impact were included. Continuous demographic variables were age, years of education and WRAT-4. Since higher levels of education and reading abilities both suggest higher cognitive reserve, a cognitive reserve score was calculated (Patel et al., 2013) by combining WRAT- 4 scores with completed years of education and dichotomized (0 = low, 1 = high) via median split. Other dichotomous demographic variables were sex (male = 0, female = 1), place of birth outside North America (Canada/U.S.A., irrespective of ethnicity), and employment. Numerical HIV- and ART-related variables were: Recent and nadir CD4 T-cell counts, recent and peak plasma HIV viral load, years since HIV diagnosis, CNS Penetrance Effectiveness (CPE) rank (Letendre et al., 2008) of current ART regimens (untreated = 0). Dichotomous HIV-related variables included: on ART, ART non-adherence within the past five days, and ART side-effects (including ART-toxicity). Regarding comorbidities, two numerical self-report measures were included: Depressive symptoms in the Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001) and health-related quality of life (QoL; Crane et al., 2006). Dichotomous comorbidity variables were: HCV co-infection, cardiovascular conditions (hypertension, peripheral vascular diseases, myocardial ischemia or infarction), diabetes, metabolic conditions (dyslipidemia and lipodystrophy), self-reported current substance use (alcohol >9 drinks/sitting, current use of marijuana, crack/cocaine, or any other illicit drug), and psychiatric diagnoses (mood and affective disorders). Furthermore, toxoplasma serostatus was included as a variable of interest. Latent toxoplasmosis has been suggested to affect brain levels of vitamin B-12 and folate and has been associated with neurocognitive functions in previous population-based studies (Berrett, Gale, Erickson, Brown, & Hedges, 2017), as well as in HIV studies (Basavaraju, 2016; Ene et al., 2016).

3.4. Statistical analyses

Analyses followed three major steps. First, empiric neurocognitive profiles were identified by LPA and compared to HAND diagnoses. Secondly, characteristics of individuals with these profiles were compared with univariate methods. Third, profile characteristics were compared with Random Forest Analyses (RFA).

LPA is a mixture-modeling technique that identifies homogenous subgroups (Nylund et al., 2007). Superior to cluster analysis, multiple rigorous statistical tests and model fit statistics identify the true number of profiles (Nylund et al., 2007). Indicator variables were the ten neurocognitive test scores. The optimal number of empirically derived profiles was based on significant Vuong-Lo-Mendell-Rubin Test (Vuong, 1989) and Bootstrapped Likelihood Ratio Test (McLachlan & Peel, 2000). Akaike's Information Criterion and Bayesian Information Criterion (Nylund et al., 2007) aided with model selection. LPA was carried out with Mplus 7.4 (Muthen & Muthen, 1998-2015). Univariate comparisons of LPA profiles were conducted using ANOVA, Kruskall-Wallis tests, or χ^2-tests as indicated. LPA profiles and HAND diagnoses were compared with χ^2-tests.

Multivariate comparison between profiles was conducted by Random Forest Analysis (RFA), a machine learning technique based on the combination of predictions of multiple non-parametric classification and regression trees (Breiman, 2001). In the current application, the goal of the RFA was to identify variables ('predictors') that successfully separated the cohort into one versus another LPA profile. Unlike in simple decision trees, the search process for identifying these variables is iterative, using an ideal subset of variables, and it is repeated multiple times (i.e., creating a 'forest' rather than a single tree). The results from all trees in the forest are then aggregated and compared, and each of the contributing predictors is ranked

according to its relative importance in correctly splitting the data into two groups (i.e., LPA profiles here). Two thirds of the original data (the training or 'bootstrap' sample) are used to develop the decision tree forest, which is then tested on the remaining cases (the test or 'out-of-bag' sample). Compared to simple decision trees, RFA solutions avoid overfitting, are more stable, and have higher prediction accuracy (Strobl et al., 2009). Another advantage of RFA is its usefulness in situations with smaller samples and large numbers of potential predictors, compared to requirements for simple decision trees (Breiman, 2001; Strobl et al., 2009). Due to statistical requirements for RFA (Kuhn & Johnson, 2013), variables with less than 10% of cases in the smaller of two compared profiles were excluded. For example, HCV-co-infection was excluded since the cohort included only few individuals with an HCV co-infection, ranging from 3.1% to 6.7% within each profile. For the same reason, ART non-adherence, and ART treatment status were eliminated from all comparisons. In order to retain additional categorical predictors that would have been excluded based on this criterion (lipodystrophy), we combined lipodystrophy with dyslipidemia to form one dichotomous variable ('metabolic conditions'). AIDS status duplicates information carried by nadir CD4 and was therefore excluded. Similarly, recent viral load replicated information in detectable viral load and was also excluded. Diabetes had to be excluded as a predictor in the **P1** vs. **P2** comparison due to low (< 10%) occurrences in these profiles. Thus, each of the three initial RFA included the maximum number of permissible variables, with slight differences between comparisons based on these statistical exclusion rules (**Table 2**).

Missing data were negligible (maximum 5.8%, for WRAT-4, included in cognitive reserve), and imputed by the missForest package in R (Stekhoven & Bühlmann, 2011). Profile sizes were balanced by Synthetic Minority Over-sampling Technique (SMOTE; Chawla,

Bowyer, Hall, & Kegelmeyer, 2002), implemented in the DmwR package. Imbalanced class sizes biases results of classifiers towards the majority class, such that predictive models obtain very high classification accuracies (AUC) despite erroneously classifying all cases into the majority class (Torgo, Ribeiro, Pfahringer, & Branco, 2013). This problem is illustrated in **Table 3**, resulting in low classification accuracies for the smaller **P3** group. To avoid this problem, SMOTE over-samples the minority class (**P3** cases here) by creating synthetic cases based on k-nearest neighbour classification and bootstrapping. To retain the maximum number of original cases and the largest possible sample size, we increased the size of **P3** to match the sizes of **P1** and **P2**, as was done in other studies (Hariharan, Tirodkar, Porwal, Bhattacharya, & Joly, 2017; McDermott et al., 2016). We did not implement SMOTE to compare **P1** and **P2** given that these profiles had approximately the same number of individuals. SMOTE parameters used to upsample (*perc.over*) and donwsample (*perc.under*) classes to balance the data for each RFA comparison in the entire cohort are shown in **Table 3**.

RFA was conducted using the cforest function from the Party package (Hothorn, Buhlmann, Dudoit, Molinaro, & van der Laan, 2006) in R (R Core Team, 2017). Forests consisted of 5000 trees and at each potential split we evaluated a random sample of two predictor variables, identified as yielding optimal prediction accuracy. Each variable's importance was evaluated by conditional permutation accuracy importance (Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008). The importance values of irrelevant variables hover around zero, therefore variables with negative, zero or small positive importance values are considered unimportant to distinguish profiles, whereas variables with larger positive values are considered important (Strobl et al., 2008; Strobl et al., 2009). Variable importance scores vary depending on the included predictors, such that by themselves, the numerical values of variable importance scores

are not meaningful and cannot provide absolute thresholds. Instead, as suggested by (Strobl et al., 2009), the threshold for variable importance was set by taking the absolute (i.e., positive) value of the lowest variable importance score obtained from each RFA. Predictors with scores lower than this value are considered unimportant, while predictors with scores higher than this value are considered important (indicated by a dotted line in the figures). Model strength was evaluated by area under the receiver operating curve (AUC) and F_1 scores. Although RFA is a robust tool to identify relevant predictors of membership, directionality of the predictors cannot be obtained from this analysis alone, as predictors may appear at different locations, and different cut-off scores for membership, across the trees. Therefore, the directionality of important predictors was determined by univariate analysis.

Table 2. List of variables considered for inclusion in the Random Forest Analyses (x = included)

Variable	P1 vs. P2	P1 vs. P3	P2 vs. P3
Age	X	X	X
Sex	X	X	X
Continent of Origin	X	X	X
Employment	X	X	X
Cognitive Reserve	X	X	X
Quality of Life	X	X	X
Duration of HIV Diagnosis	X	X	X
Recent VL	-	-	-
Detectable VL	X	X	X
Peak VL	X	X	X
Recent CD4 T-cell count	X	X	X
Nadir CD4 T-cell count	X	X	X
AIDS	-	-	-
On ART at Test Date	-	-	-
ART non-adherence	-	-	-
ART Side-effects	X	X	X
CPE	X	X	X
Alcohol (binge)	-	-	-
Marijuana Use	X	X	X
Crack/cocaine	-	-	-
Other illicit drugs	-	-	-
HCV Co-infection	-	-	-
Toxoplasma Seropositivity	X	X	X
Metabolic Conditions	X	X	X
Diabetes	-	X	X
Cardiac Conditions	X	X	X
Psychiatric Disorders	X	X	X
Domestic Violence	X	X	X
PHQ-9	X	X	X

Abbreviations. AIDS: acquired immune deficiency syndrome; ART: antiretroviral therapy; CPE: CNS-penetrance effectiveness rank; HCV: hepatitis C virus; PHQ-9: Patient Health Questionnaire-9; VL: viral load.

Table 3. Synthetic Minority Over-sampling Technique (SMOTE) parameters used to balance the size of P3 (entire cohort; N = 381)

Model	**P1 vs P3**		**P2 vs P3**	
Unbalanced	**P1**	**P3**	**P2**	**P3**
n	159	59	163	59
Classification %	100%	5.1%	100%	0%
Balanced				
perc.under:perc.over	*135*	*200*	*138.5*	*200*
n	159	177	163	177
Classification %	78.6%	83.6%	69.9%	74.6%

Abbreviations. perc.over: upsampling parameter; perc.under: downsampling parameter).

4. RESULTS

4.1. Latent profile analysis

The cohort (n = 381) consisted of 88% males, predominantly born in Canada/U.S.A (78.7%, of which 4.3% identified as Indigenous), with a mean age of 47.3±11.1 years. At test date, 94% were actively receiving ART, although 8.7% reported non-adherence within the last five days before testing. We tested two-, three-, and four-profile LPA models to define neurocognitive profiles. The three-profile model showed the best fit (**Table 4**). The first neurocognitive profile (**P1**) included 159 participants (41.7% of the cohort), the second profile (**P2**) had 163 participants (42.8%), and the third profile (**P3**) consisted of 59 individuals (15.5%). Neurocognitive impairment was present in **P3** only, indicated by z-scores below the normative mean in all test (z-score < -1), except D-KEFS Animals and GPB (**Table 5**; **Figure 3A**).

The profiles differed in mean performance ($F(2, 372) = 478.6$, $p < 0.001$; **P1**> **P2** > **P3**; **Fig. 3A**; **Table 5**). **P3** performed significantly lower than **P1** in all tests (p's < 0.05, Bonferroni-corrected). **P2** performed significantly lower than **P1** in all but simple attention (TMT-2: $p = 0.84$) and non-dominant motor functions (GPB; $p = 0.12$), and significantly higher than **P3** in all but HVLT-R delayed recall ($p = 0.12$). Within profiles, test performance also differed ($F = 19.0$, $p < 0.001$): Compared to performance in the other tests, **P1** had higher verbal memory performance (immediate HVLT-R; p's < 0.05; **Fig. 3A**; **Table 5**). **P2** had lower executive functions (WCST and FAS) and verbal memory (HVLT-R; p's < 0.05), and **P3** had lower TMT scores (p's < 0.05). Thus, LPA identified three distinct neurocognitive performance profiles, one (**P3**) with neurocognitive impairment.

Table 4. Latent Profile Analyses on the ten neurocognitive indicators

Number of Profiles	AIC	BIC	LMR	BLRT
Entire cohort (n = 381)				
2	10417.3	10539.5	<.001	<.001
3	10226.3	10391.9	<.001	<.001
4	10144.7	10353.6	.13	.14
North American-born patients (n = 300)				
2	8024.3	8139.1	<.001	<.001
3	7866.9	8022.4	.01	.01
4	7804.4	8000.7	.28	.28

Abbreviations: AIC: Akaike's Information Criterion; BIC: Bayesian Information Criterion; BLRT: Bootstrapped Likelihood Ratio Test (*p*-value); LMR: Vo-Lo-Mendell-Rubin Likelihood Ratio Test (*p*-value).

Table 5. Mean z-scaled performance (standard deviations) in the ten neuropsychological indicators used in Latent Profile Analyses in the entire cohort (n = 381 patients).

Test	P1 (n=159)	P2 (n=163)	P3 (n=59)
SDMT[a,b,c]	0.47 (0.84)	-0.08 (0.83)	-1.40 (0.83)
TMT-2[b,c]	-0.06 (0.76)	-0.17 (0.69)	-1.86 (0.90)
TMT-4[a,b,c]	-0.02 (0.65)	-0.33 (0.73)	-2.27 (0.76)
WCST p-err.[a,b,c]	-0.30 (0.86)	-0.61 (0.83)	-1.18 (0.73)
D-KEFS FAS[a,b,c]	-0.19 (0.95)	-0.64 (0.83)	-1.31 (0.83)
D-KEFS Animals[a,b,c]	0.63 (0.88)	-0.11 (0.90)	-0.99 (1.05)
HVLT-R imm.[a,b,c]	0.97 (0.76)	-0.65 (0.70)	-1.23 (0.75)
HVLT-R del.[a,b]	0.62 (0.83)	-0.82 (0.72)	-1.09 (0.88)
GPB dom.[a,b,c]	0.28 (1.05)	-0.07 (0.89)	-1.00 (1.04)
GPB non-dom.[b,c]	0.31 (1.17)	0.06 (0.93)	-0.96 (1.07)

Abbreviations. Dom. = dominant hand; D-KEFS FAS = Delis-Kaplan Executive Functions Systems letter fluency task; D-KEFS Animals = Delis-Kaplan Executive Functions Systems category fluency task; GPB = grooved pegboard; HVLT del. = Hopkins Verbal Learning Test-Revised delayed recall; HVLT imm. = Hopkins Verbal Learning Test immediate recall; Non-dom. = non-dominant hand; SDMT = Symbol Digit Modalities Test; TMT = Trail Making Task; WCST-pers.err. = Wisconsin Card Sorting Task perseverative errors.
[a]: **P1** > **P2** ($p < .05$). [b]: **P1** > **P3** ($p < .05$). [c]: **P2** > **P3** ($p < .05$).

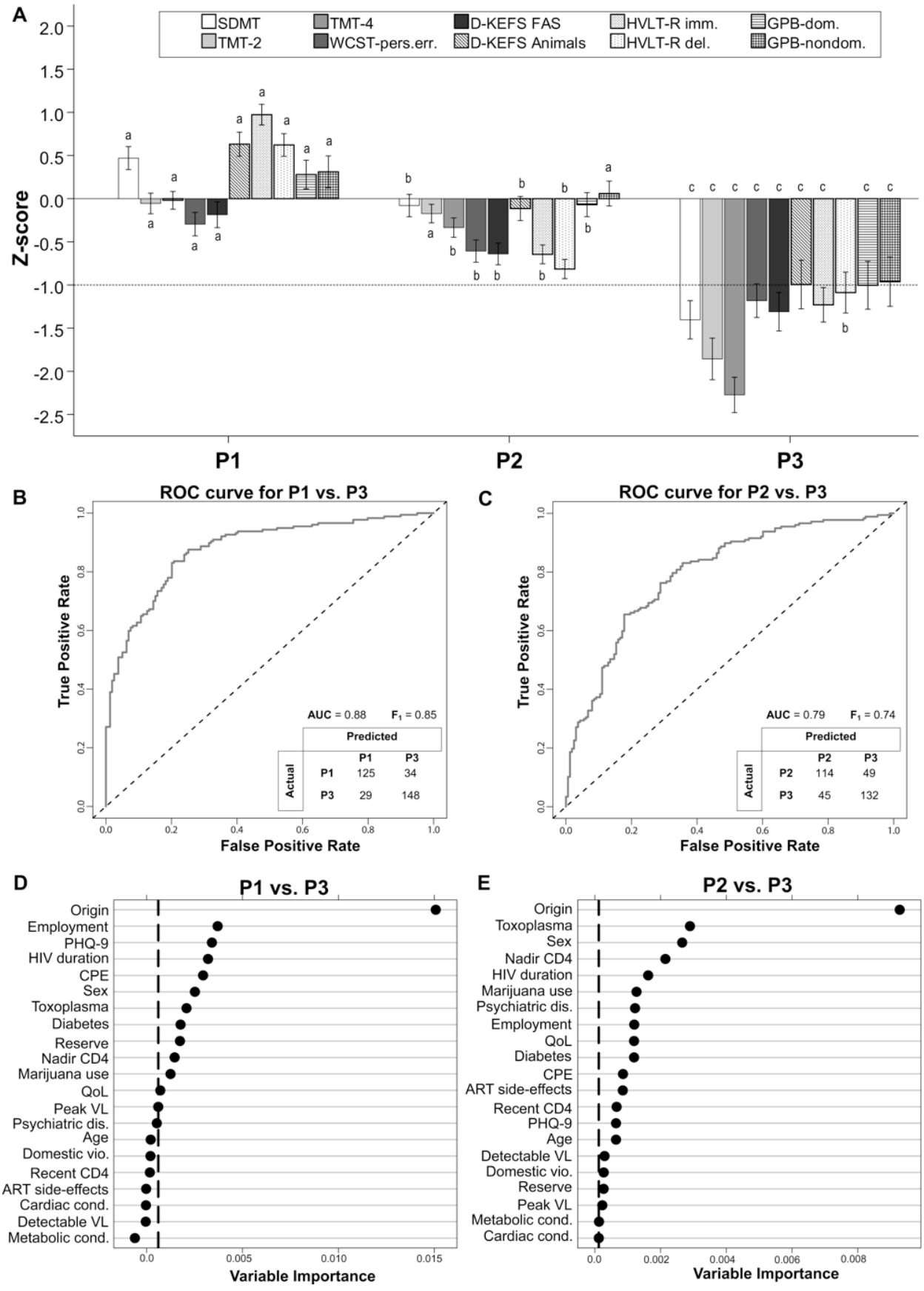
A
SDMT
TMT-2
TMT-4
WCST-pers.err.
D-KEFS FAS
D-KEFS Animals
HVLT-R imm.
HVLT-R del.
GPB-dom.
GPB-nondom.
Z-score
P1
P2
P3
B
ROC curve for P1 vs. P3
True Positive Rate
False Positive Rate
AUC = 0.88
F1 = 0.85
Predicted
Actual
P1 125 34
P3 29 148
C
ROC curve for P2 vs. P3
AUC = 0.79
F1 = 0.74
P2 114 49
P3 45 132
D
P1 vs. P3
Origin
Employment
PHQ-9
HIV duration
CPE
Sex
Toxoplasma
Diabetes
Reserve
Nadir CD4
Marijuana use
QoL
Peak VL
Psychiatric dis.
Age
Domestic vio.
Recent CD4
ART side-effects
Cardiac cond.
Detectable VL
Metabolic cond.
Variable Importance
E
P2 vs. P3
Origin
Toxoplasma
Sex
Nadir CD4
HIV duration
Marijuana use
Psychiatric dis.
Employment
QoL
Diabetes
CPE
ART side-effects
Recent CD4
PHQ-9
Age
Detectable VL
Domestic vio.
Reserve
Peak VL
Metabolic cond.
Cardiac cond.
Variable Importance

