2006

# The effects of highly active antiretroviral therapy (HAART) on neuropsychological status in HIV-infection: A prospective study.

Susan E. Hayman-Abello
*University of Windsor*

# THE EFFECTS OF HIGHLY ACTIVE ANTIRETROVIRAL THERAPY (HAART) ON NEUROPSYCHOLOGICAL STATUS IN HIV-INFECTION: A PROSPECTIVE STUDY

by

Susan E. Hayman-Abello, M. A.

A Dissertation
Submitted to the Faculty of Graduate Studies and Research
through the Department of Psychology
in Partial Fulfilment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada
2006

*Your file Votre référence*
*ISBN: 978-0-494-17124-0*
*Our file Notre référence*
*ISBN: 978-0-494-17124-0*

## Abstract

The advent of highly active antiretroviral therapy (HAART) has resulted in significant reductions in HIV morbidity and mortality but the longitudinal effects on cognition are less well known. This longitudinal study examined the effects of highly active antiretroviral therapy (HAART) on cognitive performance in adults with HIV. Three hundred eighty-six adults representing all stages of HIV disease were tested at baseline and 180 of those subjects underwent a follow-up assessment. Subjects who were on HAART outperformed subjects on a non-HAART regimen, and those taking no antiretroviral medications, on the Grooved Pegboard Test, and were less impaired overall, at Time 1 and Time 2. Trends in the expected direction were found on other tests of psychomotor speed as well as on tests of attention and learning efficiency, however the group mean methodology failed to identify the breadth and depth of individual cognitive impairment across this sample. Longitudinal analyses using Reliable Change Indices indicated that 65% of the HIV-positive subjects had a stable cognitive profile during the follow-up time frame while 10% exhibited improved cognitive functioning and 26% deteriorated. Although there was a definite relationship between HAART and immune reconstitution, as well as HAART and viral suppression, a clear association between HAART and cognitive improvement was not identified. When the sample was restricted to individuals with CD4 cell counts < 200 μl, HAART was associated with cognitive change on selected measures of attention, psychomotor speed, learning efficiency and abstraction. Cognitive improvement was predicted by a model consisting of immune system response, initial cognitive impairment level, and estimated IQ. These findings suggest that the long-term effect of HAART on cognition is mediated by multiple factors. The dissociation between systemic and neurocognitive response to HAART that was identified in the current study supports a model of immune system-mediated cognitive disruption in HIV. As individuals with HIV survive longer, and potentially experience more cognitive impairment, more longitudinal research of this type is necessary to investigate the neurocognitive sequelae of HIV and response to HAART.

Acknowledgements

I am deeply indebted to many people for their assistance in the completion of this project. The expertise and generous assistance of Dr. Sean Rourke, and his laboratory staff and students, at the Neurobehavioural Research Unit at St. Michael's Hospital in Toronto was essential in the genesis and completion of this project. I am also grateful to the hundreds of research subjects and clinical patients who were the subjects for this work. Thank you to committee members Dr. Darren Fuerst and Dr. Mary Louise Drake for their comments and critique throughout. Thank you to Dr. Lisa Rapport for her insightful and supportive comments on the final product, as well as her ideas for extending this research. Thank you, of course, to Dr. Byron Rourke, not only for his support during this research project, but also for his expectations and encouragement throughout my graduate training.

Finally, and above all, I cannot thank enough my family and friends who provided unconditional support, in more ways than I could have imagined, during my many years as a graduate student.

Table of Contents

List of Tables

# List of Figures

# Introduction

Acquired Immune Deficiency Syndrome (AIDS) was first identified in the early 1980s as a life-threatening set of immune-related diseases of otherwise unknown cause. We now know that AIDS represents the late clinical stage of infection with the human immunodeficiency virus (HIV[1]). HIV causes damage to the immune system that, in turn, causes a person to become very susceptible to other infections and cancers that would otherwise be attacked and destroyed by a normally functioning immune system. The devastating neurological complications of HIV became evident early on as HIV-infected (or HIV+) individuals began to exhibit a constellation of severe and progressive neurological and neuropsychological deficits, labelled HIV-associated dementia (HAD; Navia, Cho, Petito, & Price, 1986; Navia, Jordan, & Price, 1986).

More recently, a distinction has been made between severe neurocognitive impairment (i.e., "dementia") typically evident in later-stage HIV-infection and AIDS, and a constellation of more subtle and varied neurocognitive impairments that may be observed in early stages of HIV-infection (e.g., HIV-1-associated minor cognitive motor disorder; American Academy of Neurology [AAN], 1991). Although the distinction between these two types of neuropsychological impairment has been made on a descriptive basis, much less is known about the course of neurocognitive impairment in HIV. Longitudinal investigations of neuropsychological performance are necessary to uncover answers to questions about natural history, treatment efficacy, and prognosis.

Before reviewing the literature related to existing longitudinal neuropsychological investigations of HIV disease, a summary of epidemiology, definitions, classifications, and test battery considerations will be presented. This will be followed by an overview of known neurocognitive deficits in HIV and information on disease course as gleaned from cross-sectional research efforts. Next, the methodological problems associated with longitudinal HIV research will be highlighted in parallel with a review of longitudinal HIV research findings. Finally, variables that potentially mediate the expression of HIV-

[1] The generic term HIV will be used throughout this paper although some authors refer to the virus that causes AIDS as HIV-1 to distinguish it from a variant, HIV-2, that is predominantly found in Africa and is not typically associated with neurological problems (McArthur, 1994).

related cognitive impairment will be presented, including the impact of antiretroviral medication and the relationship between disease markers and cognitive deficit.

*Epidemiology*

The incidence of positive HIV tests among Canadians slowly decreased from 2,983 in 1995 to 2,104 in 2000. This trend then reversed and incidence rose by 20% to 2,529 in 2004, with a prevalence of 57,674 as of December 31, 2004 (Public Health Agency of Canada, 2004) . At the end of 2004, 13,111 Canadians had died from this disease. These statistics also indicate that over approximately the last 10 years there has been a relative increase in HIV incidence for heterosexually exposed individuals and for women, and a relative decrease for intravenous drug users. Whereas the incidence rates in men who have sex with men decreased from 1995 to 1999, there was a recent increase in incidence for this risk group between 2000 and 2004.

Similarly, the rates of AIDS diagnoses revealed an unexpected change in recent years. In Canada, the peak AIDS diagnosis rate occurred in the early 1990s and then exhibited a steady decline, especially in the late 1990s. This trend was thought to be reflective of the widespread use of antiretroviral therapy, and the introduction of highly active antiretroviral therapy (HAART), which was thought to delay or prevent AIDS onset (Centers for Disease Control, 1998; Health Canada, 2001). However, the strong rate of decline in AIDS diagnosis notably slowed in 1997 and in 2000, there was an actual increase across the country. Rates of reported AIDS cases have been fairly steady in recent years (Public Health Agency of Canada, 2004) but it has been speculated that this potentially worrisome recent trend is related to antiretroviral therapy resistance and/or limited duration of drug efficacy (Health Canada, 2001).

*Definitions and Classification*

The CDC released a preliminary definition of AIDS in 1982 following an unusual outbreak of opportunistic infections such as Kaposi's sarcoma (KS) and Pneumocystis carinii pneumonia (PCP) in previously healthy persons who had significantly compromised immunity without identifiable cause (Centers for Disease Control, 1982). The cause of the disease was eventually designated as infection with the human immunodeficiency (retro)virus (HIV), and in the late 1980s the CDC advanced a

common, four-level system to classify HIV disease stage (Centers for Disease Control, 1986). The four, mutually exclusive and progressive HIV-infection groups included "Stage I", an acute phase in which HIV seroconversion produces transient symptoms only; "Stage II", an asymptomatic phase; "Stage III", which included patients with generalized lymphadenopathy; and "Stage IV", which included patients with any symptoms beyond lymphadenopathy (e.g., persistent fever or significant weight loss, neurologic disease, opportunistic infections). In 1987, the CDC revised the criteria to include HIV encephalopathy as a disease that is sufficient for the diagnosis of AIDS (Centers for Disease Control, 1987).

To better account for differing levels of immunosuppression among individuals infected with HIV, the definition was expanded to its most current form in which classification depends upon both clinical symptoms and levels of immunosuppression (Centers for Disease Control, 1992). Specifically, CD4+ T-cell (or CD4 cell) counts are categorized as follows: Category 1 = ≥500 cells/μl; Category 2 = 200 to 499 cells/μl; Category 3 = <200 cells/μl. Clinical categories are defined as follows: Category A includes asymptomatic individuals or those with minimal, transient symptoms of HIV-infection (e.g., generalized lymphadenopathy); Category B includes individuals with non-category C, immune-mediated symptoms (e.g., thrush, persistent fever or diarrhea); and Category C includes known AIDS-defining illnesses (e.g., PCP, KS, HIV-related encephalopathy). A diagnosis of AIDS is given for all Category C individuals, and for any individual who meets the immunosuppression Category 3 criteria (i.e., has less than 200 CD4 cells/μl).

*Neurological and Neuropsychological Sequelae*

Central nervous system (CNS) complications are thought to be the most common and disabling complications of HIV infection (Grant, Marcotte, Heaton, & the HNRC group, 1999; Martin, & Grant, 1994). HIV neuropathology can occur at any stage of HIV infection but is most apparent with severe immune suppression. Neurological damage is thought to arise from a combination of direct and indirect pathological mechanisms following the infiltration of HIV into the CNS. The majority of changes are thought to result from the release of toxic and inflammatory molecules by HIV-infected microglia and macrophages, with some direct neuronal damage also occurring (Bell, 2004;

Williams & Hickey, 2002). Subcortical white matter and basal ganglia regions of the brain are predominantly involved but reactive astrocytosis can be evident in white and grey matter, with neuronal loss maximally in the frontal cortex (Bell, 2004).

Neurocognitive deficits have been demonstrated throughout the course of HIV-infection, including during early infection (Heaton et al., 1995) and during later disease stages (Navia, Cho, Petito, & Price, 1986). Because nomenclature differences have at times contributed to confusion regarding the incidence, nature, and severity of HIV-associated neurocognitive deficits (van Gorp, Lamb, & Schmitt, 1993), a brief summary of current classification systems follows.

The term "AIDS dementia complex" (ADC; Navia, Jordan, & Price, 1986) was introduced to describe the progressive and severe nature of the cognitive, motor, and behavioural deficits observed among individuals with AIDS. However, later research suggested that some HIV+ individuals might exhibit cognitive decline prior to diagnosis with AIDS. In addition, the cognitive deficits themselves might not be severe enough to warrant a diagnosis of "dementia." Therefore, the term *AIDS dementia complex* was not technically accurate (Bornstein, 1994).

To clarify the potential differences in level of neurocognitive impairment, the American Academy of Neurology (AAN) introduced a distinction between the mild and more severe forms of cognitive impairment using the terms "HIV-1-associated minor cognitive/motor disorder" (MCMD) and "HIV-1-associated dementia complex" (HAD), respectively (AAN, 1991). As part of these definitions, specific kinds and levels of cognitive impairment were identified for each severity level (e.g., at least two of inattention, reduced information processing speed, memory problems, etc. must be present for HAD). According to this classification scheme, the impairments associated with the more severe form noticeably interfere with activities of daily living whereas the impairments associated with MCMD are less conspicuous (AAN, 1991).

HIV-associated dementia involves significant, progressive, and disabling deficits in memory, concentration, psychomotor speed, social behaviour, and activities of daily living, that are increasingly present in individuals with advanced stages of HIV-infection (Heaton et al., 1995; Navia, Cho, Petito, & Price, 1986; Navia, Jordan, & Price, 1986; Price & Brew, 1988). A milder form of HIV-related impairment, also involving deficits

in attention, speed of information processing, and learning efficiency has also been documented among individuals with less advanced HIV disease (Heaton et al., 1995) but its expression in otherwise asymptomatic HIV+ individuals has been the subject of some debate.

A 1995 review of cross-sectional neuropsychological research conducted between 1987 and 1994 found that there was significant disagreement on the presence and nature of neurocognitive symptoms in the Asymptomatic stage of HIV-infection (White, Heaton, Monsch, & the HIV Neurobehavioral Research Center [HNRC] Group, 1995). As a whole, evidence compiled from the 57 reviewed studies yielded strikingly contradictory findings in that approximately half of all studies reported "conclusive" or "suggestive" evidence of Asymptomatic stage cognitive deficit, whereas the other half did not detect any deficits in the Asymptomatic stage of HIV. Closer analysis of the 57 studies revealed a range of methodological differences that possibly invalidate a direct comparison between investigations.

A clearer finding emerged when studies were compared that reported individual impairment prevalence, rather than group differences (White, Heaton, Monsch, & the HNRC Group, 1995). There was a 35% impairment rate for asymptomatic HIV+ subjects compared to a 12% impairment rate in HIV-negative (HIV-) control groups, an interesting finding that supported the idea that group comparisons might mask individual change (van Gorp, Lamb, & Schmitt, 1993).

As part of their review, White and her colleagues also drew attention to other potential research confounds that might have masked or clouded the comparative results of the 57 studies. These included differences between subjects' mode of HIV-infection, type and breadth of neuropsychological assessment battery, sample size, and data analysis methods. It was concluded that controlling for mode of infection and using a suitably large neuropsychological assessment battery substantially influenced the outcome and conclusions about deficit presence or absence. For example, the presence of cognitive impairment was "conclusively" identified in the majority of tests that used "large" (i.e., >14 measure) batteries. White's recommendations for future investigations included better consideration of race, gender, subject selection methods, statistical issues associated with multiple comparisons, infection mode, and test battery size. Other

investigators have also drawn attention to the need to choose neuropsychological tests that are sufficiently sensitive and broad to detect subtle and varied deficits in the early stages HIV-infection (Bornstein, 1994; van Gorp, Lamb, & Schmitt, 1993).

Bornstein (1994) argued that the use of comprehensive batteries (e.g., Butters et al., 1990) has led to a better analysis of cognitive impairment. Using a large, comprehensive neuropsychological assessment battery, Heaton and the HNRC group identified mild neurocognitive deficits in all stages of HIV infection. They found that the risk of impairment increased with each successive disease stage from approximately 30% for asymptomatic HIV+ individuals, 44% for mildly symptomatic HIV+ individuals, to 55% for individuals with AIDS (Heaton et al., 1995).

A meta-analysis of 41 studies of HIV-related neuropsychological sequelae showed increasing frequency of impairment through disease stages from Asymptomatic to Symptomatic to AIDS (Reger, Welsh, Razani, Martin, & Boone, 2002). Reger et al. noted that motor functioning was the most notable area of impairment for Asymptomatic subjects and that problems with attention and concentration accompanied motor deficit as the most prominent problems in later disease stages. More frequent impairment in the latter disease stages was evident in terms of motor functioning, executive skills, and information processing speed, leading the authors to conclude that there was a progressive decline in these domains. The pattern of deficits and deficit progression was likened to the course of a subcortical dementia, with the greatest changes in cognitive function arising at the Symptomatic HIV stage. Signs of impairment related to cortical dysfunction (e.g., memory retention and visual construction problems) were not usually observed until the latter stages of disease.

The increasing rates of impairment found in cross-sectional analyses such as these suggest a progressive course of cognitive decline but it is important to note that HIV-dementia has been observed among HIV+ individuals who were otherwise Asymptomatic (Navia & Price, 1987; Selnes et al., 1995). Some HIV+ individuals show a relapsing/remitting pattern of deficit and not all individuals with AIDS develop HIV-associated dementia (Heaton et al., 1995). Longitudinal studies are required to draw conclusions about the course of HIV-associated cognitive decline, and to assess the impact, if any, of antiretroviral medication on cognitive functioning.

*Changes in Neuropsychological Functioning Over Time*

The timing and expression of neurocognitive impairment across the course of HIV infection has been poorly understood. Because of the poor prognosis associated with HAD (Goodwin, Pretsell, Chiswick, Egan, & Brettle, 1996; Price & Brew, 1988) and the possibility of predicting cognitive decline by way of neuropsychological assessment (Sacktor et al., 1996), it is important to understand more about the nature and course of neuropsychological impairment in patients with HIV. Efforts of this kind will improve our understanding of the disease process itself, and may provide a way of identifying individuals at risk for rapid progression who can be targeted for specific neuropsychological, psychosocial, and/or medical intervention.

Challenges arise when summarizing longitudinal HIV research, because of both general variations in longitudinal research paradigms and challenges unique to conducting research with an HIV-infected population. Review of the available longitudinal neuropsychological research for HIV-infection reveals conflicting findings that show some of the discrepancies noted in White's (1995) review of impairment prevalence. For example, of 17 selected articles on the topic, over half reported progressive cognitive decline across the course of HIV-infection, five argued in favour of cognitive improvement (often due to practice effects), and three reported no significant change across time. Closer inspection of the investigations reveals significant differences in methodology, statistical analysis, battery depth, population characteristics, disease stage, and test-retest interval. In the next section, some methodological issues associated with longitudinal HIV research will be highlighted in the context of reviewing the available literature.

*Methodological Issues in Longitudinal HIV Research*

There are a variety of potential methodological differences and considerations that must be evaluated when designing HIV natural history studies and when reviewing the existing literature. Patients enrolled in a longitudinal research investigation might progress from one HIV disease classification category to another and/or experience significant fluctuations in disease classification markers and health parameters. Subject attrition can be of particular concern when investigating impairment in a population that is susceptible to progressive dysfunction. Inability to complete the original test battery at

a follow-up assessment due to increased physical or cognitive impairment can reduce the number of subjects available for longitudinal comparison, in addition to the unfortunate reality of losing baseline subjects due to significant medical complications of HIV-infection and/or death.

Other complications can arise with regard to medications. Following the introduction of antiretroviral drug therapies, natural history longitudinal studies may include a mixed or uncontrolled group of medicated subjects. Furthermore, patients may begin or end treatment in the middle of the investigation period (Karlsen, Reinvang, & Froland, 1993), perhaps because of the known pill burdens and complications of treatment adherence or because of treatment failure. Subjects enrolled in studies may also have different, or unknown, HIV histories in terms of mode of infection, and time since seroconversion.

In addition to the varying methods employed in cross-sectional research for defining neuropsychological impairment (Bornstein, 1994; van Gorp, Lamb, & Schmitt, 1993), multiple methods are found for classifying neurocognitive change. However, it can be argued that few of the methods of quantifying change accounts adequately for statistical error, practice effects, or individual differences at baseline assessment (Temkin, Heaton, Grant, & Dikmen, 1999).

One means of assessing longitudinal change is through use of group mean change scores on neuropsychological tests (i.e., group mean score at follow-up minus group mean score at baseline). One of the risks of using group mean change score methodology to assess changes in longitudinal cognitive functioning, though, is its potential to mask subtle individual changes. The need to make cognitive comparisons on an individual basis has been identified as particularly important when assessing the subtle and/or mild cognitive deficits seen in HIV-infection (Bornstein, 1994; Hamby, Bardi, & Wilkins, 1997; Heaton et al., 1995; van Gorp, Lamb, & Schmitt, 1993; White, Heaton, Monsch, & the HNRC Group, 1995). Using a Global Deficit Score (i.e., an average of impairment ratings across a battery of neuropsychological tests) for assigning individual impairment has been recommended as both sensitive and specific for individuals with HIV cognitive impairment (Carey et al., 2004; Heaton, Grant, & Matthews, 1991; Heaton, Kirson, Velin, Grant, & the HNRC Group, 1994).

Various statistical and methodological techniques or corrections might be employed to establish the presence or absence of cognitive change by way of repeated neuropsychological assessment. The identification of reliable change in cognitive function can depend on a variety of factors including an individual's initial level of performance (McSweeny, Naugle, Chelune, & Luders, 1993; Temkin, Heaton, Grant, & Dikmen, 1999), test-retest interval and test properties (Dikmen, Heaton, Grant, & Temkin, 1999; Dodrill & Troupin, 1975; Hamby, Bardi, & Wilkins, 1997), as well as education, age, and regression to the mean (Dikmen, Heaton, Grant, & Temkin, 1999). When the potential confounds of these variables are not given consideration, cognitive change, or the lack thereof, may be incorrectly diagnosed.

A frequently overlooked variable in designing longitudinal change studies has to do with the effects of practice. It is possible that some of the conflicting results observed in the HIV longitudinal research to date, particularly with respect to change in Asymptomatic HIV+ individuals, may be partially explained by differences in addressing practice effects. Specific research addressing cognitive change in HIV-infection will be presented next while continuing to highlight some methodological issues that have contributed to a lack of consensus in this area of investigation.

*Practice Effects in Longitudinal HIV Research*

Saykin et al. (1991) compared HIV- and HIV+ groups on 10 neuropsychological tests at baseline and again one and one-half years later. All subjects showed increases in scores over time, regardless of HIV status or disease stage within the HIV+ group. The improvements in test scores were attributed to practice effects. A significant test by time interaction suggested that the relative influence of practice effects depended on the level of task difficulty, and level of education, but not on HIV status.

Karlsen, Reinvang, and Froland (1993) addressed the potential for cognitive decline during the Asymptomatic stage of HIV-infection by examining within-subject differences on neuropsychological functioning as tested four times over a two year period. Significant improvements in scores were noted on all tests from Time 1 through to Time 3 but not from Time 3 to Time 4. No subject was classified as neuropsychologically impaired at any time during the two-year follow-up. Based on the increasing and/or stable test scores obtained with their five-test screening battery, it was

concluded that HIV+ Asymptomatic individuals do not experience cognitive decline over a two-year period. However, the observed improvement in domain scores may have resulted solely from practice effects rather than from stable or improving neurocognitive functioning. Furthermore, nearly one-third of the subjects began drug therapy during the two-year period, which may have positively influenced their retest scores.

McCaffrey et al. (1995) anticipated problematic effects from the repeated administration of neuropsychological tests and addressed this issue as part of a larger investigation into the relationship between neuroimaging data and neuropsychological functioning in HIV. Significant practice effects for both HIV+ and HIV- groups were evident on the California Verbal Learning Test (CVLT) and on the Paced Auditory Serial Addition Task (PASAT) when subjects were tested twice, 7 to 10 days apart. However, on the Visual Search Test, only the HIV- control group showed full practice effects, in comparison with the Asymptomatic HIV+ group who showed practice effects on just one component of the test. The Symptomatic HIV+ group did not show any practice effects on the same test. The authors concluded that practice effects were a possible confound on certain neuropsychological tests and they recommended a multiple baseline procedure to correct for this factor. However, it also should be noted that there appeared to be a differential practice effect according to the presence, absence, and stage of HIV-infection, even over a short 7- to 10-day retest interval, suggesting possible HIV-related cognitive decline, particularly for the Symptomatic HIV+ group. Rather than test all subjects at least twice before assessing the effects of change, it might be more feasible to incorporate a statistical correction for potential practice effects, particularly when conducting research with clinical samples.