Figure 3. Latent Profile Analyses and Random Forest Analyses of 381 HIV-infected participants from the Southern Alberta HIV Clinic. A) Distribution of test performance of participants across three latent neurocognitive profiles derived from Latent Profile Analysis B) & C) Receiver Operating Characteristic curve for RFA results. D) & E) Random Forest Analyses (RFA), showing the importance for specific variables distinguishing **P1**/**P3**, and **P2**/**P3**, respectively. The dotted black line represents variable importance cut-off; dots to the right of the line indicate important variables.
a,b,c: different superscripts denote differences between groups (Bonferroni post-hoc's: $p < .05$).
Abbreviations: ART = antiretroviral therapy; AUC = area under the receiver operating curve; Cond. = condition; CPE = CNS penetrance effectiveness rank; D-KEFS FAS = Delis-Kaplan Executive Functions Systems letter fluency task; D-KEFS Animals = Delis-Kaplan Executive Functions Systems category fluency task; Dis. = disorder; Dom. = dominant hand; GPB = grooved pegboard; HVLT-R del. = Hopkins Verbal Learning Test delayed recall; HVLT-R imm. = Hopkins Verbal Learning Test immediate recall; Non-dom. = non-dominant hand; Origin = continent of birth; PHQ-9 = Patient Health Questionnaire-9; QoL = health-related quality of life; ROC = receiver operating characteristic; Reserve = cognitive reserve; SDMT = Symbol Digit Modalities Test; TMT = Trail Making Task; Toxoplasma = toxoplasma seropositivity; Vio. = violence; VL = viral load. WCST-pers.err. = Wisconsin Card Sorting Task perseverative errors

4.2. Univariate differences between profiles and comparisons with HAND diagnosis

Univariate differences (**Table 6**) indicated that **P3** included more toxoplasma seropositive female participants born outside North America, compared to **P1** and **P2**. Marijuana use was significantly higher in **P2** compared to **P3.** The **P1** profile had a significantly lower CPE rank than **P2.** Trends showed non-significantly lower nadir CD4 T-cell counts and higher rates of AIDS in **P3** than in **P1/P2**.

Of the 381 participants, 81 (21.3%) were diagnosed with HAND. Of these, 31 had asymptomatic neurocognitive impairment (ANI), 42 had mild neurocognitive disorder (MND), and 8 had HIV-associated dementia (HAD). The remaining 300 patients had no neurocognitive deficits (NN). The proportions of HAND diagnoses within and across LPA profiles showed large differences ($\chi^2(2) = 219.4$, $p < 0.001$; Bonferroni post-hoc tests: p's < .05). The HAND diagnosis was more frequent in **P3** (89.8%), followed by **P2** (16.6%), and least frequent in **P1**, with only one

HAND-ANI participant. Symptomatic HAND (MND and HAD) was substantially more frequent in **P3** (76.3%) than in **P2** (8.6%). Conversely, the **P1** and **P2** profiles were mostly comprised of NN patients (**P1** = 99.3%; **P2** = 83.4%).

Table 6. Patient characteristics and differences between LPA profiles (n = 381); data are means (standard deviations) or percentages within profiles, unless stated otherwise

Variable	P1 (n = 159)	P2 (n = 163)	P3 (n = 59)	Test Statistic	p-value
Age (years)	46.3 (11.1)	48.2 (11.0)	47.3 (11.4)	$F(2, 297) = 1.07$	.34
Sex (male)	91.8%[a]	87.7%[a,b]	78.0%[b]	$\chi^2(2) = 7.80$	**.02**
Place of Origin (North America)	88.7%[a]	81.0%[a]	45.8%[b]	$\chi^2(2) = 48.20$	**<.001**
Employment (currently employed)	75.5%	70.6%	59.3%	$\chi^2(2) = 5.45$	.07
Cognitive Reserve (% high)	59.9%	54.2%	43.6%	$\chi^2(2) = 4.37$	.11
QoL (1 = poor, 5 = excellent) †	4 (1-5)	4 (1-5)	4 (1-5)	$\chi^2(2) = 1.63$	.44
HIV duration (in years)	11.4 (8.4)	11.3 (8.0)	12.4 (8.4)	$F(2, 297) = 0.42$	.66
Recent VL copies/ml (log) †	1.60 (1.60-5.44)	1.60 (1.60-5.18)	1.60 (1.60-4.95)	$\chi^2(2) = 1.57$	.46
Detectable VL	11.9%	14.7%	18.6%	$\chi^2(2) = 1.66$	.44
Peak VL copies/ml (log) †	4.87 (1.60-6.81)	4.79 (1.60-6.78)	4.93 (1.60-6.98)	$\chi^2(2) = 0.35$	.84
Recent CD4 T-cell (count/mm³)	583.3 (241.5)	586.9 (295.2)	538.5 (241.2)	$F(2, 297) = 0.78$	.46
Nadir CD4 T-cell (count/mm³)	230.0 (178.4)	223.0 (168.5)	174.3 (152.0)	$F(2, 297) = 2.40$	.09
AIDS (nadir < 200 count/mm³)	46.5%	43.6%	61.0%	$\chi^2(2) = 5.40$	.07
On ART at test date	92.5%	94.5%	96.6%	$\chi^2(2) = 1.45$	.49
ART non-adherence (last five days)	6.8%	11.0%	7.0%	$\chi^2(2) = 1.90$	.39
ART side-effects	14.5%	17.2%	15.3%	$\chi^2(2) = 0.46$	.80
CPE †	7 (0-11)[a]	7 (0-13)[b]	7 (0-11)[a,b]	$\chi^2(2) = 7.27$	**.03**
Substance use					
Alcohol (binge)	3.8%	6.1%	0%	$\chi^2(2) = 4.18$	.12
Marijuana	27.7%[a,b]	31.3%[b]	13.6%[a]	$\chi^2(2) = 6.96$	**.03**
Crack/cocaine	7.5%	3.7%	3.4%	$\chi^2(2) = 2.90$	.23
Other illicit drugs	3.1%	2.5%	1.7%	$\chi^2(2) = 0.39$	.82
HCV co-infection	3.1%	6.7%	3.4%	$\chi^2(2) = 2.60$	.27
Toxoplasma seropositive	7.5%[a]	9.2%[a]	25.4%[b]	$\chi^2(2) = 14.98$	**<.001**
Dyslipidemia	25.2%	28.2%	18.6%	$\chi^2(2) = 2.11$	.35
Lipodystrophy	8.2%	9.2%	8.5%	$\chi^2(2) = 0.11$	.95
Diabetes	5.0%	8.6%	13.6%	$\chi^2(2) = 4.52$	.11

Variable	P1 (n = 159)	P2 (n = 163)	P3 (n = 59)	Test Statistic	p-value
Cardiac conditions	12.6%	16.6%	15.3%	$\chi^2(2) = 1.04$	.60
Psychiatric disorders	28.9%	29.4%	32.2%	$\chi^2(2) = 0.23$	.89
Domestic violence	24.5%	29.4%	27.1%	$\chi^2(2) = 0.99$	.61
PHQ-9 †	4 (0-26)	5 (0-26)	5.5 (0-23)	$\chi^2(2) = 3.07$	.22

Abbreviations. ART: combination antiretroviral therapy; CPE: antiretroviral central nervous system penetration-effectiveness score (individuals not on ART received a CPE rank of zero); HCV: hepatitis C virus; PHQ-9: Patient Health Questionnaire-9; QoL: quality of life; VL: viral load.

[a,b,c]: different superscripts denote differences between groups (Bonferroni post-hoc's: $p < .05$

†: Medians (range). PHQ-9 available for n=367.

4.3. Random forest analyses

The importance of variables associated with profile membership was tested in a multivariate framework by three RFAs: Comparing **P1** with **P2**, **P1** with **P3**, and **P2** with **P3**. Directionality of important variables was determined by the univariate results in **Table 6**. RFA between **P1** and **P2** showed poor prediction performance (AUC = 0.46, F_1 = 0.46; not shown). Fits were substantially better for RFA models involving **P3** (**Figures 3B and 3C**). The most important predictor of **P3** was place of birth ('Origin' in **Figures 3D and 3E**). Other important variables in both RFAs were lower employment rates, longer HIV duration, female sex, toxoplasma seropositivity, diabetes, low nadir CD4 and lower marijuana use. Depressive symptoms (PHQ-9), higher CPE, and lower cognitive reserve differentiated **P3** from **P1** (**Fig. 3D**). Psychiatric disorders and low QoL further distinguished **P3** from **P2** (**Fig. 3E**). Several additional variables trended towards the statistical threshold and were not further considered. **Figure 4** summarizes only the most important variables and their corresponding directionality.

To verify the RFA outcomes with a conventional method, we performed separate logistic regressions. Keeping in mind that this method is not suited to assess the large number of partly collinear predictors, only the five most important variables identified by RFA were analysed. Multicollinearity between selected predictors was assessed via variable inflation factor (VIF > 3) and tolerance (< 0.20; Hair, 2006). For the **P1** vs. **P3** regression, variables included: place of origin, employment status, PHQ-9, HIV duration, and CPE. For the **P2** vs. **P3** regression, variables included were: place of origin, toxoplasma seropositivity, sex, nadir CD4, and HIV duration. Both logistic regression models were significant (**P1**/**P3**: $\chi^2 = 55.8$, df = 5; **P2**/**P3**: $\chi^2 = 39.95$, df = 5; *p*'s < .001) and none of the predictors showed VIF or tolerance scores indicative of multicollinearity (**P1**/**P3**: $VIF_{max} = 1.15$, $tolerance_{max} = 0.87$; **P2**/**P3**: $VIF_{max} = 1.34$, $tolerance_{max}$

= 0.74). In the **P1**/**P3** model, only non-North American origin (OR = 14.2, [6.3, 32.1], $p < .001$) and higher CPE (OR = 1.3, [1.0, 1.6], $p = .02$) predicted membership into **P3**. In the **P2**/**P3** model, only non-North American origin (OR = 5.3, [2.4, 11.6], $p < .001$) predicted membership into **P3**. There was also a trend towards longer HIV duration ($p = .06$) predicting **P3** membership. Thus, outcomes replicated the essential RFA findings, albeit with reduced sensitivity.

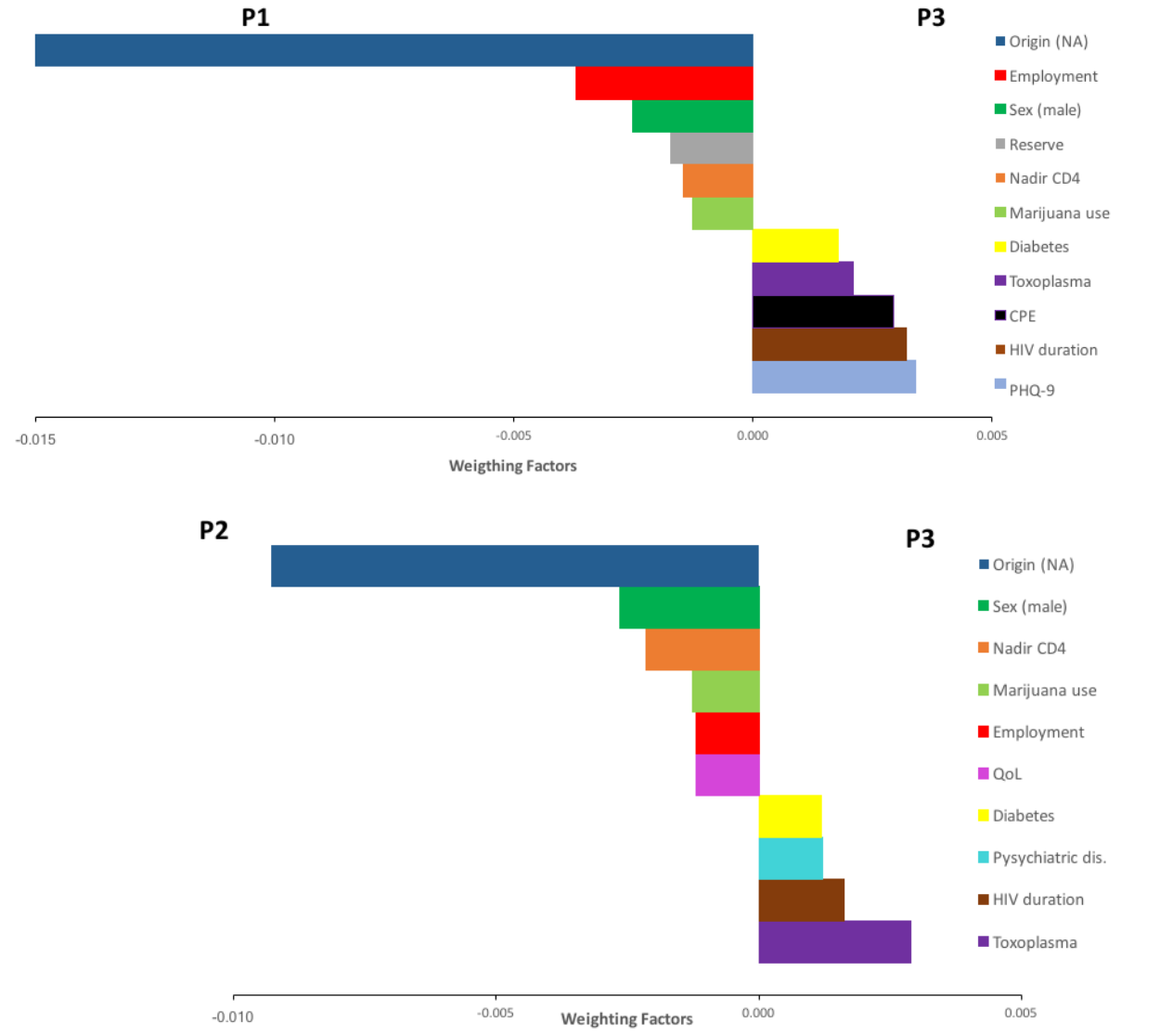

Figure 4. Summary of only important variables predicting neurocognitive performance in the Southern Alberta Clinic cohort (N= 381) identified by Random Forest Analysis. Length of bar represents relative importance.
Refer to Figure 3 for abbreviations.

The results above were dominated by demographic variables, emphasizing previous cluster-analytical findings (Devlin & Giovannetti, 2017) and Molsberry et al. (2018) findings, placing many individuals with non-North American origin into the impaired **P3** profile (**Table 6**, **Fig 4**). Instead of important predictors of cognitive profiles, the previous results may therefore be identifying predictors of where participants were born (*origin)*, mirroring the demographic differences between **P1**/**P2** and **P3**. Thus, in an attempt to isolate effects of birth continent, we conducted an additional RFA identifying important predictors of non-North American origin**.** Variables included in the RFA are summarized in **Table 7**. SMOTE parameters to balance the smaller non-North American group are in **Table 8**. North-Americans were distinguishable well from non-North Americans via RFA (AUC = 0.82, F_1 = 0.83; **Fig. 5**). As suspected, some of the top variables differentiating **P3** from the other profiles were also important variables differentiating groups by continent of origin (**Fig. 5**). Among these, toxoplasma seropositivity, female sex, and low marijuana use were by far the most important (**Fig. 5**). These results suggest that the importance of toxoplasma seropositivity, female sex, and marijuana use in distinguishing between the better performing profiles (**P1**/**P2**) from **P3** is most likely due to demographic differences and sampling bias between the profiles (as most non-North American participants clustered in **P3**).

Table 7. List of variables considered for inclusion in the Random Forest Analyses between North American (NA) and non-North American (non-NA) participants (x = included).

Variable	NA vs non-NA
Age	X
Sex	X
Continent of Origin	-
Employment	X
Cognitive Reserve	X
Quality of Life	X
Duration of HIV Diagnosis	X
Recent VL	-
Detectable VL	X
Peak VL	X
Recent CD4 T-cell count	X
Nadir CD4 T-cell count	X
AIDS	-
On ART at Test Date	-
ART non-adherence	-
ART Side-effects	X
CPE	X
Alcohol (binge)	-
Marijuana Use	X
Crack/cocaine	-
Other illicit drugs	-
HCV Co-infection	-
Toxoplasma Seropositivity	X
Metabolic Conditions	X
Diabetes	-
Cardiac Conditions	X
Psychiatric Disorders	X
Domestic Violence	X
PHQ-9	X

Refer to Table 2 abbreviations.

Table 8. Synthetic Minority Over-sampling Technique (SMOTE) parameters used to balance the size of non-North American group (entire sample, N = 381)

Model	**NA vs non-NA**	
Unbalanced	**P1**	**P3**
n	300	81
Classification %	100%	0%
Balanced		
perc.under:perc.over	*150*	*200*
n	243	243
Classification %	92.2%	77.0%

Abbreviations. NA: North American; perc.over: upsampling parameter; perc.under: downsampling parameter.

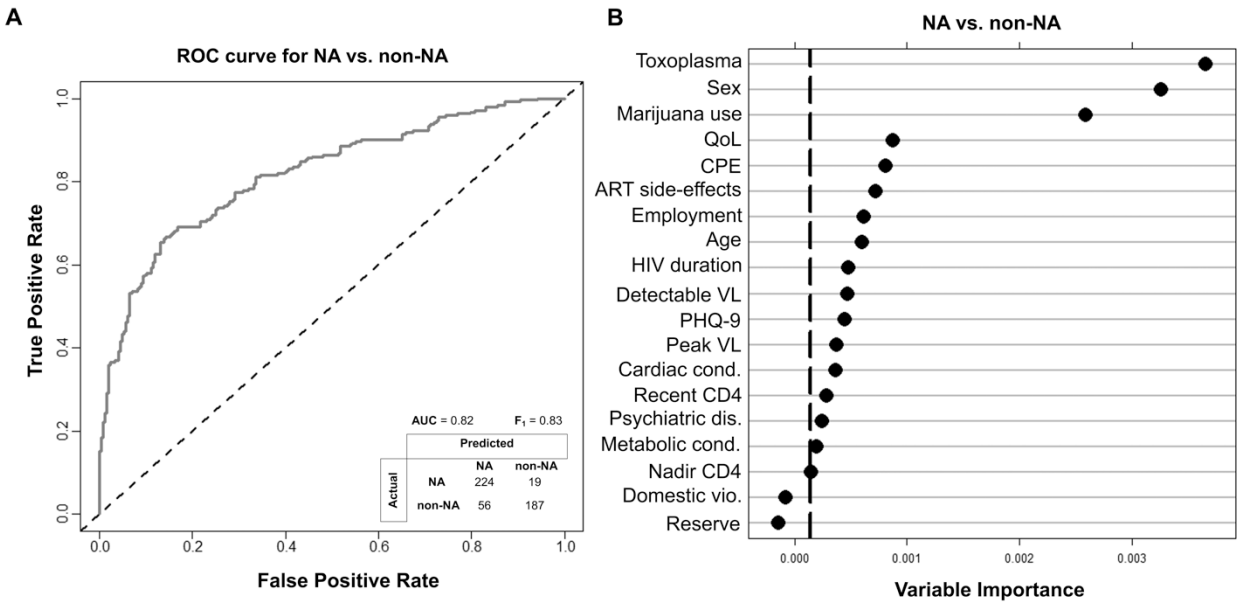

Figure 5. Random Forest Analyses (RFA), showing the importance for specific variables distinguishing North American (NA) versus non-North American-born (non-NA) participants. A) Receiver Operating Characteristic curve for RFA results. B) Random Forest Analyses (RFA) results. The dotted black line represents variable importance cut-off; dots to the right of the line indicate important variables.

Refer Figure 3 for abbreviations.

4.4. North American participants

To uncover potential risk-factors for **P3** in a demographically more homogenous group, we also repeated the analyses in North American participants only (n = 300). Like our LPA results in the entire cohort, a three-profile model had the best fit (cf. **Table 4**). The North-American **P1** (**P1$_{NA}$**) was composed of 136 patients (45.3% of the cohort), **P2 $_{NA}$** had 135 patients (45.0%), and **P3 $_{NA}$** had 29 patients (9.7%; **Fig. 6A**). Profiles again differed substantially in mean performance ($F(2, 295) = 277.5$, $p < 0.001$; **P1$_{NA}$** > **P2$_{NA}$** > **P3$_{NA}$**), and neurocognitive impairment was restricted to **P3$_{NA}$** (**Table 9**). Patterns in **P1$_{NA}$** and **P2$_{NA}$** resembled our previous results (compare **Figs. 3A and 6A**), but memory in **P3$_{NA}$** and **P2$_{NA}$** was now similar ($p > 0.05$; **Table 9**). Univariate comparisons (**Table 10**) revealed only a few profile differences: **P2 $_{NA}$** had a significantly higher proportion of Indigenous participants, and **P3$_{NA}$** had significantly higher age, proportions of retired/unemployed, and CPE, with trends pointing to longer duration of HIV-infection and higher rates of diabetes.

RFA omitted origin and toxoplasma sero-status as predictors (<10% incidence). SMOTE parameters again increased the size of **P3$_{NA}$** only (**Table 11**). Again, RFAs involving **P3$_{NA}$** showed acceptable classification (**Figure 6B/C**; **P1$_{NA}$/P2$_{NA}$**: AUC = 0.46, $F_1 = 0.46$). The most important predictors differentiating **P3$_{NA}$** from both other profiles were duration of HIV-infection, CPE, and unemployment/retirement (**Figs. 6D/E**). Older age, psychiatric comorbidities, low nadir, and QoL differentiated **P3$_{NA}$** from **P1$_{NA}$**. Lower rates of marijuana use and higher rates of ART side-effects further distinguished **P3$_{NA}$** from **P2$_{NA}$** (**Fig. 6E**). Thus, risk-factors for neurocognitive impairment in North American participants were mainly comprised of HIV-duration, age, CPE, and psychiatric comorbidities (**Figure 7**).

Logistic regressions were applied to the five most important variables identified by RFA in the North American sample to verify our results. The five most important variables of the **P1$_{NA}$** vs. **P3$_{NA}$** comparison were: duration of HIV infection, CPE rank, employment status, age, and psychiatric disorders. For the **P2$_{NA}$** vs. **P3$_{NA}$** comparison, variables were: HIV duration, employment status, CPE rank, nadir CD4, and marijuana use. Both logistic regression models were significant (**P1$_{NA}$** / **P3$_{NA}$**: $\chi^2 = 26.8$, df = 5; **P2$_{NA}$** / **P3$_{NA}$**: $\chi^2 = 20.9$, df = 5; *p*'s < .001). None of the selected predictors showed VIF or tolerant scores indicative of multicollinearity (**P1$_{NA}$** /**P3$_{NA}$**: $VIF_{max} = 1.41$, $tolerance_{max} = 0.71$; **P2$_{NA}$** /**P3$_{NA}$**: $VIF_{max} = 1.71$, $tolerance_{max} = 0.86$). In the **P1$_{NA}$** / **P3$_{NA}$** model, CPE rank (OR = 1.5, [0.9, 1.1], $p = .03$), and being unemployed/retired (OR = 0.3, [0.1, 0.7], $p < .01$) significantly predicted membership into **P3$_{NA}$**. There was a non-significant trend for older age predicting membership into **P3$_{NA}$** (OR = 1.0, [1.0, 1.1], $p = .07$). Being unemployed/retired (OR = 0.3, [0.1, 0.7], $p < .01$), and no marijuana use (OR = 0.3, [0.1, 0.9], $p = .03$) predicted **P3$_{NA}$** membership compared to **P2$_{NA}$**. There were non-significant trends towards higher CPE rank ($p = .08$) and longer HIV duration ($p = .09$) predicting **P3$_{NA}$** membership. Again, outcomes replicated the essential RFA findings, but with reduced sensitivity.

Table 9. Mean z-scaled performance (standard deviations) in the ten neuropsychological indicators used in Latent Profile Analyses in North American-born patients (n = 300)

Test	**P1$_{NA}$ (n=136)**	**P2$_{NA}$ (n=135)**	**P3$_{NA}$ (n=29)**
SDMT[a,b,c]	0.46 (0.85)	0.02 (0.84)	-1.25 (0.74)
TMT-2 [b,c]	-0.00 (0.72)	-0.09 (0.64)	-1.98 (0.84)
TMT-4[a,b,c]	0.04 (0.63)	-0.25 (0.72)	-2.01 (0.86)
WCST p-err.[b,c]	-0.29 (0.87)	-0.50 (0.73)	-1.11 (0.78)
D-KEFS FAS[a,b,c]	-0.16 (0.93)	-0.62 (0.77)	-1.3 (0.77)
D-KEFS Animals[a,b,c]	0.68 (0.84)	0.00 (0.87)	-0.71 (1.01)
HVLT-R imm.[a,b]	1.00 (0.72)	-0.61 (0.68)	-0.86 (0.75)
HVLT-R del [a,b]	0.66 (1.09)	-0.79 (0.70)	-0.66 (0.74)
GPB dom.[a,b,c]	0.27 (1.03)	-0.09 (0.90)	-1.20 (1.09)
GPB non-dom.[b,c]	0.33 (1.11)	0.03 (0.97)	-1.14 (1.16)

[a]: **P1$_{NA}$** > **P2$_{NA}$** ($p < .05$). [b]: **P1$_{NA}$** > **P3$_{NA}$** ($p < .05$). [c]: **P2$_{NA}$** > **P3$_{NA}$** ($p < .05$)

Refer to Table 2 for abbreviations.

.

Table 10. Patient characteristics and differences between latent profiles in North American-born patients (n = 300); data are means (standard deviations) or percentages within profiles, unless stated otherwise

Variable	$P1_{NA}$ (n = 136)	$P2_{NA}$ (n = 135)	$P3_{NA}$ (n = 29)	Test Statistic	p-value
Age (years)	46.5 (10.7)[a]	49.0 (10.7)[a,b]	52.5 (11.1)[b]	F(2, 297) = 4.36	**.01**
Sex (male)	91.9%	88.9%	89.7%	$\chi^2(2) = 0.73$	.70
Indigenous ethnicity (self-report)	0.7%[a]	7.4%[b]	6.9%[a,b]	$\chi^2(2) = 7.78$	**.02**
Employment (currently employed)	75.7%[a]	69.6%[a]	37.9%[b]	$\chi^2(2) = 16.08$	**< .001**
Cognitive Reserve (% high)	57.4%	52.3%	44.4%	$\chi^2(2) = 1.72$	.42
QoL (1 = poor, 5 = excellent) †	4 (1-5)	4 (1-5)	4 (1-5)	$\chi^2(2) = 0.88$	.65
HIV duration (in years)	11.6 (8.3)	11.9 (8.2)	15.6 (8.9)	F(2, 297) = 2.91	.06
Recent VL copies/ml (log) †	1.6 (1.6-5.2)	1.6 (1.6-5.2)	1.6 (1.6-4.9)	$\chi^2(2) = 0.35$	.84
Detectable VL	12.5%	14.8%	10.3%	$\chi^2(2) = 0.56$	.76
Peak VL copies/ml (log) †	4.9 (1.6-6.8)	4.8 (1.6-6.8)	4.9 (1.7-7.0)	$\chi^2(2) = 0.13$	.94
Recent CD4 T-cell (count/mm³)	584.9 (241.4)	572.9 (276.3)	548.8 (252.8)	F(2, 297) =0.25	.78
Nadir CD4 T-cell (count/mm³)	226.3 (179.0)	228.0 (166.4)	182.6 (147.2)	F(2, 297) = 0.90	.41
AIDS (nadir < 200 count/mm³)	47.1%	42.2%	55.2%	$\chi^2(2) = 1.81$	.41
On ART at test date	92.6%	93.3%	100%	$\chi^2(2) = 2.22$	.33
ART non-adherence (last five days)	7.2%	9.5%	6.9%	$\chi^2(2) = 0.52$	.77
ART side-effects	14.7%	17.0%	20.7%	$\chi^2(2) = 0.72$	.70
CPE †	7 (0-11)[a]	7 (0-13)[a]	8 (6-11)[b]	$\chi^2(2) = 8.97$	**.01**
Substance use					
Alcohol (binge)	3.7%	6.7%	3.4%	$\chi^2(2) = 1.44$	.49

Variable	$P1_{NA}$ (n = 136)	$P2_{NA}$ (n = 135)	$P3_{NA}$ (n = 29)	Test Statistic	p-value
Marijuana	29.4%	36.3%	17.2%	$\chi^2(2) = 4.46$	.11
Crack/cocaine	8.8%	4.4%	3.4%	$\chi^2(2) = 2.64$	.27
Other illicit drugs	3.7%	3.0%	0%	$\chi^2(2) = 1.11$	.57
HCV co-infection	2.9%	8.1%	3.4%	$\chi^2(2) = 3.86$	.15
Toxoplasma seropositive	7.4%	3.7%	6.9%	$\chi^2(2) = 1.78$	.41
Dyslipidemia	28.7%	27.4%	31.0%	$\chi^2(2) = 0.17$	.92
Lipodystrophy	9.6%	10.4%	10.3%	$\chi^2(2) = 0.05$	.97
Diabetes	5.1%	8.1%	17.2%	$\chi^2(2) = 5.02$	.08
Cardiac conditions	13.2%	15.6%	13.8%	$\chi^2(2) = 0.31$	.86
Psychiatric disorders	30.1%	32.6%	41.4%	$\chi^2(2) = 1.39$	.50
Domestic violence	25.0%	29.6%	37.9%	$\chi^2(2) = 2.17$	.34
PHQ-9 †	4 (0-26)	5 (0-26)	5.5 (0-21)	$\chi^2(2) = 0.94$	.63

[a,b]: different superscripts denote differences between groups (Bonferroni post-hoc's: $p < .05$). †: Non-normal variables are reported as medians (range). PHQ-9 available for n=287.
Refer Table 6 for abbreviations

Table 11. Synthetic Minority Over-sampling Technique (SMOTE) parameters used to balance the size of P3 (North American only; N = 300)

Model		**$P1_{NA}$ vs $P3_{NA}$**		**$P2_{NA}$ vs $P3_{NA}$**	
Unbalanced		**$P1_{NA}$**	**$P3_{NA}$**	**$P2_{NA}$**	**$P3_{NA}$**
	n	136	29	135	29
	Classification %	100%	0%	100%	0%
Balanced					
	perc.under:perc.over	*150*	*200*	*150*	*200*
	n	87	87	87	87
	Classification %	73.6%	78.2%	67.8%	70.1%

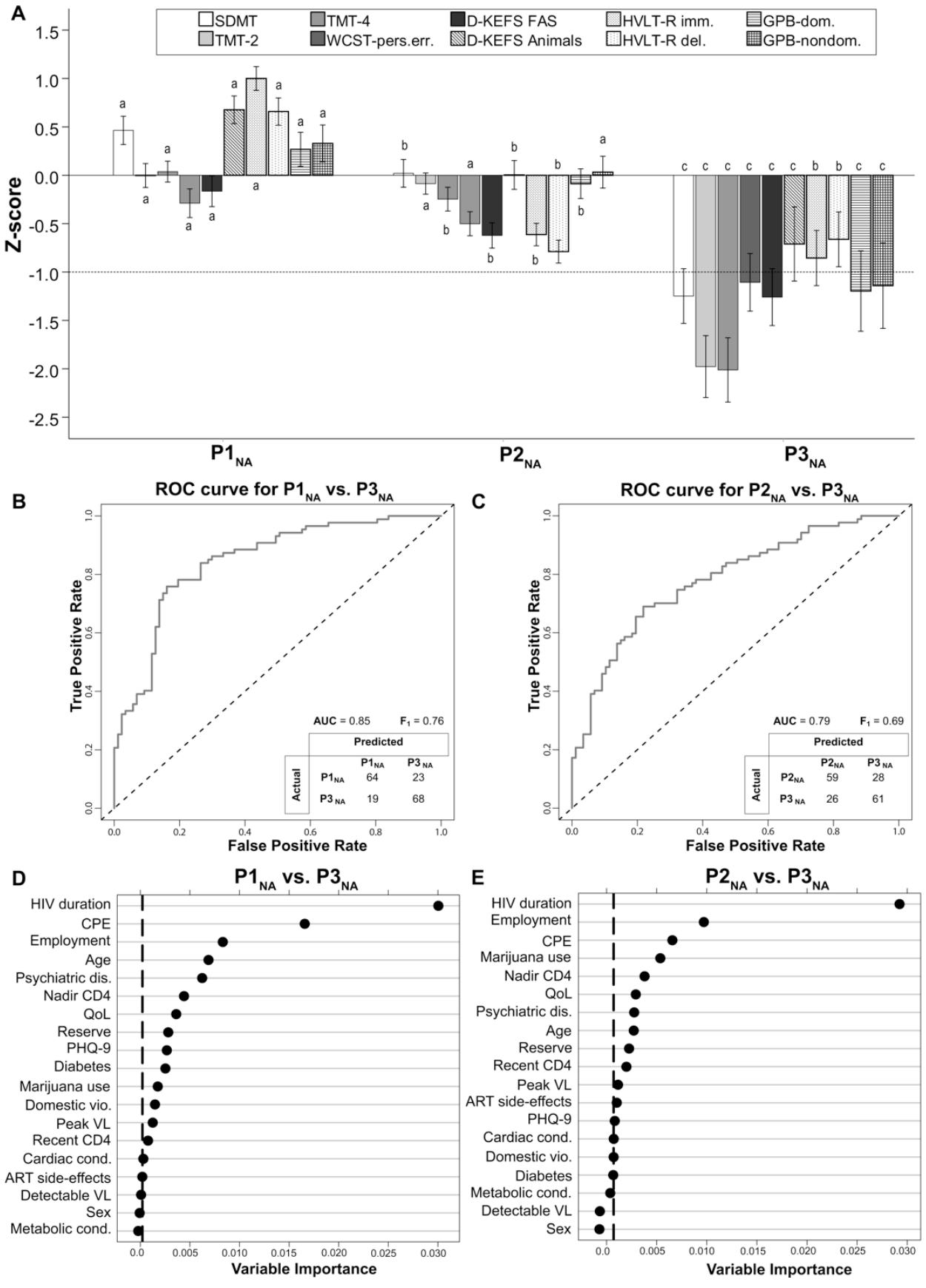
A
SDMT
TMT-4
D-KEFS FAS
HVLT-R imm.
GPB-dom.
TMT-2
WCST-pers.err.
D-KEFS Animals
HVLT-R del.
GPB-nondom.
Z-score
P1NA
P2NA
P3NA
B
ROC curve for P1NA vs. P3NA
True Positive Rate
False Positive Rate
AUC = 0.85
F1 = 0.76
Predicted
Actual
P1NA
P3NA
64
23
19
68
C
ROC curve for P2NA vs. P3NA
AUC = 0.79
F1 = 0.69
P2NA
59
28
26
61
D
P1NA vs. P3NA
HIV duration
CPE
Employment
Age
Psychiatric dis.
Nadir CD4
QoL
Reserve
PHQ-9
Diabetes
Marijuana use
Domestic vio.
Peak VL
Recent CD4
Cardiac cond.
ART side-effects
Detectable VL
Sex
Metabolic cond.
Variable Importance
E
P2NA vs. P3NA
HIV duration
Employment
CPE
Marijuana use
Nadir CD4
QoL
Psychiatric dis.
Age
Reserve
Recent CD4
Peak VL
ART side-effects
PHQ-9
Cardiac cond.
Domestic vio.
Diabetes
Metabolic cond.
Detectable VL
Sex
Variable Importance

Figure 6. Latent Profile Analyses and Random Forest Analyses of 300 HIV-infected North American participants from the Southern Alberta HIV Clinic. A) Distribution of test performance of participants across three latent neurocognitive profiles derived from Latent Profile Analysis in North American participants only. B) & C) Receiver Operating Characteristic curve for RFA results. D) & E) Random Forest Analyses (RFA), showing the importance for specific variables distinguishing **P1$_{NA}$** / **P3$_{NA}$**, and **P2$_{NA}$** / **P3$_{NA}$**, respectively. The dotted black line represents variable importance cut-off; dots to the right of the line indicate important variables.
a,b,c: different superscripts denote differences between groups (Bonferroni post-hoc's: $p < .05$).
Refer to Figure 3 for abbreviations.