Stern et al. (1995) also demonstrated differential practice effects according to HIV status but went on to interpret the finding as evidence of progressive, cognitive decline relative to a HIV- control group. As part of a four and one half year follow-up on a cross-sectional analysis of well educated, HIV+ and HIV- homosexual/bisexual men (i.e., the Multicenter AIDS Cohort Study (MACS) project), cognitive changes over time were assessed using six condensed neuropsychological factor scores. The scores represented memory, attention, language, motor speed, executive functions, and orientation. The neuropsychological factor scores were lower at baseline for the HIV+ group compared to

the HIV- group, and all factor scores, except for orientation, improved over the four years for all groups. However, using a variation of multiple regression that took into account patients' scores over multiple repeated testing, this group of investigators was also able to document an attenuated rate of change in the HIV+ individuals. In other words, although their test scores were increasing over time, they were not increasing as quickly as those of the HIV- group. Furthermore, this reduced improvement rate in the HIV+ group was observed over a four-year period and was also significantly related to indices of disease severity including CD4 cell count, disease stage, and death.

The authors concluded that subtle, yet widespread, neurocognitive impairments were evident among HIV+ individuals in the early stages of the disease and that an attenuated rate of improvement over time signalled progressive cognitive decline, reflective of HIV penetration in the CNS (Stern et al., 1995). The interpretation of a reduced practice effect as evidence of cognitive decline remains questionable to some (Bornstein, 1994). However, Stern et al. provided evidence of differences in the rate of cognitive change over a long time period that were closely associated with markers of neurological functioning and of HIV disease progression, supporting the presence of early and detrimental effects of HIV on the CNS.

Practice effects have also been shown to have different effects on longitudinal scores with respect to outcome data. Selnes et al. (1995) also examined follow-up data from the large MACS project with a specific view to assessing cognitive changes before and after AIDS diagnosis in a group of HIV+ individuals. Significant practice effects on psychomotor speed and verbal memory tests were observed prior to diagnosis of AIDS for HIV+ non-demented individuals, but only on visual memory tests for demented HIV+ individuals. As expected, the HIV+ subjects with dementia exhibited significant decline, or trends in that direction, in psychomotor speed both before and after AIDS diagnosis. Non-demented subjects did not show any declines prior to clinical AIDS diagnosis but did show a significant decline in psychomotor speed, and trends in the direction of worsening verbal memory, after diagnosis with AIDS.

The authors concluded that HIV+ persons do not demonstrate cognitive decline before diagnosis with AIDS unless they are already diagnosed with dementia, and they only show mild cognitive decline after diagnosis when they have an AIDS-defining

illness. No worsening of cognitive function was observed for individuals whose AIDS diagnosis was based solely on immunosuppression (Selnes et al., 1995). However, practice effects may have obscured subtle cognitive changes. Although the MACS project data is generated from a large sample, it has also been criticized for using a potentially insensitive screening battery (Bornstein, 1994). Additional concerns relate to the lack of generalizability beyond its highly educated, Caucasian sample (Bornstein, 1994). Nonetheless, these investigators demonstrated that subtle cognitive decline coincided with AIDS diagnosis in non-demented individuals, and they confirmed a more rapid cognitive decline over time for HIV+ individuals who have dementia.

*Subtle and Varied Impairment*

The mild and varied nature of the cognitive impairments identified in Asymptomatic and Symptomatic HIV+ individuals presents another potential problem when attempting to assess cognitive change (Butters et al., 1990). White et al. (1995) reported that the size and breadth of a neuropsychological battery significantly influenced the sensitivity of any HIV neuropsychological investigation. Bornstein (1994) also argues that neuropsychological tests that might be sensitive to severe impairment (e.g., in HAD) may not be sufficiently sensitive to more subtle presentations of cognitive deficit (e.g., MCMD). Bornstein proposes that this lack of sensitivity in a screening battery explains much of the negative findings for early HIV stages in the MACS large-scale projects.

The minority of existing longitudinal neuropsychological investigations in HIV employ neuropsychological test batteries that would be classified as "large" by the standards of previous review work in this area (White, Heaton, Monsch, & the HNRC Group, 1995). As an exception, Villa et al. (1996) used a large battery to investigate the severity and pattern of any neuropsychological deficits in a group of HIV+ Asymptomatic patients to determine possible relationships between neuropsychological functioning and medical variables (e.g., immunosuppression). In addition, they looked for early neuropsychological predictors of HIV dementia and AIDS. They used 18 neuropsychological measures, grouped into six domains and compared scores for an Asymptomatic HIV+ group with those of an HIV- control group, both at baseline and12 to 18 months later. The presence of early cognitive deficits among HIV+ Asymptomatic individuals (as measured by number of impaired test scores) was predictive of further

decline in cognitive functioning and predicted faster progression to AIDS diagnosis, compared to Asymptomatic individuals with "normal" neuropsychological profiles. Early indicators of neuropsychological impairment among HIV+ Asymptomatic individuals were also useful in predicting further cognitive and medical decline.

Marra et al. (1996) also used a "large" battery to investigate the relationship between neuropsychological functioning and neurological variables on a longitudinal basis. All of the subjects' scores were ranked relative to the performance of all other subjects and averaged across seven neuropsychological domains to provide mean ranked scores for each individual. Ranked scores were used to minimize potential practice effects. Consistent with other cross-sectional studies, neuropsychological performance at baseline was worse for the HIV+ group than for the HIV-, homosexual and bisexual control group. However, no obvious worsening of performance was observed over an average two-year period, as assessed by changes in a global ranked score from Time 1 to Time 2. Furthermore, there was also no association between neurological variables such as increasing brain atrophy or cerebral spinal fluid HIV levels with neuropsychological change. It was concluded that, although initial neuropsychological performance was lower than that of a control group, no cognitive decline was apparent in their HIV+ sample followed over a two-year period.

To investigate the relationship between structural brain changes and neuropsychological changes in HIV disease, Hall et al. (1996) assessed the relationship between brain volume variables and 19 neuropsychological measures, collapsed into eight cognitive domains, in groups of Symptomatic and Asymptomatic HIV+ individuals over a period of approximately two and one-half years. Changes in neuropsychological test or domain score were represented by the difference between raw score at Time 1 and Time 2. A significant relationship was identified between decline in overall neuropsychological performance, mental flexibility, and gross motor ability, and increasing brain atrophy. Declining performance in gross motor ability, figural memory, tactile perception, fine-motor functioning, and in general speeded functioning, were specifically correlated with loss of subcortical brain volume. Although the neuroanatomical-neuropsychological relationship was identified for both Asymptomatic and Symptomatic HIV+ individuals, it was more pronounced in later disease stage, and

more evident in subcortical brain regions. However, there was no identifiable relationship between CD4 count and changes in brain volume. The associations between neurocognitive decline and brain atrophy among HIV+ individuals in this study contradicted those of Marra et al. (1996) and supported the idea of a progressive subcortical dysfunction related to advancing HIV disease, evident even in Asymptomatic stages of HIV disease.

Overall, the use of larger neuropsychological batteries that have sufficient breadth to be sensitive to the mild and varied impairments accompanying early HIV-infection appears necessary to adequately assess changes in longitudinal functioning. Four of the five "large" battery investigations reviewed here concluded that cognitive decline was apparent, even in the Asymptomatic stages of HIV.

*Individual Versus Group Change in the Pre-HAART era*

The results of available longitudinal neuropsychological HIV research can also be examined according to individual versus group research methodology. Typically, investigators analyze change by comparing the starting and end neuropsychological performance for HIV+ groups and HIV- groups. They then report significant change if group mean scores were significantly different between the groups at Time 2. The potential insensitivity of this method of analysis to subtle individual change in HIV, compared to using individual change scores to assess relationships, can be illustrated with the work of Burgess and colleagues (Burgess et al., 1994).

Burgess et al. (1994) found normal range performances for an Asymptomatic HIV+ group at baseline and follow-up, which did not differ from the mean score of an HIV- control group. Using group mean comparisons, WAIS Block Design decline was only detected for the subjects diagnosed with AIDS. However, when multiple regression equations were used to predict individual follow-up scores from initial test performance, 55% of the HIV+ subjects obtained significantly lower scores at follow-up than would have been expected based on their initial level of performance, compared to only 15% of the HIV- subjects. This significant difference in impairment prevalence remained even when demented HIV+ subjects were excluded, suggesting that at least a subgroup of Asymptomatic HIV+ individuals experience an otherwise unexpected decline in performance over a period of one year. These authors highlighted the discrepancy

between their results and those of other authors who made comparisons solely on the basis of group means (Saykin et al., 1991; Selnes et al., 1990; 1992), and emphasized the importance of using statistical methods that are sensitive to subtle individual changes over time.

Of note, the Burgess et al. investigation was one of few longitudinal HIV investigations that employed a within-subjects, individual change index such as those recommended for use in longitudinal studies (Heaton et al., 2001). Within-subjects change analysis methods in which statistically significant decline or improvement in test scores is established for each individual include the Reliable Change Index and variants (Jacobson and Truax 1991; Chelune et al. 1993) and regression models (e.g., McSweeny et al., 1993). The apparent sensitivity of these approaches should be considered when planning data analysis for future HIV investigations.

*Subject Characteristics*

Several factors related to individual participants in HIV studies can also have an influence on research outcomes. The detection or expression of HIV-related cognitive impairment can be affected by pre-existing differences in terms of mode of infection, level of education, and age (Bornstein, 1994; Ferrando et al., 1998; Valcour et al., 2004; van Gorp, Lamb, & Schmitt, 1993; White, Heaton, Monsch, & the HNRC Group, 1995). In addition, there are potential HIV-related differences in the nature and extent of systemic and health-related issues that accompany the various stages of HIV that create heterogeneous research samples.

Mauri et al. (1993) assessed a small group of HIV+, asymptomatic homosexual and bisexual males using 9 neuropsychological measures. The subjects' scores from Time 1, Time 2 (18 months later), and Time 3 (three years later) were compared and slight improvement was demonstrated on some measures at Time 3 compared with Time 1 performance. However, a significant decline in visuo-motor speed (i.e., WAIS Digit Symbol) over the three years was detected in the HIV+ asymptomatic subjects who went on to develop AIDS. This decline was related to CD4 cell count and thought to be reflective of generalized cerebral dysfunction. These findings are in line with those of Villa et al. (1996) showing a significant decline in cognitive functioning for a subset of Asymptomatic HIV+ individuals. However, in contrast to Hall et al. (1996) , Mauri and

colleagues did find a significant relationship between immunological status and neuropsychological decline.

Goodwin, Pretsell, Chiswick, Egan, and Brettle (1996) analyzed ten-year follow-up data on a group of British, HIV+ IV drug users with a goal of identifying medical and neuropsychological predictors of HAD. Problems with patient compliance and varying rates of dementia onset reduced the possibility of equivalent data for pre- and post-dementia assessments and resulted in significant subject attrition. However, the results obtained indicated significant changes in HIV+ subjects' auditory event-related potentials that pre-dated the clinical diagnosis of dementia. The available neuropsychological assessment data also revealed significant declines in information processing speed (e.g., Trail Making Test) and verbal memory (e.g., Auditory-Verbal Learning Test) that coincided with dementia diagnosis. The authors concluded that sustained and progressive cognitive impairment was observed in a subgroup of IV drug users who progressed from Asymptomatic HIV to AIDS, and in many cases death. In contrast, a matched control group of non-demented HIV+ IV drug users who were Asymptomatic, Symptomatic, or had AIDS, showed relatively stable functioning during the same time frame. Although the mechanisms are not yet understood, it seems clear from these results that even in the later stages of HIV disease, disease progression and accompanying physical and neurocognitive impairments are not uniform.