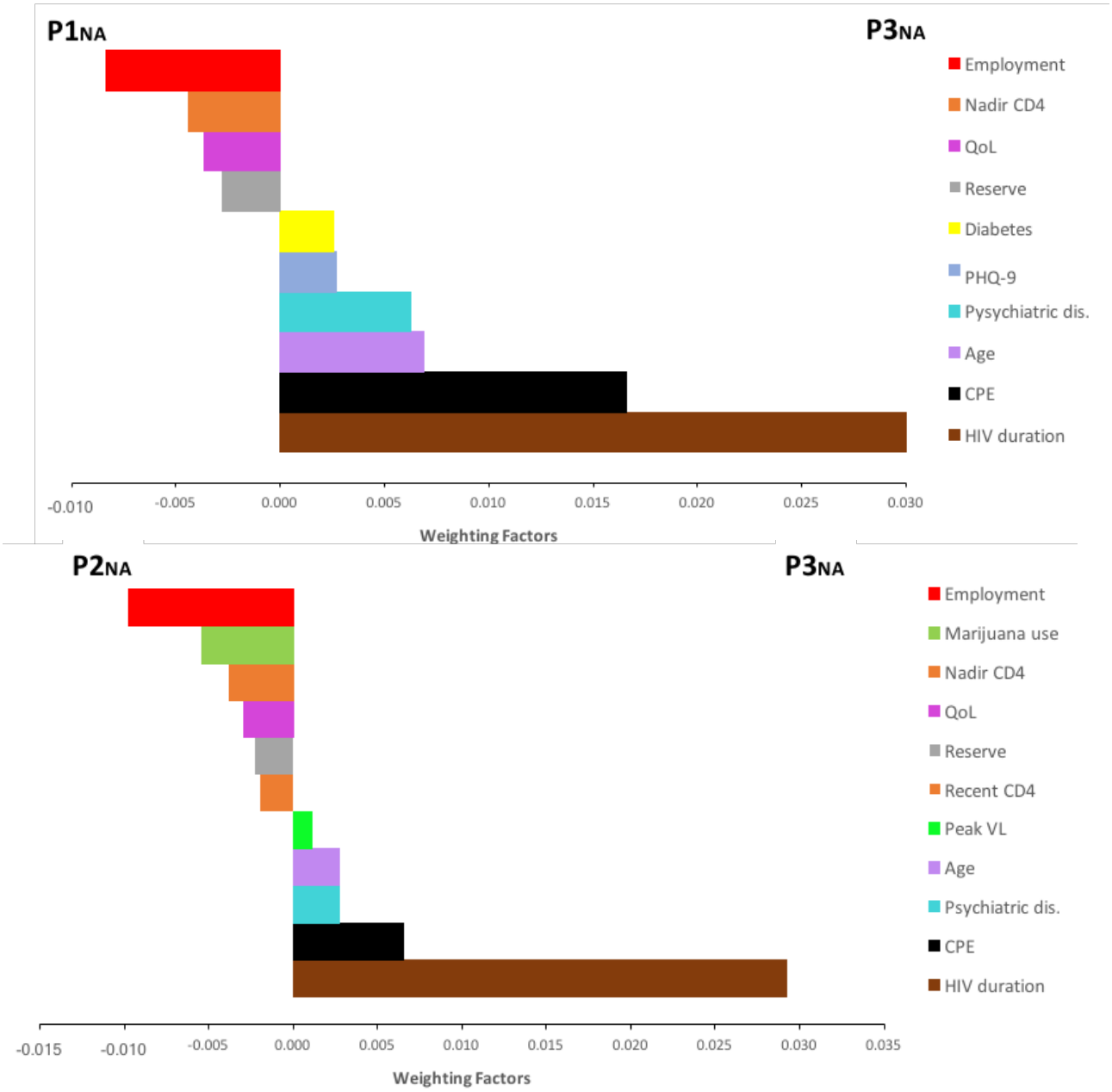

Figure 7. Summary of important variables predicting neurocognitive performance in the North American patients from the Southern Alberta Clinic cohort (n = 300) identified by Random Forest Analysis. Length of bar represents relative importance.
Refer to Figure 3 for abbreviations.

5. DISCUSSION

This is the first study implementing both LPA and RFA to identify and distinguish neurocognitive performance profiles in HIV-infected persons. The present cohort was representative of the HIV-infected population in care described in Canada (Public Health Organization, 2018) and other high-income countries. Three distinct profiles emerged: patients in **P1** (41.7%) displayed the highest overall neurocognitive performance, especially for verbal memory, while patients in **P2** (42.8%) showed relatively weaker executive and verbal memory functions. Only patients in **P3** (15.5%) displayed neurocognitive impairment, with pronounced deficits in the Trail-Making tasks (**Fig. 3A**). Random Forest Analysis revealed that North American descent, and features associated with place of origin including toxoplasmosis and female sex were the most robust predictors of membership to **P3** in the entire cohort (**Fig. 4**). Constraining analyses to North American participants revealed that highly CNS-penetrant ART regimens, or ART drugs with higher CPE ranks were important predictors of neurocognitive impairment. Older age and long-standing HIV were also amongst the top predictors of neurocognitive impairment in the North American subsample (**Fig. 7**).

These results provide a proof-of-principle for the use of LPA and RFA to delineate neurocognitive profiles and their associated risk-factors in a demographically representative cohort of patients with HIV-infection receiving contemporary care. The following discussion will compare the current findings with previous results, expand on the uncovered predictors, and will touch on future directions.

5.1. Identifying cognitive performance patterns – useful in HIV?

We expected to uncover two distinct profiles of neurocognitive impairment. Contrary to our hypothesis, we failed to identify multiple patterns of neurocognitive *impairment*. Instead, only one globally impaired group was found, both when considering the entire cohort and the North American patients only. One may think that we were unable to uncover distinct types of impairment profiles due to the relatively small size of our cohort, and small number of neurocognitive impaired patients. However, our results are remarkably similar to the LPA findings from the much larger MACS cohort (N = 2904, including 1531 HIV-positive participants) that also pointed to only one with global impaired group (Molsberry et al., 2018). In addition, Molsberry et al. (2018) also found an intermediate but neurocognitively unimpaired profile with selectively lowered memory function (**P2/P2$_{NA}$** also had relative reductions in memory as well as in executive functions), and one profile with high memory performance (similar to **P1/P1$_{NA}$** here).

We failed to find a profile with predominant motor deficits unlike in previous cluster analytical studies (Lojek & Bornstein, 2005a; van Gorp et al., 1993). In our cohort, motor functions were indeed relatively preserved. This discrepancy might be explainable by the fact that data of participants in Lojek & Bornstein (2005) and van Gorp et al (1993) were obtained prior to the introduction of highly effective ART in the mid 1990's (Palella et al., 1998). Neurocognitive deficits in the pre-ART era are known to encompass marked motor deficits due to subcortical HIV pathology as a result from uncontrolled viral replication within the brain (Carey, Woods, Rippeth, et al., 2004; Heaton et al., 2011). Therefore, the lack of a motor-impaired profile likely represents neurocognitive functions of well-treated HIV cohorts in the ART era. Interestingly, a more recent cohort receiving current ART also reported only one

globally impaired group (like **P3**), along with one higher performing group (resembling a combination of **P1**/**P2**) (Fazeli et al., 2014).

Even though multiple impairment profiles were not uncovered here, both **P2** and **P2$_{NA}$** represented a group of individuals with intermediate but unimpaired neurocognitive performance. Most patients within these intermediary profiles performed were not diagnosed with HAND and their characteristics were undistinguishable from those of individuals in **P1**/**P1$_{NA}$**. However, it is still possible that their relative reductions in executive and memory functions could be subtle subclinical neurocognitive difficulties. These would not have been captured by a HAND diagnosis. Taken together, our and previous results imply that apart from a global NCI pattern, distinguishable neurocognitive impartment profiles in well-treated HIV cohorts are unlikely. Nevertheless, determination of profiles with different longitudinal trajectories of cognitive performance may be more informative in capturing the heterogeneity of NCI in HIV.

5.2. Confounding influences of demographic variability on NCI

The most surprising results was the magnitude of the influence of place of origin on NCI here: As clearly seen in **Fig. 4**, being born outside of North America was by far the most important variable to determine membership in the globally impaired **P3** profile. When non-North American participants were excluded from our analysis, the impaired **P3** profile was reduced by half, and the overall rate of cognitive impairment in our sample decreased from 15.5% to less than 10%. Despite the demographic differences between the current and the American MACS cohort, Molsberry et al. (2018) also found that non-White ethnicity (e.g., African American) was by far the most prominent indicator distinguishing the globally impaired profile from the two

others. Country of origin and ethnic background may be influencing quality of education, literacy, income etc. which could then inflate their rates of NCI. Indeed, all patients were fluent in English and sufficiently educated to perform the neurocognitive testing. However, performance differences based on origin may be influenced by the applied normative adjustments which are mainly derived from Caucasian reference cohorts from the U.S. Unfortunately, existing adjustments for ethnic minority groups apply to minorities in the U.S. (e.g., Hispanic; Cherner et al., 2007) and would not have been suited here (e.g., for Indigenous peoples, Sub-Saharan Africans, South-East Asians). Higher rates of neurocognitive impairment in HIV-positive minority groups (particularly deficits in executive functions) is a consistent finding across many studies (Marquine et al., 2016; Mindt et al., 2008; Wojna et al., 2006) highlighting the influence of sociocultural factors such as quality of education, acculturation, and literacy (to name a few) when assessing neurocognitive functions in minority groups. Longitudinal assessment of neurocognitive performance could help overcome the lack of appropriate norms, as performance can then be compared to each participant's baseline scores. However, longitudinal assessments are lengthy, costly, and introduce other confounds that may also influence performance such as practice effects. Ethnic variability is a reality of HIV/AIDS in Canada and elsewhere and the lack of appropriate tests/norms to properly assess these patients must be addressed. There is a clear need for the development of more appropriate test norms or tests, ideally from well-matched local HIV-negative control groups (De Francesco et al., 2016).

Second to origin of birth, toxoplasma seropositivity and female sex characterized **P3**, these two factors also distinguished North American participants from others (**Figure 5**). Toxoplasma seropositivity is a recognized risk-factor for neurocognitive dysfunction in HIV-infection (Bharti et al., 2016). Negative relationships between toxoplasmosis and neurocognitive performance have

been reported in a healthy HIV-uninfected cohort (Gale, Brown, Erickson, Berrett, & Hedges, 2015), but only in combination with other risk factors, i.e., in racial/ethnic minorities, groups with higher poverty-to-income ratio, or lower education. Therefore, the importance of toxoplasma seropositivity in this profile might well reflect socio-economic circumstances of patients in **P3**. Females have been found to be particularly vulnerable to neurocognitive effects of HIV-infection (Maki & Martin-Thormeyer, 2009; Maki et al., 2015), aggravated by advanced age, lower reading levels, and non-Caucasian ethnicity (Maki et al., 2015). Women are also susceptible to neurocognitive changes due to stress, in particular memory problems (Rubin et al., 2017; Rubin et al., 2014). In the current cohort, over a third of females were born outside North America, reported current domestic violence, and may have experienced trauma (not assessed in detail). Female sex was also key predictors of non-North American origin (see **Fig. 5**), thus it is difficult to disentangle effects of female sex *per se* from those of non-North American descent. It is likely that a combination of demographic and societal disadvantages, along with biological vulnerability (e.g., to effects of stress and trauma) contributed to the overrepresentation of females in **P3**. The low proportion of female participants (n = 46, 12% of the entire cohort) precluded us from further testing of potential female-specific cognitive risk factors. Such should be tested in cohorts with larger numbers of women (e.g., the Women's Interagency HIV Study, WIHS).

It is evident from our results that rates of NCI may be biased by the inclusion of participants of diverse ethnic and sociocultural backgrounds. Non-White participants are overrepresented in cognitively impaired groups in multiple other published HIV cohorts (Heaton et al., 2015; Maki et al., 2015; Winston et al., 2013). The majority of non-North American participants in the current cohort are of sub-Saharan origin, which is a different demographic than HIV cohorts from the U.S.A. Although new norms have been developed specifically for African participants (Singh et

al., 2010), these only exist for a few select tests (e.g., Trail Making Tests and Digit Span). Additionally, these norms were obtained in African countries (Uganda, South Africa) and therefore apply to participants educated and living in these countries. This context is clearly different from the current cohort, where (recent or more remote) African immigrants have been living in Canada and may have been educated and/or part of the Canadian workforce for a while. Therefore, currently available norms are not culturally adequate to assess neurocognitive performance in our non-North American subgroup. Until more appropriate norms are established with local and demographically matched HIV-negative participants, the neurocognitive performance of foreign-born individuals should not be interpreted as NCI. Notably, normative test adjustments were available and applied to many of the U.S.-based HIV cohorts, but evidently only partly addressed the strong influence of non-White ethnicity to NCI. Multiple other factors that may be confounded with non-White ethnicity in the U.S.A. (i.e., access to care, stigma, SES) or those identified here (**Fig. 5**; toxoplasmosis, female sex) may still exert deleterious influences on test performance despite appropriate normative data. Owing to these complications in interpretations of neuropsychological test results in non-North American participants, including but not limited to the lack of appropriate normative data, the following discussion is restricted to findings from the North American subset of the cohort.

5.3. Risk factors for neurocognitive impairment in North American patients

While in the entire cohort non-North American descent was clearly predominant as the main predictor of neurocognitive impairment, analysis of only North American participants revealed multiple HIV- and treatment-related, demographic and comorbid risk-factors.

5.3.1. Demographic variables

In North Americans, older age, lower cognitive reserve and unemployment distinguished the impaired profile from (**P3$_{NA}$**) from the higher performing profiles (**P1$_{NA}$**, **P2$_{NA}$**). HIV and ageing represent independent risk factors for neuropathological changes in the CNS and their combination may be particularly detrimental (Canizares, Cherner, & Ellis, 2014). Additive or synergistic effects of multiple age-related changes with HIV have been suggested, including persistent neuroinflammation, oxidative stress, cardiovascular and metabolic disorders, increased immune system reactivity, and decreased metabolism of ART with associated potential of neurotoxicity (Soontornniyomkij & Achim, 2012). Finding age among the most important variables to distinguish P3$_{NA}$ from the other profiles follows these results. Some of these age-related factors were also assessed here and emerged as important predictors of NCI. These will be discussed in section **5.3.3.**

Cognitive reserve, here defined as the combination of reading ability and education was higher in **P1$_{NA}$** than **P3$_{NA}$**. Employment rates were also higher in the high-performing group. These observations are similar to previous findings implicating cognitive reserve as protective against age- and HIV-related neurological insults (Patel et al., 2013). For example, intact neuropsychological functions have been found to more strongly depend on cognitive reserve (education level and reading ability) in older HIV-infected individuals compared younger HIV-positive participant as well as older HIV-uninfected individuals (Foley et al., 2012). Additionally, low cognitive reserve has been associated with more severe forms of NCI in HIV (Morgan et al., 2012). Cognitive reserve-building educational, work-related or leisure-time activities along with high premorbid IQ and educational attainment have been associated with higher neurocognitive performance in older HIV cohorts (Milanini et al., 2016).

5.3.2. HIV- and treatment-related variables

Neurocognitive impairment in North American participants was associated with longstanding HIV infection, low nadir CD4 counts, and ART regimens with higher CPE ranks. Low nadir CD4 count has been consistently associated with lower neurocognitive performance in HIV (Saylor et al., 2016), perhaps reflecting potential adverse effects of immunosuppression on brain structure, i.e., irreversible brain damage ('legacy effect'; Ellis et al. (2011), possibly caused by viral replication and ensuing neuroinflammation or unrecognized brain infections. The importance of the CPE could point to ART-related neurotoxicity as a consequence of higher CNS penetrance, leading to subsequent cognitive reductions (Underwood, Robertson, & Winston, 2015) potentially associated with brain changes after long-term ART exposure (Jernigan et al., 2011). However, beneficial cognitive effects of ART-treatment with high CPE scores are also reported (Cysique et al., 2009). A causal relationship between CPE and cognition cannot be inferred in this study. That is, individuals with NCI may have been treated more aggressively in order to address NCI rather than preceding NCI. This should be clarified in longitudinal studies.

Though the CPE ranking system is widely used, it has serious limitations, including a semi-quantitative/qualitative ranking based on animal and in-vitro studies that varied in quality (Letendre, 2011; Letendre et al., 2008), potential inaccuracies in equating concentrations of ART in the CSF with those in brain tissue (Pardridge, 2011; Brewster et al., 1997); and the lack of CPE information on newer ARTs. Thus, an alternative method to assess exposure to ARTs as well as classes of different ARTs should be employed, for example, length of exposure or amount of drug consumed.

Given the rapidly growing population of older HIV-infected patients, understanding the neurocognitive consequences of the interplay between age, longstanding infection, and prolonged ART exposure is imperative.

5.3.3. Comorbidities

Psychiatric diagnoses, current depressive symptoms in the PHQ-9, diabetes, and marijuana use also differentiated neurocognitive profiles. Psychiatric comorbidities such as depression are known to reduce neurocognitive performance in patients with HIV-infection, although effects are usually mild (Fellows, Byrd, & Morgello, 2013). As PHQ-9 scores differentiated **P3$_{NA}$** only from the best performing profiles (**P1$_{NA}$**), our findings concur with this observation. They also highlight the importance of identifying and treating depression in patients with HIV/AIDS to address neurocognitive problems.

Diabetes and other metabolic conditions occur at higher than normal rates in HIV/AIDS (Tripathi et al., 2013) and have been implicated in neurocognitive deficits in HIV/AIDS by multiple studies (Canizares et al., 2014; Dufouil, Richert, Thiebaut, Bruyand, Amieva, Dauchy, Dartigues, Neau, Morlat, Dehail, Dabis, Bonnet, Chene, & ANRS CO3 Aquitaine Study Group, 2015; Vance et al., 2014). Longitudinally, comorbid diabetes has also been associated with more pronounced cognitive decline in HIV/AIDS (Dufouil, Richert, Thiebaut, Bruyand, Amieva, Dauchy, Dartigues, Neau, Morlat, Dehail, Dabis, Bonnet, Chene, & Group, 2015). Diabetes is associated with chronic inflammation (Lasselin & Capuron, 2014), the release of excess pro-inflammatory cytokines, and activation/expression of other inflammatory mediators, which in turn are associated with cognitive dysfunction (Heringa et al., 2014). Metabolic disturbances have also been attributed to long-term exposure to certain ART (Martin-Iguacel, Negredo, Peck,

& Friis-Moller, 2016), therefore the development of age-related comorbidities, such as diabetes, in the context of ART exposure, are important risk factors to address in light of the ageing HIV/AIDS demographic. Interestingly, marijuana use differentiated **P3$_{NA}$** from the higher performing profiles, with higher use in patients without neurocognitive impairment. Previous studies of the relationships between marijuana use and cognition in HIV-infection are inconclusive (Cristiani, Pukay-Martin, & Bornstein, 2004; Thames et al., 2017). Thames, Mahmood, Burggren, Karimian, and Kuhn (2016) recently compared cognitive functions between HIV-infected (n = 55) and HIV-negative (n = 34) marijuana users and reported lower verbal memory in moderate-to-heavy marijuana users with concurrent HIV-infection, compared to light users (and controls). However, in a follow up study heavy marijuana use did not increase cognitive dysfunction in HIV/AIDS (Thames et al., 2017). Deleterious effects of other substances (especially stimulants) on cognition in HIV are better established (Meyer et al., 2013; Tang et al., 2015), but could not be ascertained in this cohort due to relatively low rates of stimulant use. The potentially protective functions of marijuana should be further explored.

5.4. Strengths and limitations

This is the first study to apply machine learning to identify predictors of neurocognitive impairment in HIV. Due to the multifactorial nature of NCI, multiple predictors need to be concurrently assessed. However, many variables were multicollinear (e.g., age and HIV duration, age and employment, CD4 counts and viral load), precluding their use in standard regression models. The RFA represents an alternative technique to investigate risk-factors of NCI in HIV as it can overcome many statistical limitations of standard methods (mixing of continuous and categorical data, missing data, unbalanced class sizes, and moderate multicollinearity). Therefore,

the application of RFA can help address the demographic and clinical complexity of NCI in the context of HIV. Another strength of this study was that we avoided the relatively arbitrary definition of neurocognitive domains by instead using individual test scores and assessing performance patterns across those scores instead of 'domains'. The application of LPA represents a relatively novel approach in this area as well, and allowed the empirical identification of cognitive performance clusters. Furthermore, since the LPA profiles were derived directly from primary neuropsychological data, they are 'blind' to set thresholds of impairment severity, but rather identify similarities across performance patterns. Therefore, LPA is able to overcome some of the measurement ambiguities surrounding the current diagnostic guidelines and offers a more objective alternative to interpret neurocognitive performance and impairment in HIV/AIDS.

The main limitations of our study were the lack of HIV-uninfected controls and the limited size of the cohort. Furthermore, although established normative adjustments were applied to all tests, reference data accommodating the unique background of the current cohort were unavailable. To address this limitation, analyses were repeated in the North Americans with globally impaired performance, but this subgroup was small. The current results resemble the cognitive profiles found by applying LPA in a larger cohort that included HIV-negative controls (Molsberry et al., 2018). However, Molsberry et al. (2018) combined both groups and did not compare profiles and their predictors in HIV-infected patients with those in non-infected controls. The similarity between Molsberry et al. (2018) LPA profiles and ours suggests that the presence or absence of HIV-infection may not influence performance patterns across neuropsychological tests. A comparison of the current results, including both LPA and RFA, across multiple cohorts would be desirable to validate the profiles, and to verify HIV-specific and non-specific features associated with them.

5.5. Future directions

This study demonstrated the utility of data-driven and machine learning techniques to help delineate neurocognitive impairment subgroups and predictors in a well-treated HIV cohort. As recently noted also by others (Devlin & Giovannetti, 2017), the controversies surrounding the definitions and measurements of HAND might benefit from integration of data-driven statistical approaches like LPA. While these techniques cannot solve the clinical and etiological complexities underpinning HAND, they offer an objective alternative to investigating neurocognitive patterns among patients and might help corroborate HAND diagnoses. Data-driven techniques such as LPA can also be used for longitudinal modeling of change in neurocognitive performance (Nylund et al. 2007). A longitudinal analysis of LPA profiles (i.e., analysis of trajectories) would help validate the stability or change of neurocognitive profiles in HIV.

Neurocognitive impairment in HIV is not one-dimensional, but rather a complex and multidimensional process driven by both viral and host factors. Identification of genetic predictors of neurocognitive profiles was outside of the scope of the current study. Future studies may benefit from the incorporation of genetic data to identify the most important genetic predictors of neurocognitive profiles in HIV, perhaps first targeting the importance of CCR5 and APOe genotypes. Owing to the importance of CPE in the current outcomes, it may also be worthwhile to target CYP3A4 genotypes in conjunction with neurocognitive profiles in HIV. However, even though investigating the genetic (and epigenetic) susceptibility of patients to NCI has become a relevant area of interest given the heterogeneity across patients, there currently are no clear (epi-)genetic markers for NCI in HIV. This ambiguity may be driven by the lack of clear phenotypic definition of NCI. To broaden the current approach and establish consensus empirical subtypes of

neurocognition in HIV, one would require multiple large cohorts. This may result in more robust phenotypes. Machine learning could then aid in the identification of genetic and epigenetic biomarkers for the resulting neurocognitive phenotypes. These findings may help inform targeted in-vitro and animal studies to elucidate the mechanisms by which certain genes or epigenetic changes predispose individuals to NCI. For these types of approaches, shared public and collaborative global data repositories, including neuropsychological, clinical-demographic, genetic, and epigenetic data, will be needed to maximize statistical power.

Additionally, future studies could investigate further the association between ART treatment-factors and neurocognitive impairment. Highly penetrant ART regimens were important predictors of neurocognitive impairment in the current study, along with older age and longer HIV diagnosis duration. These results point to the potential detrimental effect of longstanding ART exposure to cognitive performance. Many individuals in the current cohort have been on antiretroviral medication for over two decades. In vitro studies point to potential neurotoxic properties of certain ART drugs (K. Robertson et al., 2012; Vivithanaporn et al., 2016). Given that people living with HIV depend on ART medication for life, it is imperative to understand their potential contribution to the development of NCI.

5.6. Conclusion

Between 20% to 60% of people living with HIV have cognitive problems, and these are concerning to both clinicians and patients. However, cognitive problems are variable across patients, and the etiology as well as staging of these problems are highly debated, including disagreements around the clinical criteria to define HIV-associated neurocognitive disorders (HAND). Heterogeneity of cognitive performance/impairment could also point to different risk

factors underlying different outcomes. Data-driven techniques offer an approach to complement current HAND staging, accommodate the multifactorial etiologies, and reflect the heterogeneity of outcomes. Implementing LPA and RFA in a well-characterized cohort of people living with HIV in Alberta, this proof of concept study demonstrates the feasibility and utility of these techniques to uncover cognitive profiles and their associated predictors. Here, we empirically derived two unimpaired profiles, and discovered only one profile with NCI. Accounting for the demographic confound of finding NCI associated with non-North American descent, main risk factors included a combination of known risk factors (older age, low cognitive reserve, age-related and psychiatric comorbidities, nadir CD4 T-cell count/immunosuppression) for NCI in HIV, along with treatment-related factors which could point to neurotoxic effects of long-standing ART exposure. The combination of the current data-driven statistical methods provides a novel methodological approach to quantifying neurocognitive performance and its associated characteristics in patients with treated HIV-infection.

References

Abutaleb, A., Kattakuzhy, S., Kottilil, S., O'Connor, E., & Wilson, E. (2018). Mechanisms of neuropathogenesis in HIV and HCV: Similarities, differences, and unknowns. *Journal of Neurovirology, 24*(6), 670-678. doi:10.1007/s13365-018-0678-5

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. Beverly Hills: Sage Publications.

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, 59*(5), 469-477. doi:10.1097/QAI.0b013e318249db17

Andrade, A. S., Deutsch, R., Celano, S., Duarte, N. A., Marcotte, T. D., Umlauf, A., . . . Collier, A. C. (2013). Relationships among neurocognitive status, medication adherence measured by pharmacy refill records, and virologic suppression in HIV-infected persons. *Journal of Acquired Immune Deficiency Syndromes, 62*(3), 282-292. doi:10.1097/QAI.0b013e31827ed678

Andres, M. A., Feger, U., Nath, A., Munsaka, S., Jiang, C. S., & Chang, L. (2011). APOE epsilon 4 allele and CSF APOE on cognition in HIV-infected subjects. *Journal of Neuroimmune Pharmacology, 6*(3), 389-398. doi:10.1007/s11481-010-9254-3

Antinori, A., Arendt, G., Becker, J. T., Brew, B. J., Byrd, D. A., Cherner, M., . . . Wojna, V. E. (2007). Updated research nosology for HIV-associated neurocognitive disorders. *Neurology, 69*(18), 1789-1799. doi:10.1212/01.WNL.0000287431.88658.8b

Asahchop, E. L., Akinwumi, S. M., Branton, W. G., Fujiwara, E., Gill, M. J., & Power, C. (2016). Plasma microRNA profiling predicts HIV-associated neurocognitive disorder. *AIDS, 30*(13), 2021-2031. doi:10.1097/QAD.0000000000001160

Basavaraju, A. (2016). Toxoplasmosis in HIV infection: An overview. *Tropical Parasitology, 6*(2), 129-135. doi:10.4103/2229-5070.190817

Becker, J. T., Martinson, J. J., Penugonda, S., Kingsley, L., Molsberry, S., Reynolds, S., . . . Sacktor, N. (2015). No association between Apoε4 alleles, HIV infection, age, neuropsychological outcome, or death. *Journal of Neurovirology, 21*(1), 24-31. doi:10.1007/s13365-014-0290-2

Berrett, A. N., Gale, S. D., Erickson, L. D., Brown, B. L., & Hedges, D. W. (2017). Toxoplasma Gondii moderates the association between multiple folate-cycle factors and cognitive function in U.S. adults. *Nutrients, 9*(6), 564. doi:10.3390/nu9060564

Bharti, A. R., McCutchan, A., Deutsch, R., Smith, D. M., Ellis, R. J., Cherner, M., . . . Letendre, S. L. (2016). Latent toxoplasma infection and higher toxoplasma gondii immunoglobulin G levels are associated with worse neurocognitive functioning in HIV-infected adults. *Clinical Infectious Diseases, 63*(12), 1655-1660. doi:10.1093/cid/ciw655

Blackstone, K., Moore, D. J., Franklin, D. R., Clifford, D. B., Collier, A. C., Marra, C. M., . . . Heaton, R. K. (2012). Defining neurocognitive impairment in HIV: Deficit scores versus clinical ratings. *The Clinical Neuropsychologist, 26*(6), 894-908. doi:10.1080/13854046.2012.694479

Bol, S. M., Booiman, T., van Manen, D., Bunnik, E. M., van Sighem, A. I., Sieberer, M., . . . van 't Wout, A. B. (2012). Single nucleotide polymorphism in gene encoding transcription factor Prep1 is associated with HIV-1-associated dementia. *PLoS One, 7*(2), e30990. doi:10.1371/journal.pone.0030990

Bonnet, F., Amieva, H., Marquant, F., Bernard, C., Bruyand, M., Dauchy, F. A., . . . ANRS CO3 Aquitaine Cohort. (2013). Cognitive disorders in HIV-infected patients: Are they HIV-related? *AIDS, 27*(3), 391-400. doi:10.1097/QAD.0b013e32835b1019

Boven, L. A., van der Bruggen, T., van Asbeck, B. S., Marx, J. J., & Nottet, H. S. (1999). Potential role of CCR5 polymorphism in the development of AIDS dementia complex. *FEMS Immunology and Medical Microbiology, 26*(3-4), 243-247. doi:10.1111/j.1574-695X.1999.tb01395.x

Brandt, J., & Benedict, R. H. (2001). *Hopkins Verbal Learning Test—Revised*. Lutz, FL: Psychological Assessment Resources, Inc.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. doi:10.1023/A:1010933404324

Bryant, V. E., Whitehead, N. E., Burrell, L. E., Dotson, V. M., Cook, R. L., Malloy, P., . . . Cohen, R. A. (2015). Depression and apathy among people living with HIV: Implications for treatment of HIV-associated neurocognitive disorders. *AIDS and Behavior, 19*(8), 1430-1437. doi:10.1007/s10461-014-0970-1

Burt, T. D., Agan, B. K., Marconi, V. C., He, W., Kulkarni, H., Mold, J. E., . . . Ahuja, S. K. (2008). Apolipoprotein (apo) E4 enhances HIV-1 cell entry in vitro, and the APOE epsilon4/epsilon4 genotype accelerates HIV disease progression. *Proceedings of the National Academy of Science, 105*(25), 8718-8723. doi:10.1073/pnas.0803526105

Canizares, S., Cherner, M., & Ellis, R. J. (2014). HIV and aging: Effects on the central nervous system. *Seminars in Neurology, 34*(1), 27-34. doi:10.1055/s-0034-1372340

Carey, C. L., Woods, S. P., Gonzalez, R., Conover, E., Marcotte, T. D., Grant, I., . . . HNRC Group. (2004). Predictive validity of global deficit scores in detecting neuropsychological

impairment in HIV infection. *Journal of Clinical and Experimental Neuropsychology, 26*(3), 307-319. doi:10.1080/13803390490510031

Carey, C. L., Woods, S. P., Rippeth, J. D., Gonzalez, R., Moore, D. J., Marcotte, T. D., . . . HNRC Group. (2004). Initial validation of a screening battery for the detection of HIV-associated cognitive impairment. *The Clinical Neuropsychologist, 18*(2), 234-248. doi:10.1080/13854040490501448

Chang, L., Andres, M., Sadino, J., Jiang, C. S., Nakama, H., Miller, E., & Ernst, T. (2011). Impact of apolipoprotein E epsilon4 and HIV on cognition and brain atrophy: Antagonistic pleiotropy and premature brain aging. *Neuroimage, 58*(4), 1017-1027. doi:10.1016/j.neuroimage.2011.07.010

Chang, L., Jiang, C., Cunningham, E., Buchthal, S., Douet, V., Andres, M., & Ernst, T. (2014). Effects of APOE epsilon4, age, and HIV on glial metabolites and cognitive deficits. *Neurology, 82*(24), 2213-2222. doi:10.1212/WNL.0000000000000526

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research, 16*(1), 321-357. doi:10.1613/jair.953

Cherner, M., Suarez, P., Lazzaretto, D., Fortuny, L. A., Mindt, M. R., Dawes, S., . . . HNRC Group. (2007). Demographically corrected norms for the Brief Visuospatial Memory Test-revised and Hopkins Verbal Learning Test-revised in monolingual Spanish speakers from the U.S.-Mexico border region. *Archives of Clinical Neuropsychology, 22*(3), 343-353. doi:10.1016/j.acn.2007.01.009

Chiang, M. C., Dutton, R. A., Hayashi, K. M., Lopez, O. L., Aizenstein, H. J., Toga, A. W., . . . Thompson, P. M. (2007). 3D pattern of brain atrophy in HIV/AIDS visualized using

tensor-based morphometry. *Neuroimage, 34*(1), 44-60. doi:10.1016/j.neuroimage.2006.08.030

Clifford, D. B., & Ances, B. M. (2013). HIV-associated neurocognitive disorder. *The Lancet Infectious Diseases, 13*(11), 976-986. doi:10.1016/S1473-3099(13)70269-X

Cohen, R. A., Seider, T. R., & Navia, B. (2015). HIV effects on age-associated neurocognitive dysfunction: premature cognitive aging or neurodegenerative disease? *Alzheimer's Research and Therapy, 7*(1), 37. doi:10.1186/s13195-015-0123-4

Collier, A. C., Coombs, R. W., Schoenfeld, D. A., Bassett, R. L., Timpone, J., Baruch, A., . . . Corey, L. (1996). Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. *New England Journal of Medicine, 334*(16), 1011-1017. doi:10.1056/NEJM199604183341602

Corder, E. H., Robertson, K., Lannfelt, L., Bogdanovic, N., Eggertsen, G., Wilkins, J., & Hall, C. (1998). HIV-infected subjects with the E4 allele for APOE have excess dementia and peripheral neuropathy. *Nature Medine, 4*(10), 1182-1184. doi:10.1038/2677

Corley, M. J., Dye, C., D'Antoni, M. L., Byron, M. M., Yo, K. L., Lum-Jones, A., . . . Maunakea, A. K. (2016). Comparative DNA methylation profiling reveals an immunoepigenetic signature of HIV-related cognitive impairment. *Scientific Report, 6*, 33310. doi:10.1038/srep33310

Crane, H. M., Van Rompaey, S. E., Dillingham, P. W., Herman, E., Diehr, P., & Kitahata, M. M. (2006). A single-item measure of health-related quality-of-life for HIV-infected patients in routine clinical care. *AIDS Patient Care and STDs, 20*(3), 161-174. doi:10.1089/apc.2006.20.161

Cristiani, S. A., Pukay-Martin, N. D., & Bornstein, R. A. (2004). Marijuana use and cognitive function in HIV-infected people. *Journal of Neuropsychiatry and Clinical Neurosciences, 16*(3), 330-335. doi:10.1176/jnp.16.3.330

Cysique, L. A., Heaton, R. K., Kamminga, J., Lane, T., Gates, T. M., Moore, D. M., . . . Brew, B. J. (2014). HIV-associated neurocognitive disorder in Australia: A case of a high-functioning and optimally treated cohort and implications for international neuroHIV research. *Journal of Neurovirology, 20*(3), 258-268. doi:10.1007/s13365-014-0242-x

Cysique, L. A., Maruff, P., & Brew, B. J. (2004). Prevalence and pattern of neuropsychological impairment in human immunodeficiency virus-infected/acquired immunodeficiency syndrome (HIV/AIDS) patients across pre- and post-highly active antiretroviral therapy eras: A combined study of two cohorts. *Journal of Neurovirology, 10*(6), 350-357. doi:10.1080/13550280490521078

Cysique, L. A., Vaida, F., Letendre, S., Gibson, S., Cherner, M., Woods, S. P., . . . Ellis, R. J. (2009). Dynamics of cognitive change in impaired HIV-positive patients initiating antiretroviral therapy. *Neurology, 73*(5), 342-348. doi:10.1212/WNL.0b013e3181ab2b3b

D'Aquila, R. T., Hughes, M. D., Johnson, V. A., Fischl, M. A., Sommadossi, J. P., Liou, S. H., . . . Hirsch, M. S. (1996). Nevirapine, zidovudine, and didanosine compared with zidovudine and didanosine in patients with HIV-1 infection. A randomized, double-blind, placebo-controlled trial. National Institute of Allergy and Infectious Diseases AIDS Clinical Trials Group Protocol 241 Investigators. *Annals of Internal Medicine, 124*(12), 1019-1030. doi:10.7326/0003-4819-124-12-199606150-00001

Davis, L. E., Hjelle, B. L., Miller, V. E., Palmer, D. L., Llewellyn, A. L., Merlin, T. L., . . . Wiley, C. A. (1992). Early viral brain invasion in iatrogenic human immunodeficiency virus infection. *Neurology, 42*(9), 1736-1739. doi:10.1212/WNL.42.9.1736

Dawes, S., Suarez, P., Casey, C. Y., Cherner, M., Marcotte, T. D., Letendre, S., . . . HNRC Group. (2008). Variable patterns of neuropsychological performance in HIV-1 infection. *Journal of Clinical and Experimental Neuropsychology, 30*(6), 613-626. doi:10.1080/13803390701565225

De Francesco, D., Underwood, J., Post, F. A., Vera, J. H., Williams, I., Boffito, M., . . . POPPY Study Group. (2016). Defining cognitive impairment in people-living-with-HIV: The POPPY study. *BMC Infectious Diseases, 16*(1), 617. doi:10.1186/s12879-016-1970-8

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: The Psychological Corporation.