*Evidence for Subgroup Hypothesis*

Variations in the systemic and neurological course of HIV-infection (McArthur, 1994) and evidence of differing rates of cognitive decline within HIV+ individuals (Goodwin, Pretsell, Chiswick, Egan, & Brettle, 1996; Selnes et al., 1995) together point to the possibility of differing mechanisms of pathogenesis, and consequently differing subgroups. As reported earlier, Silberstein et al. (1993) used a prospective research design to evaluate changes in cognitive functioning in methadone-maintained HIV+ and HIV- subjects over a four-year period. They predicted that there would be more cognitive decline in individuals who developed HIV-related medical symptoms compared to those HIV+ individuals who remained asymptomatic during the 4-year research interval.

Subtle and varied declines in cognitive functioning were evident for some HIV+ individuals, particularly on psychomotor speed tests such as Trail Making Parts A and B,

and Finger Tapping. Interestingly, the declines were not consistently progressive according to disease stage. That is, HIV+ asymptomatic individuals showed significant declines in simple motor speed over the four-year period, whereas HIV+ symptomatic individuals maintained already low scores on simple motor speed and evidenced significant decline in more complex psychomotor ability over four years.

The investigators interpreted the inconsistent declines in functioning as reflective of differential CNS HIV processes but recognized the limitations posed by their small sample size and limited generalizability beyond an inner city, IV drug-using HIV risk group. Nonetheless, these investigators supported previous research results of cognitive decline in both symptomatic and symptomatic HIV+ groups. They also highlighted the sensitivity of psychomotor measures in detecting HIV-related cognitive decline.

Sacktor et al. (1996) employed a specific methodology to assess both one-time and sustained decline for HIV+ individuals (enrolled in the MACS project) who were assessed using a screening battery at six month intervals and followed for up to nine years. A "one-time" decline was defined as a decreased Time 2 test score that was greater than two standard deviations lower than their own best performance, and a "sustained decline" was said to occur if the retest score met this criteria on at least two consecutive assessments. Sixty-two percent of the 234 HIV+ individuals assessed using the "full" screening battery exhibited at least one episode of decline, and 32% of the same subjects experienced a sustained cognitive decline. It was noted that HIV+ subjects who evidenced cognitive decline were more likely to develop clinical dementia, and to develop it earlier, compared to HIV+ subjects who did not demonstrate the one-time decline. Furthermore, HIV+ subjects who demonstrated sustained psychomotor decline on the "mini" screening battery (i.e., Trail Making Test Part A and Symbol Digit Modalities) had an increased risk of dementia, AIDS, and death, compared to the HIV+ individuals who did not show a sustained psychomotor decline.

Given the interest in determining the natural course of neurocognitive impairment and the potential progression of MCMD to HAD, one potential method of analysis involves categorizing and following individuals by level of cognitive deficit, rather than by disease stage, and then predicting risk of other disease or death. The HNRC group followed 414 HIV+ men, tested annually or semi-annually, for up to seven years using a

large neuropsychological battery. Subjects were categorized at baseline as normal, neurocognitively impaired, or as meeting criteria for MCMD[2] (Ellis et al., 1997). The risk of death was calculated to be significantly higher for those HIV+ individuals with any level of cognitive impairment compared to those who showed normal neuropsychological functioning. Furthermore, the risk was statistically significant in subjects with MCMD but manifested only as a trend for the group with subclinical neuropsychological impairment. In other words, there was a higher risk of mortality for mildly neuropsychologically impaired (i.e., non-demented) subjects. In addition, the risk of mortality was not accounted for by immune status (i.e., CD4 cell count), viral load, or disease stage, suggesting that the course of HIV-related neurocognitive impairment and decline might be mediated by an independent neurocognitive factor. Of significant importance, this group linked a one-time occurrence of neurocognitive impairment with an increased risk of death, whereas Sacktor et al. (1996) did not find evidence of increased mortality risk unless subjects had experienced sustained decline.

*Antiretroviral Therapy*

The advent of antiretroviral therapy, and the more recent introduction of highly active antiretroviral therapy (HAART; Carpenter, Cooper, & Fischl, 2000), has significantly reduced HIV mortality and morbidity (Centers for Disease Control, 1998). Improvements in neurocognitive functioning have also been associated with pre-HAART era antiretroviral medications (Schmitt, Dickson, & Brouwers, 1994; Yarchoan et al., 1987).

Since 1997, the effects of HAART (also called potent or combination antiretroviral therapy) on cognitive impairment, deterioration, and recovery have been the focus of great interest. Although the connection between reduced morbidity and mortality in the post-HAART era has been established, the relationship between HAART and cognitive functioning is not as clear. The clinical benefits appear to be variable and the effects on the brain are not universally apparent. Current HAART investigations aim to

[2] Ellis et al. (1997) defined impairment as 1 *SD* below demographically corrected mean on any test; "Neuropsychological impairment" was identified by abnormal test performance that did not interfere with daily functioning, whereas MCMD diagnosis required significant disability; see Ellis et al. for their adaptations of the AAN (1991) diagnostic guidelines for further details.

disentangle the systemic, neurological, and neurocognitive effects of HIV and of antiretroviral medication.

Early HAART studies examined men and women at varying stages of HIV disease but tended to focus on symptomatic and late stage illness (Ferrando, Rabkin, van Gorp, Lin, & McElhiney, 2003; Sacktor et al., 2000; Tozzi et al., 1999), and on patients with at least mild cognitive problems (Suarez et al., 2001; Tozzi et al., 1999). With respect to advanced HIV disease and cognitive impairment, HAART has been shown to improve or prevent dementia (Marra et al., 2003; Sacktor et al., 2001a). More recent studies have documented improved neuropsychological performance among individuals on HAART with subtle cognitive impairment (Robertson et al., 2004).

Some longitudinal HAART research suffered from methodological limitations as well, including relatively small sample sizes (Sacktor et al., 2000; Suarez et al., 2001; Tozzi et al., 1999). In some of the MACS research over 200 subjects were analyzed, but only with respect to psychomotor ability (Sacktor et al., 1999). Often studies employed brief or screening batteries only (e.g., Cohen et al., 2001; Marra et al., 2003; Sacktor et al., 1999; Sacktor et al., 2000; Suarez et al., 2001; Tozzi et al., 1999) instead of a more comprehensive battery designed to assess a variety of neuropsychological domains (Ferrando et al., 1998; Tozzi et al., 1999). In addition, length of follow-up in some studies was as short as 4 weeks without accounting for potential practice effects (Marra et al., 2003).

With respect to neuropsychological abilities, several longitudinal post-HAART studies have suggested overall improvement in neuropsychological impairment status (Ferrando, Rabkin, van Gorp, Lin, & McElhiney, 2003; Marra et al., 2003; Tozzi et al., 1999). Aside from fairly consistent findings in terms of HAART-induced improvement psychomotor ability (Cohen et al., 2001; Ferrando et al., 1998; 2003; Sacktor et al., 1999; Sacktor et al., 2000; Suarez et al., 2001; Tozzi et al., 1999), there is some variability in what specific cognitive areas appeared to be most influenced by HAART status. Use of HAART was reported to be associated with improvements in attention and concentration (Ferrando et al., 1998; Tozzi et al., 1999) but less frequently in verbal fluency, executive function, and visual-spatial construction (Cohen et al., 2001; Tozzi et al., 1999). Ferrando et al. (1998) identified superior scores for HAART subjects on three measures

of verbal learning and memory, and Tozzi et al. (1999) reported a HAART-associated improvement in story recall over a 6 month period. However, other researchers failed to identify any HAART-related improvements in verbal learning and memory (Cohen et al., 2001; Tozzi et al., 2001). Suarez (2001) reported an improvement in verbal list learning that was finite compared to continuous improvements in other cognitive domains. Cohen (2001) reported that the majority of neuropsychological test variance could be accounted for by changes in processing speed.

The benefits of HAART and cognitive response to antiretroviral therapy seem to vary substantially between individuals (Letendre et al., 2004). There are many potential mediating variables on the relationship between HAART and neuropsychological performance including level of education (i.e., higher education associated with lack of cognitive impairment) (Ferrando et al., 1998), age, duration of HAART (Cohen et al., 2001; Ferrando et al., 1998; Suarez et al., 2001), and the ability of HAART regimens to cross the blood brain barrier (Cysique, Maruff, & Brew, 2004).

*Relationship to Laboratory Markers and Other Health Indicators*

Researchers have attempted to discern the links between cognitive decline and medical, immunological, or neuroanatomical variables, both prior to the introduction of HAART and after. The results have not been consistent, and they continue to challenge in the post-HAART era. Disease stage, including diagnosis with AIDS, has been associated with cognitive decline (Sacktor et al., 1996; Selnes et al., 1995; Stern et al., 1995; Villa et al., 1996) but not consistently (Ellis et al., 1997; Silberstein et al., 1993). Hall and colleagues (1996) found an association between cognitive decline and brain atrophy, particularly loss of subcortical brain volume, but no relationship was identified by Marra et al. (1996).

An association between CD4 cell count and cognitive worsening was identified in the pre-HAART era by some researchers (Stern et al., 1995) and more recently (Marcotte et al., 2003). On the other hand, some recent research has not supported as strong an association with CD4 immune functioning (McArthur et al., 2004), but has identified a relationship between other indicators of immune activity and cognitive deterioration (Sevigny et al., 2004). Increased plasma viral levels, particularly early elevations in viral activity, have been directly linked to changes in cognitive functioning (Ferrando et al.,

1998; Marcotte et al., 2003; Marra et al., 2003; Tozzi et al., 2005), as have elevations of viral RNA in cerebrospinal fluid (CSF) (Eggers, Hertogs, Sturenburg, van Lunzen, & Stellbrink, 2003; Letendre et al., 2004; Marcotte et al., 2003). However, other researchers have not found direct links between cognitive decline and plasma viral levels (Ellis et al., 1997) or a direct relationship between changes in CSF viral activity and cognitive change (Chang et al., 2003; Sevigny et al., 2004).

McArthur et al. (2004) reported that the relationships between cognitive functioning and markers of HIV disease or immune activation that were identified in the pre-HAART era seem to differ for current antiretroviral regimens. Cognitive decline had been identified as a significant predictor for risk of death (Ellis et al., 1997; Sacktor et al., 1996; Stern et al., 1995), but Tozzi et al. (2005) recently reported that the predictive relationship held only for those with unsuppressed viral load. McArthur et al. (2004) suggested that the observed changes in the relationships between neurological status and markers of HIV disease are related to changing neuropathological response to HIV in the post-HAART era. (i.e., reductions in HIV encephalitis frequency and severity). Brew (2004) also argues that different forms of dementia have evolved since the introduction of HAART. The diversity in research findings related to cognitive change in the post-HAART era highlights the potential role of individual response to treatment as a moderator of the antiretroviral-cognition interaction. Much of the work seems to point to dissociations in pathology and in treatment response between systemic and CNS infection.