Devlin, K. N., & Giovannetti, T. (2017). Heterogeneity of neuropsychological impairment in HIV infection: Contributions from mild cognitive impairment. *Neuropsychology Review, 27*(2), 101-123. doi:10.1007/s11065-017-9348-2

Diaz-Arrastia, R., Gong, Y., Kelly, C. J., & Gelman, B. B. (2004). Host genetic polymorphisms in human immunodeficiency virus-related neurologic disease. *Journal of Neurovirology, 10*(1), 67-73. doi:10.1080/753312755

Ding, J., Zhao, J., Zhou, J., Li, X., Wu, Y., Ge, M., & Cen, S. (2018). Association of gene polymorphism of SDF1(CXCR12) with susceptibility to HIV-1 infection and AIDS disease progression: a meta-analysis. *PLoS One, 13*(2), e0191930. doi:10.1371/journal.pone.0191930

Doyle, K. L., Woods, S. P., Morgan, E. E., Iudicello, J. E., Cameron, M. V., Gilbert, P. E., . . . HNRP Group. (2016). Health-related decision-making in HIV disease. *Journal of Clinical Psychology in Medical Settings, 23*(2), 135-146. doi:10.1007/s10880-016-9455-x

Dufouil, C., Richert, L., Thiebaut, R., Bruyand, M., Amieva, H., Dauchy, F. A., . . . ANRS CO3 Aquitaine Study Group. (2015). Diabetes and cognitive decline in a French cohort of patients infected with HIV-1. *Neurology, 85*(12), 1065-1073. doi:10.1212/WNL.0000000000001815

Ellis, R. J., Badiee, J., Vaida, F., Letendre, S., Heaton, R. K., Clifford, D., . . . CHARTER Group. (2011). CD4 nadir is a predictor of HIV neurocognitive impairment in the era of combination antiretroviral therapy. *AIDS, 25*(14), 1747-1751. doi:10.1097/QAD.0b013e32834a40cd

Ene, L., Marcotte, T. D., Umlauf, A., Grancea, C., Temereanca, A., Bharti, A., . . . Ruta, S. M. (2016). Latent toxoplasmosis is associated with neurocognitive impairment in young adults with and without chronic HIV infection. *Journal of Neuroimmunology, 299*, 1-7. doi:10.1016/j.jneuroim.2016.08.003

Fazeli, P. L., Crowe, M., Ross, L. A., Wadley, V., Ball, K., & Vance, D. E. (2014). Cognitive functioning in adults aging with HIV: A cross-sectional analysis of cognitive subtypes and influential factors. *Journal of Clinical Research in HIV AIDS and Prevention, 1*(4), 155-169. doi:10.14302/issn.2324-7339.jcrhap-13-191

Fellows, R. P., Byrd, D. A., & Morgello, S. (2013). Major depressive disorder, cognitive symptoms, and neuropsychological performance among ethnically diverse HIV+ men and

women. *Journal of the International Neuropsychological Society, 19*(2), 216-225. doi:10.1017/S1355617712001245

Fialho, R. M., Pereira, M., Mendonca, N., & Ouakinin, S. (2013). Depressive symptoms and neurocognitive performance among HIV-infected women. *Women & Health, 53*(2), 117-134. doi:10.1080/03630242.2013.767301

Fine, E. M., Delis, D. C., & Holdnack, J. (2011). Normative adjustments to the D-KEFS trail making test: Corrections for education and vocabulary level. *The Clinical Neuropsychologist, 25*(8), 1331-1344. doi:10.1080/13854046.2011.609838

Flensborg Damholdt, M., Shevlin, M., Borghammer, P., Larsen, L., & Ostergaard, K. (2012). Clinical heterogeneity in Parkinson's disease revisited: A latent profile analysis. *Acta Neurologica Scandinavica, 125*(5), 311-318. doi:10.1111/j.1600-0404.2011.01561.x

Fletcher, N. F., & McKeating, J. A. (2012). Hepatitis C virus and the brain. *Journal of Viral Hepatitis, 19*(5), 301-306. doi:10.1111/j.1365-2893.2012.01591.x

Foley, J. M., Ettenhofer, M. L., Kim, M. S., Behdin, N., Castellon, S. A., & Hinkin, C. H. (2012). Cognitive reserve as a protective factor in older HIV-positive patients at risk for cognitive decline. *Applied Neuropsychology: Adult, 19*(1), 16-25. doi:10.1080/09084282.2011.595601

Frndak, S. E., Smerbeck, A. M., Irwin, L. N., Drake, A. S., Kordovski, V. M., Kunker, K. A., . . . Benedict, R. H. (2016). Latent profile analysis of regression-based norms demonstrates relationship of compounding MS symptom burden and negative work events. *The Clinical Neuropsycholist, 30*(7), 1050-1062. doi:10.1080/13854046.2016.1200144

Gale, S. D., Brown, B. L., Erickson, L. D., Berrett, A., & Hedges, D. W. (2015). Association between latent toxoplasmosis and cognition in adults: A cross-sectional study. *Parasitology, 142*(4), 557-565. doi:10.1017/S0031182014001577

Garvey, L., Surendrakumar, V., & Winston, A. (2011). Low rates of neurocognitive impairment are observed in neuro-asymptomatic HIV-infected subjects on effective antiretroviral therapy. *HIV Clinical Trials, 12*(6), 333-338. doi:10.1310/hct1206-333

Giesbrecht, C. J., Thornton, A. E., Hall-Patch, C., Maan, E. J., Cote, H. C., Money, D. M., . . . Pick, N. (2014). Select neurocognitive impairment in HIV-infected women: Associations with HIV viral load, hepatitis C virus, and depression, but not leukocyte telomere length. *PLoS One, 9*(3), e89556. doi:10.1371/journal.pone.0089556

Gisslen, M., Price, R. W., & Nilsson, S. (2011). The definition of HIV-associated neurocognitive disorders: are we overestimating the real prevalence? *BMC Infectious Diseases, 11*, 356. doi:10.1186/1471-2334-11-356

Gomez, D., Power, C., & Fujiwara, E. (2018). Neurocognitive Impairment and Associated Genetic Aspects in HIV Infection. In: L. Cysique & S. Rourke (Eds.), *Neurocognitive complications of HIV-infection: Neuropathogenesis to implications for clinical practice for the Current Topics in Behavioral Neurosciences series*, Berlin: Springer. doi:10.1007/7854_2018_69

Gonzalez, E., Bamshad, M., Sato, N., Mummidi, S., Dhanda, R., Catano, G., . . . Ahuja, S. K. (1999). Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *Proceedings of the National Academy of Sciences, 96*(21), 12004-12009. doi:10.1073/pnas.96.21.12004

Gonzalez, E., Kulkarni, H., Bolivar, H., Mangano, A., Sanchez, R., Catano, G., . . . Ahuja, S. K. (2005). The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science, 307*(5714), 1434-1440. doi:10.1126/science.1101160

Gonzalez, E., Rovin, B. H., Sen, L., Cooke, G., Dhanda, R., Mummidi, S., . . . Ahuja, S. K. (2002). HIV-1 infection and AIDS dementia are influenced by a mutant MCP-1 allele linked to increased monocyte infiltration of tissues and MCP-1 levels. *Proceedings of the National Academy of Science, 99*(21), 13795-13800. doi:10.1073/pnas.202357499

Gould, T. J. (2010). Addiction and cognition. *Addiction Science and Clinical Practice, 5*(2), 4-14.

Gray, K. R., Aljabar, P., Heckemann, R. A., Hammers, A., Rueckert, D., & Alzheimer's Disease Neuroimaging Initiative. (2013). Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *NeuroImage, 65*, 167-175. doi:10.1016/j.neuroimage.2012.09.065

Guaraldi, G., Orlando, G., Zona, S., Menozzi, M., Carli, F., Garlassi, E., . . . Palella, F. (2011). Premature age-related comorbidities among HIV-infected persons compared with the general population. *Clinical Infectious Diseases, 53*(11), 1120-1126. doi:10.1093/cid/cir627

Gupta, S., Bousman, C. A., Chana, G., Cherner, M., Heaton, R. K., Deutsch, R., . . . Everall, I. P. (2011). Dopamine receptor D3 genetic polymorphism (rs6280TC) is associated with rates of cognitive impairment in methamphetamine-dependent men with HIV: Preliminary findings. *Journal of Neurovirology, 17*(3), 239-247. doi:10.1007/s13365-011-0028-3

Hair, J. F. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, N.J.: Pearson Prentice Hall.

Hariharan, S., Tirodkar, S., Porwal, A., Bhattacharya, A., & Joly, A. (2017). Random forest-based prospectivity modelling of greenfield terrains using sparse deposit data: An example from the Tanami Region, Western Australia. *Natural Resources Research, 26*(4), 489-507.

Haas, D. W., Ribaudo, H. J., Kim, R. B., Tierney, C., Wilkinson, G. R., Gulick, R. M., . . . Acosta, E. P. (2004). Pharmacogenetics of efavirenz and central nervous system side effects: An Adult AIDS Clinical Trials Group study. *AIDS, 18*(18), 2391-2400.

Heaton, R. K., Clifford, D. B., Franklin, D. R., Jr., Woods, S. P., Ake, C., Vaida, F., . . . CHARTER Group. (2010). HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. *Neurology, 75*(23), 2087-2096. doi:10.1212/WNL.0b013e318200d727

Heaton, R. K., Cysique, L. A., Jin, H., Shi, C., Yu, X., Letendre, S., . . . HNRC Group. (2008). Neurobehavioral effects of human immunodeficiency virus infection among former plasma donors in rural China. *Journal of Neurovirology, 14*(6), 536-549. doi:10.1080/13550280802378880

Heaton, R. K., Franklin, D. R., Ellis, R. J., McCutchan, J. A., Letendre, S. L., Leblanc, S., . . . HNRC Group. (2011). HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: Differences in rates, nature, and predictors. *Journal of Neurovirology, 17*(1), 3-16. doi:10.1007/s13365-010-0006-1

Heaton, R. K., Franklin, D. R., Jr., Deutsch, R., Letendre, S., Ellis, R. J., Casaletto, K., . . . CHARTER Group. (2015). Neurocognitive change in the era of HIV combination antiretroviral therapy: The longitudinal CHARTER study. *Clinical Infectious Diseases, 60*(3), 473-480. doi:10.1093/cid/ciu862

Heaton, R. K., Marcotte, T. D., Mindt, M. R., Sadek, J., Moore, D. J., Bentley, H., . . . HNRC Group. (2004). The impact of HIV-associated neuropsychological impairment on everyday functioning. *Journal of the International Neuropsychological Society, 10*(3), 317-331. doi:10.1017/S1355617704102130

Heringa, S. M., van den Berg, E., Reijmer, Y. D., Nijpels, G., Stehouwer, C. D., Schalkwijk, C. G., . . . Biessels, G. J. (2014). Markers of low-grade inflammation and endothelial dysfunction are related to reduced information processing speed and executive functioning in an older population - the Hoorn Study. *Psychoneuroendocrinology, 40*, 108-118. doi:10.1016/j.psyneuen.2013.11.011

Horvath, S., & Levine, A. J. (2015). HIV-1 infection accelerates age according to the epigenetic clock. *Journal of Infectious Diseases, 212*(10), 1563-1573. doi:10.1093/infdis/jiv277

Horvath, S., Phillips, N., Heany, S. J., Kobor, M. S., Lin, D. T., Myer, L., . . . Hoare, J. (2018). Perinatally acquired HIV infection accelerates epigenetic aging in South African adolescents. *AIDS*. doi:10.1097/QAD.0000000000001854

Hothorn, T., Buhlmann, P., Dudoit, S., Molinaro, A., & van der Laan, M. J. (2006). Survival ensembles. *Biostatistics, 7*(3), 355-373. doi:10.1093/biostatistics/kxj011

Ioannidis, J. P., Rosenberg, P. S., Goedert, J. J., Ashton, L. J., Benfield, T. L., Buchbinder, S. P., . . . Kostrikis, L. G. (2001). Effects of CCR5-Delta32, CCR2-64I, and SDF-1 3'A alleles on HIV-1 disease progression: An international meta-analysis of individual-patient data. *Annals of Internal Medicine, 135*(9), 782-795. doi:10.7326/0003-4819-135-9-200111060-00008

Iudicello, J. E., Morgan, E. E., Hussain, M. A., Watson, C. W.-M., & Heaton, R. K. (2019). HIV-Associated Neurocognitive Disorders. In M. L. Alosco & R. A. Stern (Eds.), *The Oxford Handbook of Adult Cognitive Disorders*. New York: Oxford University Press.

Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Taylor, M. J., Theilmann, R. J., Julaton, M. D., . . . CHARTER Group. (2011). Clinical factors related to brain structure in HIV: The CHARTER study. *Journal of Neurovirology, 17*(3), 248-257. doi:10.1007/s13365-011-0032-7

Jernigan, T. L., Gamst, A. C., Archibald, S. L., Fennema-Notestine, C., Mindt, M. R., Marcotte, T. D., . . . Grant, I. (2005). Effects of methamphetamine dependence and HIV infection on cerebral morphology. *American Journal of Psychiatry, 162*(8), 1461-1472. doi:10.1176/appi.ajp.162.8.1461

Joska, J. A., Combrinck, M., Valcour, V. G., Hoare, J., Leisegang, F., Mahne, A. C., . . . Stein, D. J. (2010). Association between apolipoprotein E4 genotype and human immunodeficiency virus-associated dementia in younger adults starting antiretroviral therapy in South Africa. *Journal of Neurovirology, 16*(5), 377-383. doi:10.3109/13550284.2010.513365

Kadri, F., LaPlante, A., De Luca, M., Doyle, L., Velasco-Gonzalez, C., Patterson, J. R., . . . Peruzzi, F. (2016). Defining plasma microRNAs associated with cognitive impairment in HIV-infected patients. *Journal of Cellular Physiology, 231*(4), 829-836. doi:10.1002/jcp.25131

Kaul, M., Garden, G. A., & Lipton, S. A. (2001). Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature, 410*(6831), 988-994. doi:10.1038/35073667

Kohler, S., Hamel, R., Sistermans, N., Koene, T., Pijnenburg, Y. A., van der Flier, W. M., . . . Ramakers, I. (2013). Progression to dementia in memory clinic patients without dementia: A latent profile analysis. *Neurology, 81*(15), 1342-1349. doi:10.1212/WNL.0b013e3182a82536

Kongs, S. K., Thompson, L. L., Iverson, G. L., & Heaton, R. K. (2000). *Winsconsin Card Sorting-64 Card Version*. Odessa, FL: Psychological Assessment Resources.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613.

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling.* Retrieved from http://dx.doi.org/10.1007/978-1-4614-6849-3

Lasselin, J., & Capuron, L. (2014). Chronic low-grade inflammation in metabolic disorders: relevance for behavioral symptoms. *Neuroimmunomodulation, 21*(2-3), 95-101. doi:10.1159/000356535

Letendre, S. (2011). Central nervous system complications in HIV disease: HIV-associated neurocognitive disorder. *Topics in Antiviral Medicine, 19*(4), 137-142.

Letendre, S., Marquie-Beck, J., Capparelli, E., Best, B., Clifford, D., Collier, A. C., . . . CHARTER Group. (2008). Validation of the CNS penetration-effectiveness rank for quantifying antiretroviral penetration into the central nervous system. *Archives of Neurology, 65*(1), 65-70. doi:10.1001/archneurol.2007.31

Levine, A. J., Quach, A., Moore, D. J., Achim, C. L., Soontornniyomkij, V., Masliah, E., . . . Horvath, S. (2016). Accelerated epigenetic aging in brain is associated with pre-mortem HIV-associated neurocognitive disorders. *Journal of Neurovirology, 22*(3), 366-375. doi:10.1007/s13365-015-0406-3

Levine, A. J., Reynolds, S., Cox, C., Miller, E. N., Sinsheimer, J. S., Becker, J. T., . . . Neuropsychology Working Group of the Multicenter AIDS Cohort Study. (2014). The longitudinal and interactive effects of HIV status, stimulant use, and host genotype upon neurocognitive functioning. *Journal of Neurovirology, 20*(3), 243-257. doi:10.1007/s13365-014-0241-y

Levine, A. J., Service, S., Miller, E. N., Reynolds, S. M., Singer, E. J., Shapshak, P., . . . Freimer, N. (2012). Genome-wide association study of neurocognitive impairment and dementia in HIV-infected adults. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 159B*(6), 669-683. doi:10.1002/ajmg.b.32071

Levine, A. J., Singer, E. J., Sinsheimer, J. S., Hinkin, C. H., Papp, J., Dandekar, S., . . . Shapshak, P. (2009). CCL3 genotype and current depression increase risk of HIV-associated dementia. *Neurobehavioral HIV Medicine, 1*, 1-7. doi:10.2147/NBHIV.S6820

Lojek, E., & Bornstein, R. A. (2005). The stability of neurocognitive patterns in HIV infected men: classification considerations. *Journal of Clinical and Experimental Neuropsychology, 27*(6), 665-682. doi:10.1081/13803390490918426

Maki, P. M., & Martin-Thormeyer, E. (2009). HIV, cognition and women. *Neuropsychology Review, 19*(2), 204-214. doi:10.1007/s11065-009-9093-2

Maki, P. M., Rubin, L. H., Valcour, V., Martin, E., Crystal, H., Young, M., . . . Anastos, K. (2015). Cognitive function in women with HIV: Findings from the Women's Interagency HIV Study. *Neurology, 84*(3), 231-240. doi:10.1212/WNL.0000000000001151

Malan-Muller, S., Hemmings, S. M., Spies, G., Kidd, M., Fennema-Notestine, C., & Seedat, S. (2013). Shorter telomere length - A potential susceptibility factor for HIV-associated

neurocognitive impairments in South African women. *PLoS One, 8*(3), e58351. doi:10.1371/journal.pone.0058351

Marquine, M. J., Sakamoto, M., Dufour, C., Rooney, A., Fazeli, P., Umlauf, A., . . . HNRC Group. (2016). The impact of ethnicity/race on the association between the Veterans Aging Cohort Study (VACS) Index and neurocognitive function among HIV-infected persons. *Journal of Neurovirology, 22*(4), 442-454. doi:10.1007/s13365-015-0411-6

Martin-Iguacel, R., Negredo, E., Peck, R., & Friis-Moller, N. (2016). Hypertension is a key feature of the metabolic syndrome in subjects aging with HIV. *Current Hypertension Reports, 18*(6), 46. doi:10.1007/s11906-016-0656-3

McCombe, J. A., Vivithanaporn, P., Gill, M. J., & Power, C. (2013). Predictors of symptomatic HIV-associated neurocognitive disorders in universal health care. *HIV Medicine, 14*(2), 99-107. doi:10.1111/j.1468-1293.2012.01043.x

McDermott, K. L., McFall, G. P., Andrews, S. J., Anstey, K. J., & Dixon, R. A. (2016). Memory resilience to Alzheimer's genetic risk: Sex effects in predictor profiles. *Journals of Gerontology. Series B: Psychological Sciences and Social Sciences*. doi:10.1093/geronb/gbw161

McGuinness, B., Barrett, S. L., McIlvenna, J., Passmore, A. P., & Shorter, G. W. (2015). Predicting conversion to dementia in a memory clinic: A standard clinical approach compared with an empirically defined clustering method (latent profile analysis) for mild cognitive impairment subtyping. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(4), 447-454. doi:10.1016/j.dadm.2015.10.003

McLachlan, G. J., & Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.

Mesaros, S., Rocca, M. A., Kacar, K., Kostic, J., Copetti, M., Stosic-Opincal, T., . . . Filippi, M. (2012). Diffusion tensor MRI tractography and cognitive impairment in multiple sclerosis. *Neurology, 78*(13), 969-975. doi:10.1212/WNL.0b013e31824d5859

Meyer, L., Magierowska, M., Hubert, J. B., Theodorou, I., van Rij, R., Prins, M., . . . Schuitemaker, H. (1999). CC-chemokine receptor variants, SDF-1 polymorphism, and disease progression in 720 HIV-infected patients. SEROCO Cohort. Amsterdam Cohort Studies on AIDS. *AIDS, 13*(5), 624-626. doi:10.1097/00002030-199904010-00015

Meyer, V. J., Rubin, L. H., Martin, E., Weber, K. M., Cohen, M. H., Golub, E. T., . . . Maki, P. M. (2013). HIV and recent illicit drug use interact to affect verbal memory in women. *Journal of Acquired Immune Deficiency Syndromes, 63*(1), 67-76. doi:10.1097/QAI.0b013e318289565c

Milanini, B., Ciccarelli, N., Fabbiani, M., Limiti, S., Grima, P., Rossetti, B., . . . Di Giambenedetto, S. (2016). Cognitive reserve and neuropsychological functioning in older HIV-infected people. *Journal of Neurovirology*. doi:10.1007/s13365-016-0426-7

Mindt, M. R., Byrd, D., Ryan, E. L., Robbins, R., Monzones, J., Arentoft, A., . . . Morgello, S. (2008). Characterization and sociocultural predictors of neuropsychological test performance in HIV+ Hispanic individuals. *Cultural Diversity and Ethnic Minority Psycholy, 14*(4), 315-325. doi:10.1037/a0012615

Modi, W. S., Lautenberger, J., An, P., Scott, K., Goedert, J. J., Kirk, G. D., . . . Winkler, C. (2006). Genetic variation in the CCL18-CCL3-CCL4 chemokine gene cluster influences HIV Type 1 transmission and AIDS disease progression. *American Journal of Human Genetics, 79*(1), 120-128. doi:10.1086/505331

Molsberry, S. A., Cheng, Y., Kingsley, L., Jacobson, L., Levine, A. J., Martin, E., . . . Neuropsychology Working Group of the MulticenterAIDS Cohort Study. (2018). Neuropsychological phenotypes among men with and without HIV disease in the multicenter AIDS cohort study. *AIDS, 32*(12), 1679-1688. doi:10.1097/QAD.0000000000001865

Morgan, E. E., Woods, S. P., Letendre, S. L., Franklin, D. R., Bloss, C., Goate, A., . . . CHARTER Group. (2013). Apolipoprotein E4 genotype does not increase risk of HIV-associated neurocognitive disorders. *Journal of Neurovirology, 19*(2), 150-156. doi:10.1007/s13365-013-0152-3

Morgan, E. E., Woods, S. P., Smith, C., Weber, E., Scott, J. C., Grant, I., & HNRP Group. (2012). Lower cognitive reserve among individuals with syndromic HIV-associated neurocognitive disorders (HAND). *AIDS and Behavior, 16*(8), 2279-2285. doi:10.1007/s10461-012-0229-7

Mukerji, S. S., Locascio, J. J., Misra, V., Lorenz, D. R., Holman, A., Dutta, A., . . . Gabuzda, D. (2016). Lipid profiles and APOE4 allele impact midlife cognitive decline in HIV-infected men on antiretroviral therapy. *Clinical Infectious Diseases, 63*(8), 1130-1139. doi:10.1093/cid/ciw495

Muthen, B. O., & Muthen, L. K. (1998-2015). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthen & Muthen.

Navia, B. A., Jordan, B. D., & Price, R. W. (1986). The AIDS dementia complex: I. Clinical features. *Annals of Neurology, 19*(6), 517-524. doi:10.1002/ana.410190602

Nightingale, S., Winston, A., Letendre, S., Michael, B. D., McArthur, J. C., Khoo, S., & Solomon, T. (2014). Controversies in HIV-associated neurocognitive disorders. The *Lancet Neurology, 13*(11), 1139-1151. doi:10.1016/S1474-4422(14)70137-1

Noorbakhsh, F., Ramachandran, R., Barsby, N., Ellestad, K. K., LeBlanc, A., Dickie, P., . . . Power, C. (2010). MicroRNA profiling reveals new aspects of HIV neurodegeneration: caspase-6 regulates astrocyte survival. *The FASEB Journal, 24*(6), 1799-1812. doi:10.1096/fj.09-147819

Nylund, K. L., Asparouhov, T., & Muthen, B. (2007). Decising on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*, 535-569.

Palella, F. J., Jr., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., . . . HIV Outpatient Study Investigators. (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *New England Journal of Medicine, 338*(13), 853-860. doi:10.1056/NEJM199803263381301

Panos, S. E., Hinkin, C. H., Singer, E. J., Thames, A. D., Patel, S. M., Sinsheimer, J. S., . . . Levine, A. J. (2013). Apolipoprotein-E genotype and human immunodeficiency virus-associated neurocognitive disorder: The modulating effects of older age and disease severity. *Neurobehavioral HIV Medicine, 5*, 11-22. doi:10.2147/NBHIV.S39573

Patel, S. M., Thames, A. D., Arbid, N., Panos, S. E., Castellon, S., & Hinkin, C. H. (2013). The aggregate effects of multiple comorbid risk factors on cognition among HIV-infected individuals. *Journal of Clinical and Experimental Neuropsychology, 35*(4), 421-434. doi:10.1080/13803395.2013.783000

Pemberton, L. A., Stone, E., Price, P., van Bockxmeer, F., & Brew, B. J. (2008). The relationship between ApoE, TNFA, IL1a, IL1b and IL12b genes and HIV-1-associated dementia. *HIV Medicine, 9*(8), 677-680. doi:10.1111/j.1468-1293.2008.00614.x

Pomara, N., Belzer, K. D., Silva, R., Cooper, T. B., & Sidtis, J. J. (2008). The apolipoprotein E epsilon4 allele and memory performance in HIV-1 seropositive subjects: Differences at baseline but not after acute oral lorazepam challenge. *Psychopharmacology, 201*(1), 125-135. doi:10.1007/s00213-008-1253-1

Public Health Agency of Canada. (2018). Summary: Estimates of HIV incidence, prevalence and Canada's progress on meeting the 90-90-90 HIV targets. *Retrieved from https://www.canada.ca*

Quasney, M. W., Zhang, Q., Sargent, S., Mynatt, M., Glass, J., & McArthur, J. (2001). Increased frequency of the tumor necrosis factor-alpha-308 A allele in adults with human immunodeficiency virus dementia. *Annals of Neurology, 50*(2), 157-162.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rickabaugh, T. M., Baxter, R. M., Sehl, M., Sinsheimer, J. S., Hultin, P. M., Hultin, L. E., . . . Jamieson, B. D. (2015). Acceleration of age-associated methylation patterns in HIV-1-infected adults. *PLoS One, 10*(3), e0119201. doi:10.1371/journal.pone.0119201

Robertson, K., Liner, J., & Meeker, R. B. (2012). Antiretroviral neurotoxicity. *Journal of Neurovirology, 18*(5), 388-399. doi:10.1007/s13365-012-0120-3

Robertson, K. R., Smurzynski, M., Parsons, T. D., Wu, K., Bosch, R. J., Wu, J., . . . Ellis, R. J. (2007). The prevalence and incidence of neurocognitive impairment in the HAART era. *AIDS, 21*(14), 1915-1921. doi:10.1097/QAD.0b013e32828e4e27

Rom, S., Rom, I., Passiatore, G., Pacifici, M., Radhakrishnan, S., Del Valle, L., . . . Peruzzi, F. (2010). CCL8/MCP-2 is a target for mir-146a in HIV-1-infected human microglial cells. *The FASEB Journal, 24*(7), 2292-2300. doi:10.1096/fj.09-143503

Rubin, L. H., Cook, J. A., Springer, G., Weber, K. M., Cohen, M. H., Martin, E. M., . . . Maki, P. M. (2017). Perceived and post-traumatic stress are associated with decreased learning, memory, and fluency in HIV-infected women. *AIDS, 31*(17), 2393-1401. doi:10.1097/QAD.0000000000001625

Rubin, L. H., Sundermann, E. E., Cook, J. A., Martin, E. M., Golub, E. T., Weber, K. M., . . . Maki, P. M. (2014). Investigation of menopausal stage and symptoms on cognition in human immunodeficiency virus-infected women. *Menopause, 21*(9), 997-1006. doi:10.1097/GME.0000000000000203

Sacktor, N., Skolasky, R. L., Seaberg, E., Munro, C., Becker, J. T., Martin, E., . . . Miller, E. (2016). Prevalence of HIV-associated neurocognitive disorders in the Multicenter AIDS Cohort Study. *Neurology, 86*(4), 334-340. doi:10.1212/WNL.0000000000002277

Saiyed, Z. M., Gandhi, N., Agudelo, M., Napuri, J., Samikkannu, T., Reddy, P. V., . . . Nair, M. P. (2011). HIV-1 Tat upregulates expression of histone deacetylase-2 (HDAC2) in human neurons: implication for HIV-associated neurocognitive disorder (HAND). *Neurochemisytry International, 58*(6), 656-664. doi:10.1016/j.neuint.2011.02.004

Sandkovsky, U., Podany, A. T., Fletcher, C. V., Owen, A., Felton-Coleman, A., Winchester, L. C., . . . Swindells, S. (2017). Impact of efavirenz pharmacokinetics and pharmacogenomics on neuropsychological performance in older HIV-infected patients. *Journal of Antimicrobial Chemotherapy, 72*(1), 200-204. doi:10.1093/jac/dkw403

Sanford, R., Fernandez Cruz, A. L., Scott, S. C., Mayo, N. E., Fellows, L. K., Ances, B. M., & Collins, D. L. (2017). Regionally specific brain volumetric and cortical thickness changes in HIV-infected patients in the HAART era. *Journal of Acquired Immune Deficiency Syndromes, 74*(5), 563-570. doi:10.1097/QAI.0000000000001294

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, 12*(5), 309. doi:10.1038/nrneurol.2016.53

Schretlen, D., Testa, S. M., & Pearlson, G. D. (2010). *Calibrated Neuropsychological Normative System*. Lutz, FL: Psychological Assessment Resources.

Schrier, R. D., Gupta, S., Riggs, P., Cysique, L. A., Letendre, S., Jin, H., . . . the HNRC Group. (2012). The influence of HLA on HIV-associated neurocognitive impairment in Anhui, China. *PLoS One, 7*(5), e32303. doi:10.1371/journal.pone.0032303

Segal, M. R., Barbour, J. D., & Grant, R. M. (2004). Relating HIV-1 sequence variation to replication capacity via trees and forests. *Statistical Applications in Genetics and Molecular Biology, 3*(1), 1-18. doi:10.2202/1544-6115.1031

Sheppard, D. P., Iudicello, J. E., Morgan, E. E., Kamat, R., Clark, L. R., Avci, G., . . . HNRP Group. (2017). Accelerated and accentuated neurocognitive aging in HIV infection. *Journal of Neurovirology, 23*(3), 492-500. doi:10.1007/s13365-017-0523-2

Simioni, S., Cavassini, M., Annoni, J. M., Rimbault Abraham, A., Bourquin, I., Schiffer, V., . . . Du Pasquier, R. A. (2010). Cognitive dysfunction in HIV patients despite long-standing suppression of viremia. *AIDS, 24*(9), 1243-1250. doi:10.1097/QAD.0b013e3283354a7b

Singh, D., Joska, J. A., Goodkin, K., Lopez, E., Myer, L., Paul, R. H., . . . Sunpath, H. (2010). Normative scores for a brief neuropsychological battery for the detection of HIV-

associated neurocognitive disorder (HAND) among South Africans. *BMC Research Notes, 3*, 28. doi:10.1186/1756-0500-3-28

Singh, K. K., Barroga, C. F., Hughes, M. D., Chen, J., Raskino, C., McKinney, R. E., & Spector, S. A. (2003). Genetic influence of CCR5, CCR2, and SDF1 variants on human immunodeficiency virus 1 (HIV-1)-related disease progression and neurological impairment, in children with symptomatic HIV-1 infection. *Journal of Infectious Diseases, 188*(10), 1461-1472. doi:10.1086/379038

Singh, K. K., Ellis, R. J., Marquie-Beck, J., Letendre, S., Heaton, R. K., Grant, I., & Spector, S. A. (2004). CCR2 polymorphisms affect neuropsychological impairment in HIV-1-infected adults. *Journal of Neuroimmunology, 157*(1-2), 185-192. doi:10.1016/j.jneuroim.2004.08.027

Singh, K. K., Hughes, M. D., Chen, J., & Spector, S. A. (2006). Impact of MCP-1-2518-G allele on the HIV-1 disease of children in the United States. *AIDS, 20*(3), 475-478. doi:10.1097/01.aids.0000200540.09856.58

Smith, A. (1973). *Symbol Digit Modalities*. Los Angeles, CA: Western Psychological Services.

Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., . . . O'Brien, S. J. (1997). Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. *Science, 277*(5328), 959-965. doi:10.1126/science.277.5328.959

Soontornniyomkij, V., & Achim, C. L. (2012). Aging. In H. E. Gendelman, I. P. Everall, & I. Grant (Eds.), *The Neurology of AIDS* (pp. 567-580). New York: The Oxfor University Press.

Spector, S. A., Singh, K. K., Gupta, S., Cystique, L. A., Jin, H., Letendre, S., . . . the HNRC Group. (2010). APOE epsilon4 and MBL-2 O/O genotypes are associated with neurocognitive impairment in HIV-infected plasma donors. *AIDS, 24*(10), 1471-1479. doi:10.1097/QAD.0b013e328339e25c

Stekhoven, D. J., & Bühlmann, P. (2011). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics, 28*(1), 112-118.

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics, 9*(1), 307. doi:10.1186/1471-2105-9-307

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods, 14*(4), 323-348. doi:10.1037/a0016973

Su, T., Schouten, J., Geurtsen, G. J., Wit, F. W., Stolte, I. G., Prins, M., . . . Schmand, B. A. (2015). Multivariate normative comparison, a novel method for more reliably detecting cognitive impairment in HIV infection. *AIDS, 29*(5), 547-557. doi:10.1097/QAD.0000000000000573

Sun, B., Abadjian, L., Rempel, H., Calosing, C., Rothlind, J., & Pulliam, L. (2010). Peripheral biomarkers do not correlate with cognitive impairment in highly active antiretroviral therapy-treated subjects with human immunodeficiency virus type 1 infection. *Journal of Neurovirology, 16*(2), 115-124. doi:10.3109/13550280903559789

Sundermann, E. E., Bishop, J. R., Rubin, L. H., Little, D. M., Meyer, V. J., Martin, E., . . . Maki, P. M. (2015). Genetic predictor of working memory and prefrontal function in women with HIV. *Journal of Neurovirology, 21*(1), 81-91. doi:10.1007/s13365-014-0305-z

Tang, V. M., Lang, D. J., Giesbrecht, C. J., Panenka, W. J., Willi, T., Procyshyn, R. M., . . . Barr, A. M. (2015). White matter deficits assessed by diffusion tensor imaging and cognitive dysfunction in psychostimulant users with comorbid human immunodeficiency virus infection. *BMC Research Notes, 8*, 515. doi:10.1186/s13104-015-1501-5

Thames, A. D., Briones, M. S., Magpantay, L. I., Martinez-Maza, O., Singer, E. J., Hinkin, C. H., . . . Levine, A. J. (2015). The role of chemokine C-C motif ligand 2 genotype and cerebrospinal fluid chemokine C-C motif ligand 2 in neurocognition among HIV-infected patients. *AIDS, 29*(12), 1483-1491. doi:10.1097/QAD.0000000000000706

Thames, A. D., Kuhn, T. P., Williamson, T. J., Jones, J. D., Mahmood, Z., & Hammond, A. (2017). Marijuana effects on changes in brain structure and cognitive function among HIV+ and HIV- adults. *Drug Alcohol Dependence, 170*, 120-127. doi:10.1016/j.drugalcdep.2016.11.007

Thames, A. D., Mahmood, Z., Burggren, A. C., Karimian, A., & Kuhn, T. P. (2016). Combined effects of HIV and marijuana use on neurocognitive functioning and immune status. *AIDS Care, 28*(5), 628-632. doi:10.1080/09540121.2015.1124983

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 Science, 102*(43), 15647-15652. doi:10.1073/pnas.0502548102

Tierney, S. M., Sheppard, D. P., Kordovski, V. M., Faytell, M. P., Avci, G., & Woods, S. P. (2017). A comparison of the sensitivity, stability, and reliability of three diagnostic schemes for HIV-associated neurocognitive disorders. *Journal of Neurovirology*. doi:10.1007/s13365-016-0510-z

Torgo, L., Ribeiro, R. P., Pfahringer, B., & Branco, P. (2013). *SMOTE for Regression*. Paper presented at the Progress in Artificial Intelligence, Berlin, Heidelberg.