To date, the findings across studies remain convoluted. Some investigations have provided evidence in support of a direct link between systemic HIV and neurocognitive decline but others argue in favour of separate systemic and CNS, or direct and indirect, mechanisms of pathogenesis. By elucidating the relationships between cognitive decline and measures of immune functioning and/or other biological markers of HIV disease, perhaps more will be understood about the pathogenesis and mechanisms of HIV on neurocognitive function.

*Summary*

There are many obstacles to understanding the natural history of HIV infection in humans. Some of the difficulties stem from the nature of HIV itself in terms of varied

modes of infection, heterogeneity of research samples, and poorly understood individual differences in disease expression and course. The mild and changing nature of cognitive impairment in the majority of individuals with HIV necessitates the use of a particularly broad and sensitive neuropsychological test battery. Longitudinal research with a population that progresses from well to medically unstable over an uncertain time period holds challenges for researchers with respect to subject recruitment and maintenance. There are also specific methodological and statistical issues associated with repeated neuropsychological testing that may affect the sensitivity and specificity of detecting true cognitive change.

Detecting changes in neuropsychological functioning in this population appears to depend upon several factors. These include: (a) the use of a sufficiently broad and sensitive neuropsychological test battery; (b) the use of statistical techniques that allow assessment of individual change; (c) the length of follow-up; (d) the proper use of demographic normative data, where possible; and (e) an appreciation of individual mediators, such as education, drug use, neurological injury and older age, which might positively or negatively affect individual response to treatment.

Overall, it appears that there are indeed mild and varied neurocognitive impairments associated with HIV that can present in individuals at any stage of the disease and are found to increase in prevalence with disease course. Typical impairments are evident in information processing speed, psychomotor speed, and attention, which may be indicators of generalized cerebral dysfunction. For a small subset of individuals, a more significant cognitive impairment interferes with daily living and is accompanied by a more rapid cognitive decline and poor prognosis. The relationships between these two subtypes of cognitive impairment, and between cognitive impairment and neuromedical indices, are not yet well understood. In the last eight to ten years, the advent of widespread antiretroviral use has improved mortality rates, and has made a significant impact with respect to preventing or postponing the onset of dementia. However, as individuals live longer with HIV infection and antiretroviral medications evolve, the prevalence of cognitive impairment is increasing (Sacktor et al., 2001a), and the use of antiretroviral medication will need to be investigated further with respect to long-term benefits and risks.

When reflecting on these investigations, it should be borne in mind that the cognitive deficits in question are typically mild. Levels of cognitive impairment that are sufficient to interfere with daily living do occur but are less common. Nonetheless, researchers are working towards identifying predictors and early markers of mortality, which include cognitive decline, dementia, and changes in medical health (i.e., disease stage). By understanding more about the course of HIV infection, and the relationship to antiretroviral therapy, we can better identify at-risk individuals who can be targeted for therapeutic intervention. Martin and Grant (1994) cautioned that "the neuropsychology of HIV presented, and continues to present, a formidable challenge to our research capability, ingenuity, and creativity" (Martin & Grant, 1994, p. 360) and this remains true a decade later. The obstacles highlighted in the current review should be addressed when designing and interpreting future research investigations.

*Current Research*

The major objectives of the current study were to determine: (1) the effect of HAART on specific neuropsychological abilities (i.e., attention, language, learning efficiency and memory, motor skills, psychomotor efficiency, and executive skills) compared to other antiretroviral regimens; (2) the dose-response relationship of HAART-mediated neuropsychological improvement over time; and (3) the impact of neuromedical comorbidity (e.g., history of head injury, substance abuse, liver complications, lower educational achievement and intellectual functioning) and demographic factors (e.g., estimated IQ) and HIV-related variables (e.g., CD4 count, viral load) on neuropsychological outcome. An additional goal of this research was to compare two methods of identifying cognitive deficits and longitudinal change among subjects with HIV (i.e., group mean comparison and individual change methodology).

*Hypotheses*

Based on the literature reviewed above pertaining to the effects of HAART on neuropsychological functioning in adults with HIV, the following hypotheses were formed:

1. HAART (a) will be associated with better neuropsychological performance, compared to non-HAART regimens, at baseline and follow-up, and (b) will be associated with improvements in neuropsychological functioning over time. These improvements are expected to be most evident on tests of psychomotor efficiency, which best reflect overall improvements in cerebral efficiency. In addition, some individuals on HAART will show improvements in other neuropsychological domains, particularly on tests of attention and learning ability.

2. Continuous HAART is predicted to be positively associated with improvement in neuropsychological functioning compared to treatment with non-continuous HAART and non-HAART medication regimens.

3. Differential changes in neuropsychological functioning over time will be influenced by the presence of neuromedical risk factors, and by some demographic variables, but less so by systemic markers of HIV disease. Specifically, it is predicted that participants with a history of neuromedical risk factors (e.g., history of head injury, substance abuse, liver pathology), and those with lower educational achievement or lower estimated IQ scores, will show no or minimal improvements on HAART. It is also predicted that there will be no relationship between neuropsychological improvement over time and changes in systemic immune functioning (e.g., as reflected by CD4 count and plasma viral load).

## Method

*Participants*

Four hundred and six adults with HIV infection completed a baseline (Time 1) neuropsychological evaluation between March 1996 and June 2004. Study participants were obtained through consecutive neuropsychology referrals to the Neurobehavioural Research Unit at St. Michael's Hospital in Toronto, Ontario from medical, infectious disease and psychiatric clinics (84.3%) and by recruitment for research (15.7%). Subjects were excluded from the current analysis of HAART effectiveness if they were involved in another, simultaneous clinical trial ($n$ = 20). The results of one-way analyses of variance (ANOVAs) showed no significant differences between recruitment sources in terms of age, education, IQ estimate, CD4 cell count, CDC93 stage, viral load status, or antiretroviral medication regimen.

A summary of the demographic characteristics of the Time 1 sample is presented in Table 1. Eighty-five percent of the subjects were known to be Caucasian with Afro-Canadians making up the next highest group (6.8%). Participants had a mean age of approximately 42 years and approximately 14 years of education. The mean estimated verbal IQ for the group was 115 ($SD$ = 7.5). Just over half of the subjects had an undetectable plasma viral load (51.0%). The CDC93 classifications of HIV infection were represented as follows: Asymptomatic (i.e., CDC93 A1, A2): 9.8%; Symptomatic (i.e., CDC93 B1, B2): 40.2%; AIDS diagnosis (i.e., CDC93 A3, B3, C1-3): 50.0%.

The sample was predominantly male (94%) and either gay or bisexual. Of the 240 subjects for whom mode of HIV infection was known, almost 80% ($n$ = 189) reported HIV acquisition through homosexual activity. Other single modes of infection reported by the subjects included heterosexual contact (5.0%), IV drug use (5.8%), and receipt of contaminated blood products (6.6%). There were 9.9% subjects who reported dual routes of infection or other causes. Non-parametric analyses with Bonferroni correction showed that subjects who used IV drugs had fewer years of schooling ($M$ = 10.5, $SD$ = 2.6) than subjects who identified homosexual ($M$ = 14.2, $SD$ = 2.7) or blood products ($M$ = 14.9, $SD$ = 3.0) transmission routes, H(4) = 21.32, $p < .001$; $U$ = 461.50. Subjects who identified IV drug use as their primary mode of transmission also had a higher mean initial CD4 cell count ($M$ = 491 $SD$ = 334) than the group of subjects who indicated a

homosexual mode of transmission ($M$ = 201, $SD$ = 189; H(4) = 13.93, $p$ = .008) but did not differ from the other groups. There were no differences between infection mode groups in terms of other HIV markers, antiretroviral status, age, estimated IQ, or baseline level of cognitive functioning (i.e., Global Deficit Score; see below for calculation). Therefore, mode of infection was not deemed to be a significant confound within the current sample and the remaining analyses were conducted with subjects representing all modes of infection combined.

Table 1.
*Demographic Characteristics of the Total Time 1 Sample (N = 386)*

| Variable | Mean | Standard Deviation |
|---|---|---|
| Age (years) | 41.6 | 8.4 |
| Education (years) | 13.9 | 2.8 |
| Estimated IQ (ANART)[a] | 115.0 | 7.5 |

| | Percentage |
|---|---|
| Race – Caucasian | 84.6 |
| Gender – Male | 93.5 |
| Undetectable viral load (<500)[b] | 51.0 |
| Asymptomatic (CDC93 A1, A2)[c] | 9.8 |
| Symptomatic (CDC93 B1, B2)[c] | 40.2 |
| AIDS (CDC93 A3, B3, C1-C3)[c] | 50.0 |

*Note*. ANART = Adult North American Reading Test. CDC93 = Centre for Disease Control 1993 HIV Classification criteria (A = Asymptomatic; B = Symptomatic; C = AIDS. A1, B1, C1 = CD4 count >500 CD4 cells; A2, B2, C2 = 200 to 499 CD4 cells; A3, B3, C3 = <200 CD4 cells).
[a] $n = 360$. [b] $n = 353$. [c] $n = 328$.

Of the 386 subjects who underwent one assessment between March 1996 and June 2004, 180 had also completed the second assessment by June 2005. Efforts were made to contact subjects between 12 and 18 months after the initial testing for a follow-up evaluation. Some subjects had moved or were otherwise unavailable for study participation at that time. In other cases, subjects rescheduled or cancelled appointments because of HIV-related and other illnesses and were unable to complete the second assessment within the planned time frame. Twenty-two subjects declined follow-up testing and 21 subjects died before follow-up appointments were conducted.

There were no group differences between follow-up participants and those who completed only a single assessment in terms of markers of HIV illness (e.g., CDC stage, CD4 count, viral load) and antiretroviral medication status. Subjects who participated in the second assessment session had completed on average one more year of education ($t = 3.33$, $p = .001$) and had slightly higher estimated IQ scores ($U = 12734.5$, $p < .001$; see Table 2). The differences are not considered clinically significant and suggest that the

follow-up group was a reasonable representation of the subjects who underwent baseline testing. However, the differences between the two groups will be borne in mind when generalizing the current findings to the broader HIV population.

Table 2.
*Comparison Demographic Data for Follow-Up Sample and Single Assessment Sample*

| | Subjects with Second Assessment ($n$ = 180) | | Subjects with Single Assessment ($n$ = 386) | |
|---|---|---|---|---|
| | Mean | *SD* | Mean | *SD* |
| Age (years) | 42.3 | 8.0 | 41.0 | 8.7 |
| Education (years)[a] | 14.4 | 2.6 | 13.5 | 2.9 |
| Estimated IQ (ANART) [a] | 116.5 | 7.2 | 113.6 | 7.6 |
| Recent CD4 Count | 391.4 | 263.5 | 365.8 | 17.8 |
| | % | | % | |
| CDC93 Stage = AIDS | 51.2 | | 48.8 | |
| Undetectable Viral Load (< 500) | 46.4 | | 55.1 | |
| Medication status (taking HAART at T1) | 65.6 | | 68.1 | |

*Note.* ANART = Adult North American Reading Test.
[a] significant difference between groups at $p < .01$

Participants were contacted for follow-up assessments approximately eighteen months after the initial neuropsychological evaluation. Variations in test-retest interval occurred due to a variety of reasons including clinical change requiring more rapid re-assessment, specific referral requests, and scheduling difficulties. The average time between first and second assessments was 22.2 months (*SD* = 11.7) and the overall follow-up interval ranged from 4 months to 5 years. The majority of subjects (80%) were evaluated between 6 months and 2.5 years after the initial assessment. One-way ANOVA and chi-square tests showed no significant relationships between test-retest interval and demographic variables, markers of HIV infection (e.g., CDC stage, CD4 count, viral load), or neuropsychological test scores. Also, differences in test-retest interval were not

related to the referral source, medication status, or to a history of neuromedical risk factors.