Touw, W. G., Bayjanov, J. R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., & van Hijum, S. A. (2013). Data mining in the life sciences with Random Forest: A walk in the park or lost in the jungle? *Briefings in Bioinformatics, 14*(3), 315-326. doi:10.1093/bib/bbs034

Tripathi, A., Jerrell, J. M., Liese, A. D., Zhang, J., Rizvi, A. A., Albrecht, H., & Duffus, W. A. (2013). Association of clinical and therapeutic factors with incident dyslipidemia in a cohort of human immunodeficiency virus-infected and non-infected adults: 1994-2011. *Metabolic Syndrome and Related Disorders, 11*(6), 417-426. doi:10.1089/met.2013.0017

Trites, R. L. (1977). *Grooved Pegboard*. Ottawa, ON: Royal Ottawa Hospital.

Underwood, J., Robertson, K. R., & Winston, A. (2015). Could antiretroviral neurotoxicity play a role in the pathogenesis of cognitive impairment in treated HIV disease? *AIDS, 29*(3), 253-261. doi:10.1097/QAD.0000000000000538

Valcour, V., Shikuma, C., Shiramizu, B., Watters, M., Poff, P., Selnes, O. A., . . . Sacktor, N. (2004). Age, apolipoprotein E4, and the risk of HIV dementia: the Hawaii Aging with HIV Cohort. *Journal of Neuroimmunology, 157*(1-2), 197-202. doi:10.1016/j.jneuroim.2004.08.029

van Gorp, W. G., Hinkin, C., Satz, P., Miller, E., Weisman, J., Holston, S., . . . Dixon, W. (1993). Subtypes of HIV-related neuropsychological functioning: A cluster analysis approach. *Neuropsychology, 7*(1), 62-72.

van Rij, R. P., Portegies, P., Hallaby, T., Lange, J. M., Visser, J., de Roda Husman, A. M., . . . Schuitemaker, H. (1999). Reduced prevalence of the CCR5 delta32 heterozygous

genotype in human immunodeficiency virus-infected individuals with AIDS dementia complex. *Journal of Infectious Diseases, 180*(3), 854-857. doi:10.1086/314940

Vance, D. E., Fazeli, P. L., Dodson, J. E., Ackerman, M., Talley, M., & Appel, S. J. (2014). The synergistic effects of HIV, diabetes, and aging on cognition: Implications for practice and research. *Journal of Neuroscience Nursing, 46*(5), 292-305. doi:10.1097/JNN.0000000000000074

Vassallo, M., Durant, J., Biscay, V., Lebrun-Frenay, C., Dunais, B., Laffon, M., . . . Dellamonica, P. (2014). Can high central nervous system penetrating antiretroviral regimens protect against the onset of HIV-associated neurocognitive disorders? *AIDS, 28*(4), 493-501. doi:10.1097/QAD.0000000000000096

Villalba, K., Devieux, J. G., Rosenberg, R., & Cadet, J. L. (2015). DRD2 and DRD4 genes related to cognitive deficits in HIV-infected adults who abuse alcohol. *Behavioral and Brain Functions, 11*, 25. doi:10.1186/s12993-015-0072-x

Vivithanaporn, P., Asahchop, E. L., Acharjee, S., Baker, G. B., & Power, C. (2016). HIV protease inhibitors disrupt astrocytic glutamate transporter function and neurobehavioral performance. *AIDS, 30*(4), 543-552. doi:10.1097/QAD.0000000000000955

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57*(2), 307-333. doi:Doi 10.2307/1912557

Wade, B. S., Valcour, V. G., Wendelken-Riegelhaupt, L., Esmaeili-Firidouni, P., Joshi, S. H., Gutman, B. A., & Thompson, P. M. (2015). Mapping abnormal subcortical brain morphometry in an elderly HIV+ cohort. *Neuroimage: Clinical, 9*, 564-573. doi:10.1016/j.nicl.2015.10.006

Wendelken, L. A., Jahanshad, N., Rosen, H. J., Busovaca, E., Allen, I., Coppola, G., . . . Valcour, V. (2016). ApoE epsilon4 is associated with cognition, brain integrity, and atrophy in HIV over age 60. *Journal of Acquired Immune Deficit Syndromes, 73*(4), 426-432. doi:10.1097/QAI.0000000000001091

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test (4th ed.)*. Lutz, FL: Psychological Assessment Resources.

Winkler, C., Modi, W., Smith, M. W., Nelson, G. W., Wu, X., Carrington, M., . . . O'Brien, S. J. (1998). Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. *Science, 279*(5349), 389-393. doi:10.1126/science.279.5349.389

Winston, A., Arenas-Pinto, A., Stohr, W., Fisher, M., Orkin, C. M., Aderogba, K., . . . Dunn, D. (2013). Neurocognitive function in HIV infected patients on antiretroviral therapy. *PLoS One, 8*(4), e61949. doi:10.1371/journal.pone.0061949

Wojna, V., Skolasky, R. L., Hechavarria, R., Mayo, R., Selnes, O., McArthur, J. C., . . . Nath, A. (2006). Prevalence of human immunodeficiency virus-associated cognitive impairment in a group of Hispanic women at risk for neurological impairment. *Journal of Neurovirology, 12*(5), 356-364. doi:10.1080/13550280600964576

Woods, S. P., Moore, D. J., Weber, E., & Grant, I. (2009). Cognitive neuropsychology of HIV-associated neurocognitive disorders. *Neuropsychology Review, 19*(2), 152-168. doi:10.1007/s11065-009-9102-5

Woods, S. P., Rippeth, J. D., Frol, A. B., Levy, J. K., Ryan, E., Soukup, V. M., . . . Heaton, R. K. (2004). Interrater reliability of clinical ratings and neurocognitive diagnoses in HIV. *Journal of Clinical and Experimental Neuropsychology, 26*(6), 759-778. doi:10.1080/13803390490509565

World Health Organization. (2018). HIV Basics. *Retrieved from https://www.who.int*

Wright, E. J., Grund, B., Cysique, L. A., Robertson, K. R., Brew, B. J., Collins, G., . . . INSIGHT START Study Group. (2015). Factors associated with neurocognitive test performance at baseline: A substudy of the INSIGHT Strategic Timing of AntiRetroviral Treatment (START) trial. *HIV Medicine, 16*, 97-108. doi:10.1111/hiv.12238

Wright, E. J., Grund, B., Robertson, K., Brew, B. J., Roediger, M., Bain, M. P., . . . INSIGHT SMART Study Group. (2010). Cardiovascular risk factors associated with lower baseline cognitive performance in HIV-positive persons. *Neurology, 75*(10), 864-873. doi:10.1212/WNL.0b013e3181f11bd8

Wyczechowska, D., Lin, H. Y., LaPlante, A., Jeansonne, D., Lassak, A., Parsons, C. H., . . . Peruzzi, F. (2017). A miRNA signature for cognitive deficits and alcohol use disorder in persons living with HIV/AIDS. *Frontiers in Molecular Neuroscience, 10*, 385. doi:10.3389/fnmol.2017.00385

Xu, S., Huang, X., Xu, H., & Zhang, C. (2007). Improved prediction of coreceptor usage and phenotype of HIV-1 based on combined features of V3 loop sequence using random forest. *Journal of Microbiology, 45*(5), 441-446.

Xu, Z., Asahchop, E. L., Branton, W. G., Gelman, B. B., Power, C., & Hobman, T. C. (2017). MicroRNAs upregulated during HIV infection target peroxisome biogenesis factors: Implications for virus biology, disease mechanisms and neuropathology. *PLoS Pathogy, 13*(6), e1006360. doi:10.1371/journal.ppat.1006360

Yelamanchili, S. V., Chaudhuri, A. D., Chen, L. N., Xiong, H., & Fox, H. S. (2010). MicroRNA-21 dysregulates the expression of MEF2C in neurons in monkey and human SIV/HIV neurological disease. *Cell Death Dis, 1*, e77. doi:10.1038/cddis.2010.56

Zimmerman, P. A., Buckler-White, A., Alkhatib, G., Spalding, T., Kubofcik, J., Combadiere, C., . . . Murphy, P. M. (1997). Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Mol Med, 3*(1), 23-36.

APPENDIX

Appendix A. Summary of studies investigating genetic polymorphisms related to HIV progression and neurocognitive impairment in HIV from Gomez et al, 2018

Gene	Variant	Mechanisms	Effect on HIV progression and/or HAND	Evidence	
				Supportive	**Negative**
			Immune-related Genes		
CCR5	*CCR5 Δ-32* (rs333)	Truncated CCR5 receptor protein	Homozygosity for the *Δ-32* variant is associated with decreased disease progression. It has also been associated with low risk for AIDS dementia complex and protective against neurocognitive impairment. Nevertheless, some suggest the protective effect may only be found in individuals who developed AIDS prior to 1991	Zimmerman et al. (1997), Boven, van der Bruggen, van Asbeck, Marx, and Nottet (1999), Gonzalez et al. (1999), van Rij et al. (1999), Ioannidis et al. (2001), Singh et al. (2003) and Bol et al. (2012)	Singh et al. (2004) and Spector et al. (2010)
CCR2	*CCR2-V64I*	Linkage disequilibrium with *CCR5*. Heterologous receptor desensitization of CCR5 and CXCR4.	Associated with slowed progression to AIDS (by 2-4 years) in a cohort of over 3000 HIV-infected participants. However, *CCR2-V64I* was also associated with rapid progression to cognitive decline in adults.	Smith et al. (1997) Singh et al. (2004)	Singh, Hughes, Chen, and Spector (2006)
CCL3	rs1130371	Linkage disequilibrium with	TT genotype associated with a twofold increased risk for HAD	Levine et al. (2009)	

Gene	Variant	Mechanisms	Effect on HIV progression and/or HAND	Evidence	
				Supportive	Negative
		CCL3, *CCL4* and *CCL18*			
	rs1719134	High linkage disequilibrium with rs1130371	Carriers of the A allele demonstrate steeper cognitive decline than infected noncarriers and uninfected carriers	Levine et al. (2014)	
CCL3L1	*CCL3L1* low copy number	Dominant HIV suppressant. Modulation of CCR5 expression.	Possession of low copy numbers in conjunction with *CCR5 Δ-32* genotype associated with over a threefold increased risk for HAD	Gonzalez et al. (2005)	
CXCL12	*SDF1-3'-A* (rs1801157)	Increased expression of CXCL12	The *SDF1-3'A/A* genotype associated with delayed progression to AIDS in European American participants. Conversely, homozygosity for the A allele is also associated with faster disease progression and cognitive decline in African American children	Winkler et al. (1998), Singh et al. (2003), Modi et al. (2006) and Ding et al. (2018)	Meyer et al. (1999), van Rij et al. (1999), Ioannidis et al. (2001), Levine et al. (2009), Spector et al. (2010), and Levine et al. (2014)
	rs754618	Potential regulator role of CXCL12 levels	Associated with increased risk for disease progression in African American participants only	Modi et al. (2006) and Ding et al. (2018)	

Gene	Variant	Mechanisms	Effect on HIV progression and/or HAND	Evidence	
				Supportive	Negative
	rs2297630	Potential regulator role of CXCL12 levels	Associated with increased risk for disease progression in African American participants only	Modi et al. (2006) and Ding et al. (2018)	
CCL2	*CCL2-2578A>G* (rs1024611)	Increased transcription and production of CCL2. Associated with increased CSF CCL2.	Associated with rapid disease progression, decreased cognitive performance, and 4.7-fold increased risk for HAD	Gonzalez et al. (2002) and Levine et al. (2014)	Levine et al. (2009), Bol et al. (2012) and Thames et al. (2015)
PREP1	rs2839619	Preferentially binds to *CCL2-2578G* allele; influences CCL2 transcription	Potentially protective against HAND; however study replication is required	Bol et al. (2012)	Levine et al. (2012)
MBL2	*MBL-2 O/O* genotype	Structural deficit of MBL, thus reduced levels of functional MBL	The "O" alleles have been associated with higher risk for cognitive decline	Spector et al. (2010) and Levine et al. (2014)	
HLA	*HLA-DR*04*	Low CD4+ T-cell response	Associated with larger cognitive impairment at baseline and steeper cognitive decline over a year	Schrier et al. (2012)	
	HLA Class I alleles (*B*27, 57, 58, A*03, 33*)	Specify CD8+ T-cell response	Potentially neuroprotective; associated with higher neurocognitive performance at baseline and lower rates of cognitive decline over a year	Schrier et al. (2012)	

Gene	Variant	Mechanisms	Effect on HIV progression and/or HAND	Evidence	
				Supportive	Negative
TNFA	*TNFA-308G>A* (rs1800629)	Higher transcriptional activity leading to increased production of TNF-α	The A allele was associated with an increased risk for ADC/HAD	Quasney et al. (2001) and Pemberton, Stone, Price, van Bockxmeer, and Brew (2008)	Diaz-Arrastia, Gong, Kelly, and Gelman (2004), Levine et al. (2009) and Bol et al. (2012)
Genes Associated with Neurocognitive Performance					
APOE	*APOE* ε4 (rs429358; rs7412)	Dysregulated lipid and sterol metabolism. Enhanced HIV cell entry in vitro.	Conflicting results; some suggest *APOE* ε4 has been associated with HAND.	Corder et al. (1998), Valcour et al. (2004)[a], Burt et al. (2008), Pomara, Belzer, Silva, Cooper, and Sidtis (2008), Spector et al. (2010), Andres et al. (2011), Chang et al. (2011), Chang et al. (2014), Panos et al. (2013)[a], Mukerji et al. (2016)[a] and Wendelken et al. (2016)[a]	Sun et al. (2010), Joska et al. (2010), Morgan et al. (2013) and Becker et al. (2015)
COMT	*COMT Val158Met* (rs4680)	40% less metabolically active enzyme	*Val/Val* genotype associated with worse neurocognitive performance	Sundermann et al. (2015)	Levine et al. (2014)

Gene	Variant	Mechanisms	Effect on HIV progression and/or HAND	Evidence	
				Supportive	Negative
DRD2	rs6277	Alters dopamine D2 receptor expression	Associated with executive dysfunction	Villalba, Devieux, Rosenberg, and Cadet (2015)	Levine et al. (2014)
DRD4	*DRD4 VNTR-7*	Enhanced dopamine response	Associated with executive dysfunction	Villalba et al. (2015)	Levine et al. (2014)
DRD3	rs6280	Increased release of dopamine, however exact role not known	Linked to neurocognitive impairment in HIV-infected patients with concurrent stimulant addiction	Gupta et al. (2011)	
Genes Associated with Susceptibility to ART CNS Toxicity					
CYP	*CYP2B6* G516T	Reduced CYP2B6 function; well-known link to higher risk of neuropsychiatric complications with efavirenz	No association with neurocognition in two studies; however, higher levels of metabolite 8-OH-efavirenz have been associated with higher neurocognitive scores		Haas et al. (2004) and Sandkovsky et al. (2017)
	*CYP3A5 *1*	Faster CYP3A5 clearance of protease inhibitors	Risk for neurocognitive deficits unknown		

Abbreviations: ANI = asymptomatic neurocognitive impairment; HAD = HIV-associated dementia; HAND= HIV-associated neurocognitive impairment; MND = mild neurocognitive disorder. [a] Studies that found a moderating effect of older age (>50 years)

Appendix B. Summary of epigenetic studies investigating mechanisms associated with neurocognitive impairment in HIV in Gomez et al, 2018

Epigenetic mechanism	Affected processes associated with neurocognitive impairment in HIV	References
miRNA	Downregulation of multiple miRNAs associated with effector caspases involved in cell death pathways in HIV encephalitis	Noorbakhsh et al. (2010)
	Upregulation of miR-146a leading to translational suppression of pro-inflammatory cytokines associated with viral entry inhibition in HIV encephalitis	Rom et al. (2010)
	Upregulation of miRNAs (miR-500a-5p, miR-34c-3p, miR-93-3p, and miR-381-3p) that target peroxisomal genes in HAND patients	Xu et al. (2017)
	Upregulation/downregulation of miRNAs broadly associated with synaptic and neuronal functions	Yelamanchili, Chaudhuri, Chen, Xiong, and Fox (2010), Kadri et al. (2016), Asahchop et al. (2016) and Wyczechowska et al. (2017)
Histone modification	Upregulation of HDAC2 associated with the inhibition of translation of genes broadly synaptic and neuronal functions	Saiyed et al. (2011)
Telomere length	Shortened telomere length (LTL) in HIV-infected females compared to controls and LTL positively correlated to memory (Malan-Muller et al. 2013). However, results have not been replicated (Giesbercht et al. 2014)	Malan-Muller et al. (2013) and Giesbrecht et al. (2014)
DNA methylation	Increased DNA methylation in PBMCs and brain tissue in HAND patients compared to controls. Increased methylation associated with lower neurocognitive function in HIV-infected children.	Horvath and Levine (2015), Rickabaugh et al. (2015), Levine et al. (2016), Corley et al. (2016) and Horvath et al. (2018)