Among subjects who were tested twice, 44 subjects presented with a potential medical or psychiatric confound (i.e., prior neurological disease, significant developmental problems, CNS opportunistic infection, moderate or worse traumatic brain injury, or recent substance abuse or dependence). Preliminary analyses of Time 1 neuropsychological performance did not produce significantly different results when these subjects were included or excluded. However, when a larger sample size became available, differences between the two groups emerged. Specifically, the effects of antiretroviral medication on neuropsychological performance were better illuminated when subjects with neuromedical comorbidity were excluded from analysis. Therefore, they were excluded from all of the analyses and this fact will be considered in the interpreting the generalizability of these results to the broader HIV population.

*Antiretroviral Classification*

At the time of baseline testing, antiretroviral classification was available for all but two of the baseline subjects. At that time, 23.4% ($n$ = 90) of the subjects were not taking antiretroviral medications and 76.5% ($n$ = 294) were on some form of antiretroviral therapy (ART). The ART regimen consisted of various combinations of nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), and protease inhibitors (PIs). HAART regimens have been defined with some variation in the neuropsychological literature but most commonly refer to a combination of three or more antiretrovirals. For the purposes of this study, HAART referred to any antiretroviral drug regimen that included at least three drugs in combinations as follows: (a) $\geq$ 2 NRTIs, together with 1 NNRTI and/or 1 PI; or (b) 1 drug from each of the three classes.

Subjects were grouped according to HAART status (on or off) at both Time 1 and Time 2. Time 1 medication status was available for 384 subjects. Of that group, 257 subjects (66.9%) were on HAART, 90 subjects (23.4%) were taking no antiretroviral medications, and 37 subjects (9.6%) were taking a non-HAART or "other" antiretroviral regimen (including saquinavir-based combinations, whose potency has been described as sub-optimal;Yeni et al., 2002). Of the 180 subjects who were assessed at Time 2,

medication status was known for all but four subjects. Within the follow-up group, there were 123 subjects (68.3%) on HAART, 40 subjects (22.2%) taking no antiretrovirals, and 13 subjects (7.2%) taking a non-HAART or "other" antiretroviral regimen.

Four subgroups were generated in an attempt to capture the changes or consistency in HAART status over time, labelled according to their status at Time 1 and Time 2: HAART off/off; HAART off/on; HAART on/off; and HAART on/on. Of the 11 subjects who had stopped taking HAART by Time 2, 4 subjects had switched to sub-optimal combination regimens (e.g., 3 NRTIs or saquinavir-based regimen) and 1 reported that his previous ARV regimen had failed and he "wasn't ready" to begin another regimen. The reasons were not available for 7 subjects but typical reasons could include self-imposed treatment interruptions, problems with tolerance and/or treatment failure (i.e., virologic, immunologic, or clinical failure; Yeni et al., 2002).

Preliminary data had indicated comparable cell sizes for these four categories (i.e., off/off - 20%, off/on - 25%, on/off - 20%, on/on - 35% respectively) but analysis of the final data set revealed a much larger inequality between these groups (i.e., off/off - 22%, off/on -14%, on/off - 8%, on/on 56%). Therefore, the two non-continuous HAART groups were collapsed into one group that contained subjects who were taking HAART at some point, but not continuously from Time 1 to Time 2. Subjects' longitudinal medication status was redefined to the following: (a) taking HAART continuously (i.e., HAART at T1 & T2), labelled HAARTc ($n$ = 99, 56.3%); (b) taking HAART non-continuously (i.e., HAART at T1 or T2) labelled HAARTnc (n = 39, 22.2%); or (c) taking a non-HAART regimen [including no antiretrovirals] at Time 1 and Time 2, labelled non-HAART ($n$ = 38, 21.6%). For certain analyses, the groups were collapsed further into HAART-any ($n$ = 138) versus non-HAART ($n$ = 38) to meet the demands of statistical procedures.

It was planned that the duration of HAART could be included as an independent variable but this information was not accessible. Instead, two alternate approaches to assessing potential dose-response and potency relationships were investigated. First, a comparison was made between subjects taking HAART both at baseline and follow-up (i.e., continuous HAART) versus those on HAART only at one time or the other (i.e., non-continuous HAART). Second, the percentage of antiretroviral drugs in each subjects'

drug regimen most likely to penetrate the blood-brain barrier was calculated and used as a representation of CNS potency. For this calculation, the following drugs were considered capable of penetrating the CNS as reflected by a high CSF penetrance index: Zidovudine, Stavudine, Efavirenz, Nevirapine, Abacavir, and Indinavir (Sacktor et al., 2001b; Evers et al., 2004).

*HIV Moderators*

To confirm the validity of the antiretroviral classification scheme, subjects' responses to medication were assessed in terms of immune system functioning (i.e., initial CD4 count and change in CD4 cell count over time) and viral activity (i.e., initial status and change in undetectable viral status). Specifically, subjects were stratified into six groups based on their initial CD4 cell count at Time 1 (i.e., high [CD4 ≥ 200] or low [CD4 < 200]) and on their change in CD4 cell count from Time 1 to Time 2 (i.e., increase of at least 50 cells, decrease of at least 50 cells, or minimal/no change; see Table 3). The value for initial disease stage was extrapolated from the CDC93 categories and requirement for a change of 50 cells was derived from a description of the immune system's predicted response to antiretroviral medication (Yeni et al., 2002). Thus, six categories reflecting a combination of initial disease stage and change over time were formed as follows:

Table 3.
*CD4 Stratification by Initial Counts and Changes over Time*

| **Category** | **Time 1 CD4** | **Change in CD4** | **Percentage of Subjects (*n* = 197)** |
|---|---|---|---|
| Low Risers | <200 cells/μl | Increase by >=50 cells | 11.5% |
| Low Decliners | <200 cells/μl | Decrease by >=50 cells | 3% |
| Low Stable | <200 cells/μl | Change less than 50 cells | 9.0% |
| High Risers | >=200 cells/μl | Increase by >=50 cells | 30.5% |
| High Decliners | >=200 cells/μl | Decrease by >=50 cells | 24.0% |
| High Stable | >=200 cells/μl | Change less than 50 cells | 20.5% |

For the purposes of assessing antiretroviral efficacy, the groups were amalgamated on the basis of improvement, worsening or maintenance of threshold CD4 cell and viral copy levels. That is, Low Risers, High Risers, and High Stable subjects were combined into one group of subjects who could be said to have improving CD4 cell counts or maintenance of threshold levels (i.e., Better/High Stable group). A second group (i.e., Worse/Low Stable) group was comprised of the remaining categories.

Significantly more subjects who were on HAART continuously had threshold improvements in immune functioning (i.e., increase to 200 cells/μl) or maintained immune functioning above the threshold level (i.e., 200 cells/μl), compared to subjects not on HAART, and compared to subjects who took HAART non-continuously ($\chi^2$ (2, $n$ = 172) =13.9; $p$ = .001; see Figure 1). Similarly, more subjects who were on HAART acquired or maintained an undetectable viral load status (i.e., <500 copies/ml) compared to subjects in the other two medication groups ($\chi^2$ (2, $n$ = 165) =23.9; p < .001; see Figure 2).

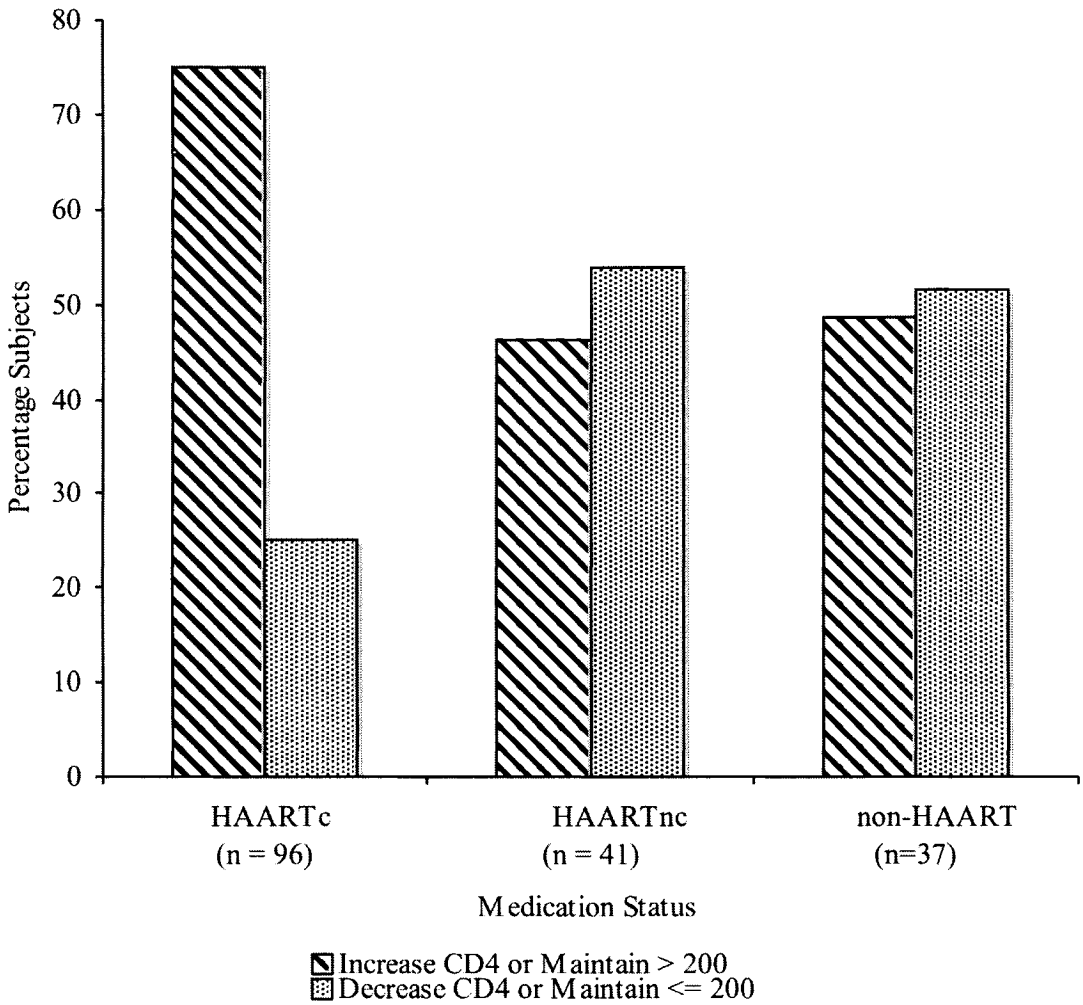


Figure 1. CD4 Cell Progression for Subjects According to Medication Status A.

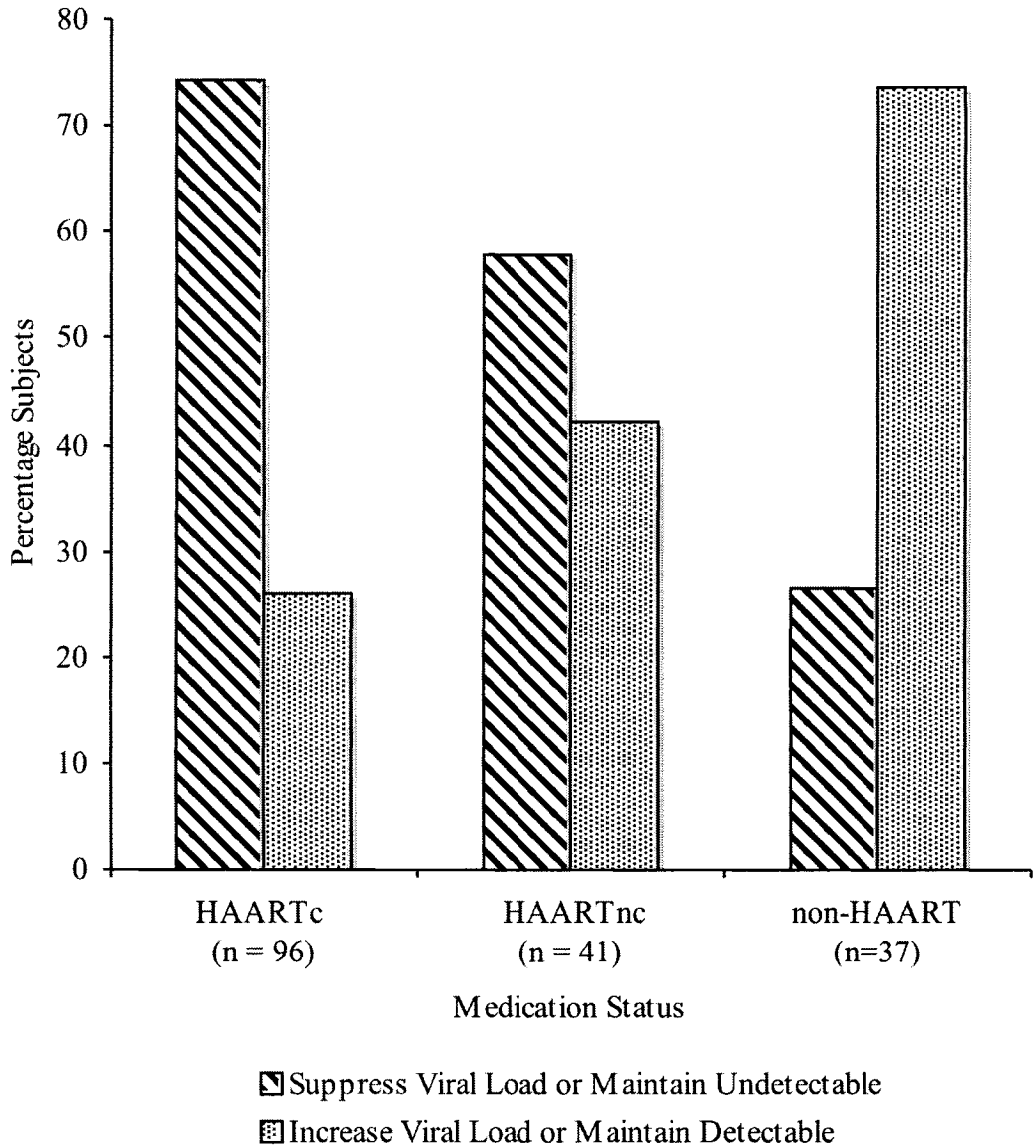


Figure 2. Viral Load Status Changes for Subjects According to Medication Status A.

These findings were maintained whether the subjects with neuromedical comorbidity were included or excluded, and also applied to the dichotomous medication split of being on HAART at any point in time versus never being on HAART (see Figures 3 and 4). All of these findings suggested that HAART was having the expected effect with respect to maintaining or reconstituting immune functioning and in terms of suppressing viral load. In particular, there seemed to be more benefit for subjects who were taking HAART for a longer duration compared to those who were not on HAART continuously.

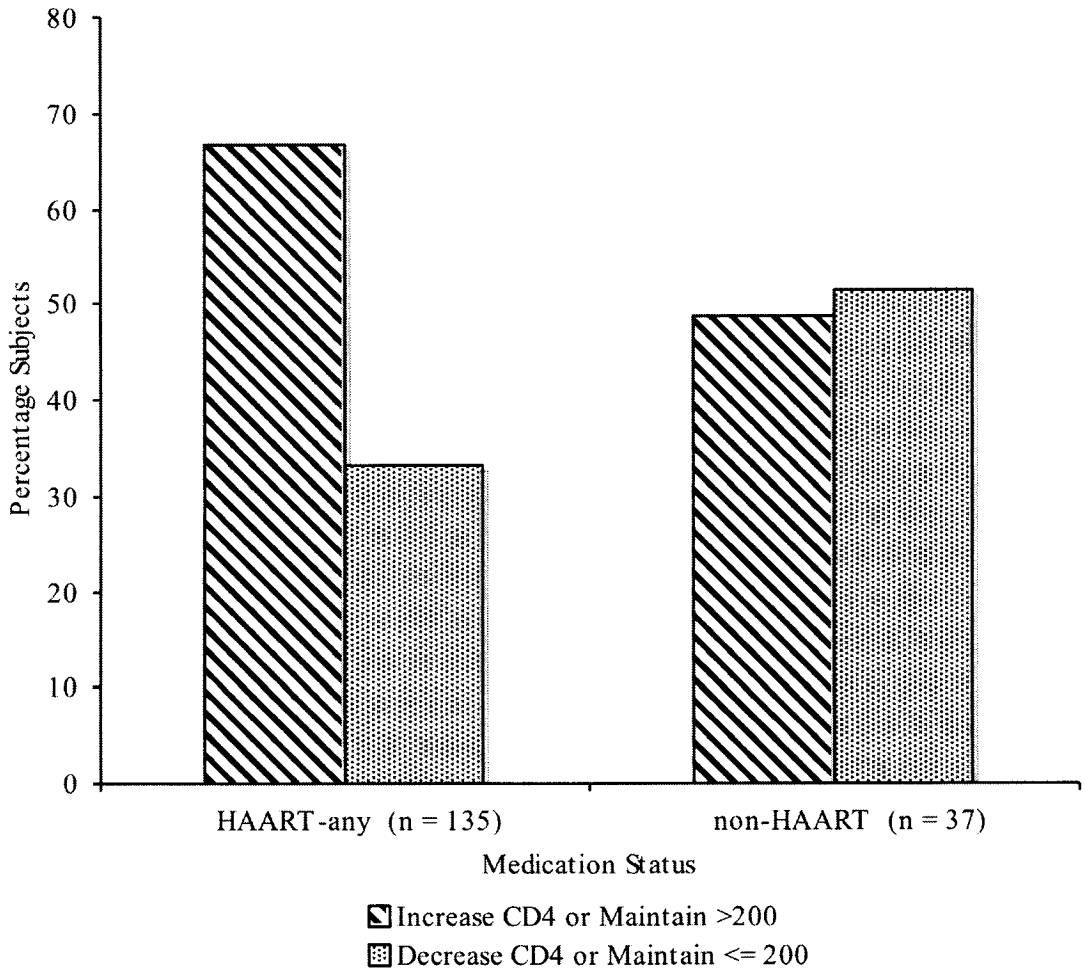


Figure 3. CD4 Cell Progression for Subjects According to Medication Status B.

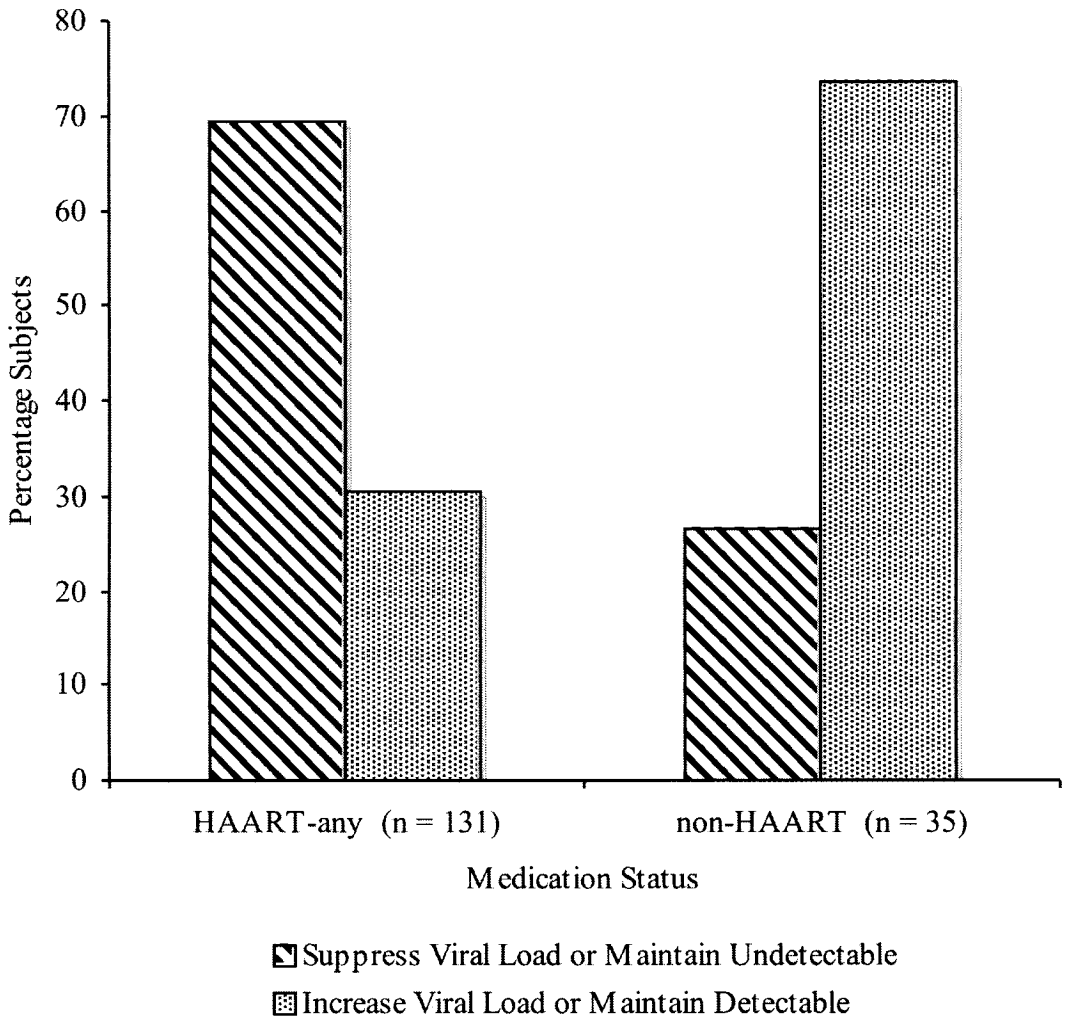


Figure 4. Viral Load Status Changes for Subjects According to Medication Status B.

At the time of this study's proposal, there was conflicting information in the literature about the potential relationship between change in cognitive function and markers of immune functioning, with many studies unable to elucidate a direct relationship. For the purposes of this study it was proposed that systemic markers of HIV disease (e.g., patient self-report of recent CD4 cell count and viral load) would not have a direct influence on changes in cognitive functioning over time for individuals on HAART.

As data collection proceeded, it became apparent that individuals in the three or four antiretroviral classifications seemed more heterogeneous with respect to initial immune status and ability to profit from therapy than first anticipated. As the data collection process for the current study evolved over the last four years, it became possible to acquire directly measured CD4 cell counts and viral load status from some subjects' lab-based blood work, rather than rely on patient report[3].

For these reasons, it was decided that a secondary set of Reliable Change Index analyses could be undertaken with subjects stratified by initial level of immune functioning and reconstitution progress (i.e., initial level and changes in CD4 cell count) in the hopes of better separating the heterogeneous HIV sample. Given the low frequency of subjects in some of the six categories described earlier (e.g., Low Decliners), the six groups were collapsed into two groups based on Low initial CD4 cell count or High initial CD4 cell count in an attempt to manage low cell size in subsequent analyses.

*Measures*

The neuropsychological assessment battery was selected on the basis of the National Institute of Mental Health workshop on Neuropsychological Assessment Approaches of AIDS-related Cognitive Changes and included the following: (1) Adult North American Reading Test (ANART; Spreen, & Strauss, 1991) (2) Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983); (3) California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987 List A Trial 1, List A Trial 5, Semantic

[3] It was assumed that the direct lab-based data for CD4 counts and viral load would be more accurate than the patients' self-report but this has been targeted for direct analysis in a future study. A decision was made to use the lab-based HIV markers where possible at Time 1 (CD4 n = 135 [67.5%]; and Time 2 (CD4 n = 132 [66%]; viral load n = 144 [72%]) and to use patient self-report of HIV markers otherwise.

Clustering, and Long-Delay Cued Recall); (4) Figure Memory Test (FMT; Heaton, Grant, & Matthews, 1991; Learning Efficiency and Percent Loss); (5) Grooved Pegboard Test (GPT; Heaton, Grant, & Matthews, 1991); (6) Stroop Neuropsychological Screening Test (Stroop; Trennerry, Crosson, DeBoe, & Leber, 1989 Colour-Word Score); (7) Symbol Digit Modalities Test (SDMT Smith, 1991); (8) Trail Making Test: Part A & B; (Reitan, & Wolfson, 1993); (9) Wechsler Adult Intelligence Scale – Revised (WAIS-R; Wechsler, 1981) Digit Span & Digit Symbol; and (10) Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtis, 1993; perseverative responses).

The tests were chosen to cover a broad range of psychological abilities, with an emphasis on the purported subcortical nature of HIV deficit. Most have been found to be sensitive to a wide range of mild impairments within an HIV-infected population (Heaton et al., 1995). The identified subtests from the CVLT were chosen based on the results of previous HIV subtyping research with this sample (Murji et al., 2003). The four chosen CVLT subtests (i.e., List A Trial 1, List A Trial 5, Semantic Clustering, and Long-Delay Cued Recall) had the highest loadings on factors of Attention Span, Learning Efficiency, and Delayed Recall. All of the tests were also categorized into one of seven domains (i.e., Language, Attention, Learning, Memory, Abstraction, Motor, and Psychomotor) based on known test characteristics (Lezak, 1995; Spreen, & Strauss, 1991) and following previous convention in HIV research (see Reger, Welsh, Razani, Martin, & Boone, 2002).

*Procedure*

The Research Ethics Board at St. Michael's Hospital, and the University of Windsor Research Ethics Board, approved the collection of data from study participants. All participants provided written, informed consent prior to the assessment and were provided opportunity to receive feedback about their results. Research subjects were paid $30 for their participation. All participants were given the opportunity to receive feedback on their performance and a summary report could be sent to their primary physician, infectious disease specialist, psychiatrist, and/or other individuals involved in their care upon request.

Neuropsychological testing, together with a brief interview and questionnaire completion, took approximately 3 to 4 hours and was conducted over two sessions,

separated by a 30 to 45 minute break. Under the supervision of a licensed clinical neuropsychologist, trained psychometrists or psychologists administered all tests. Most subjects were contacted approximately 12 to 18 months after the Time 1 assessment for repeat neuropsychological testing. The procedure for the follow-up assessment was identical to that of the first assessment.

*Statistical Analyses and Data Transformation*

Data analysis was conducted using SPSS for Windows (ver. 12.0) and for Mac (ver. 11.0). An alpha level of .05 was used for all statistical tests unless otherwise mentioned. Effect sizes were calculated using Cohen's *d* (Cohen, 1992). Data were screened for accuracy, missing points, and normality using techniques described by Tabachnick and Fidell (1996). Univariate outliers were defined as scores that exceeded an absolute, standardized *z*-score of 3.29 relative to the group mean (Tabachnick & Fidell, 1996).

As stated previously, a secondary goal of the current research was to compare group mean change methodology (Method A) to the Heaton Impairment Ratings method of assessing individual change (Method B). Reliable Change Index methodology (Method C) was used to determine longitudinal change because of its sound statistical properties and relevance to investigating disorders with significant individual variation.

*Group Change – Method A*

For Method A, raw scores on the neuropsychological test variables were converted to demographically corrected standardized scores, based on age, education, and sex using either Heaton norms (Heaton, Grant, & Matthews, 1991; Boston Naming Test, Digit Span, Digit Symbol, Verbal Fluency, Grooved Pegboard Test, Trail Making Test, Wisconsin Card Sorting Test) or test manuals (CVLT: Delis, Kramer, Kaplan, & Ober, 1987; SDMT: Smith, 1991; Stroop: Trennerry, Crosson, DeBoe, & Leber, 1989).

Scores across tests representative of traditional theoretical domains (i.e., language, attention, learning, memory, abstraction, psychomotor, motor) were averaged to address questions of cognitive domain performance. For domain-based analyses, *z*-scores on the CVLT were converted to *T*-scores with a mean of 50 and a standard deviation of 10 in order to make comparisons on a common metric with Heaton *T*-scores.

Domain-based analyses were run without Stroop data, which did not convert well to the *T*-score metric. These scores were subjected to t-tests and ANOVAs where appropriate.

*Individual Impairment Ratings – Method B*

In order to quantify the number and degree of impaired performances for each individual, standardized test scores (described above) were converted to impairment rating scores for each neuropsychological test per Heaton's automated classification approach (Heaton, Grant, & Matthews, 1991; Heaton, Kirson, Velin, Grant, & the HNRC Group, 1994; Heaton et al., 1995). When using this impairment rating approach, less weight is assigned to test scores that are within or above the normal range (see Table 4). Standardized scores for the CVLT and percentile scores for the Stroop tests were converted to impairment ratings based on the equivalent values of one, two, and three standard deviations below the mean. These impairment ratings were also averaged across the profile to yield an average impairment rating or "Global Deficit Score" (GDS) as described by Heaton et al. (1981). This approach has proven to be a useful adjunct or alternate to clinical ratings in summarizing HIV cognitive impairment across a large battery (Heaton et al., 1995). Recently, it has been recommended as a sensitive and specific way to summarize the neuropsychology test results for individuals with HIV (Carey et al., 2004). In the current study, the calculated Global Deficit Score correlated highly with blind clinical profile ratings from a qualified neuropsychologist ($r = 0.822$, $p < .001$), confirming the validity of using the automated classification approach with the current sample.

These classifications were employed to assist with categorization of individual performances at Time 1 and at Time 2 and to provide benchmarks for judging relative clinical change over time. Subjects' neuropsychological profiles were also assigned blind clinical ratings, on a scale of 1 (*Above Average*) to 9 (*Severe Impairment*), by an experienced neuropsychologist who was 'blind' to the participants' antiretroviral status (Heaton, Kirson, Velin, Grant, & the HNRC Group, 1994). When following this approach, a score of 4 indicates Borderline impairment and scores at 5 or above indicate Definite impairment. These ratings were made at Time 1 only in an effort to judge the

external validity of the Global Deficit Score calculated from the Heaton impairment ratings.

Table 4.
*Heaton Impairment Rating Systems*

| *T* Score | Impairment Rating | Clinical Classification | Clinical Rating |
|---|---|---|---|
| > 54 | | Above Average | 1 |
| 45 to 54 | | Average | 2 |
| 40 to 44 | | Below Average/ Borderline | 3 |
| >39 | 0 | Not Impaired | 4 |
| 35 to 39 | 1 | Mild | 5 |
| 30 to 34 | 2 | Mild to Mod. | 6 |
| 25 to 29 | 3 | Moderate | 7 |
| 20 to 24 | 4 | Mod. to Sev. | 8 |
| < 20 | 5 | Severe | 9 |

*T*-scores were interpreted in terms of Heaton impairment rating nomenclature (e.g., mild, moderate, severe impairment) and categorized as "normal" (i.e., $T > 39$) or impaired (i.e., $T \leq 40$). CVLT scores were assigned "mild", "moderate", and "severe" impairment labels when the standardized scores fell one, two, and three standard deviations below the mean, respectively. For the Stroop scores, mild impairment was defined as falling below the 16$^{th}$ percentile and moderate impairment was defined as falling below the 2$^{nd}$ percentile.

*Reliable Change Index Calculations - Method C*

Change Scores on Neuropsychological Tests

Reliable Change Index (RCI) methodology was used to assess the statistical and clinical significance of any changes in neuropsychological test scores from Time 1 to Time 2. Variations on reliable change computations have emerged since an early formula proposed by Jacobson & Truax (1991)[4], including a method that attempts to account for practice effects (RCI-PE; Chelune, Naugle, Lueders, Sedlak, & Awad, 1993). The reliable change methodology allows one to calculate a confidence interval around a retest score that is thought to include all possible scores that might have occurred by chance alone. Scores that fall outside of this predetermined range are judged to do so because of something beyond the effects of chance alone. This method of change analysis has been promoted in epilepsy surgery research (Chelune et al., 1993) and has been shown to reliably detect improving, worsening, and stable neurocognitive profiles in a study of neurological recovery and deterioration (Heaton et al., 2001). A brief description of the methodology used in the current study is presented next.

To obtain an RCI-PE value for each neuropsychological test administered, the difference between individuals' Time 1 and Time 2 test scores was compared against a calculated confidence interval, which took practice effects into consideration, beyond which only 10% of subjects would be expected to increase or decrease by chance alone. The practice effects were determined by comparing test-retest mean scores obtained from normal populations at test intervals that most closely matched the follow-up interval in the current study.

There is some debate regarding the specific formula used to calculate the standard error term for the RCI equation (Abramason, 2000; Chelune, 2003; Hinton-Bayre, 2000; Moritz, Iverson, & Woodward, 2003; Temkin, Heaton, Grant, & Dikmen, 2000; Temkin, 2004). For the purposes of the current research, the preferred RCI method was that suggested by Temkin et al. (1999; 2000) and endorsed by Abramason (2000) and Chelune (2003). RCI confidence intervals were defined as 1.64 times the standard error as defined: $\underline{SE} = \text{sqrt}\,[s_{12} + s_{22} - 2s_1s_2r_{12}]$ where $s_x$ is the standard deviation of a normative

[4] Jacobson & Truax formula: RC = sqrt [2* $(SEM_1)_2$], where SEM = standard error of measurement of scores at Time 1.

test scores at Time X; and $r_{12}$ is the test-retest reliability in a normative sample (Temkin, 2004). The resulting RCI value was the absolute increase or decrease in a score that would be expected by chance alone for 90% of a normal sample.

Test-retest reliability coefficients and normative group means were taken from test manuals or from published longitudinal research. Time 1 and Time 2 test scores were retained in their raw form, or in the standardized form compatible with the available reliable change equation building blocks (i.e., baseline means and standard deviations, follow-up means and standard deviations, and test-retest reliability coefficients). Every effort was made to use RCI equation building blocks generated from the same normative sample (Temkin, 2004) and this was possible for 11 of the 15 neuropsychological tests. When choosing RCI building blocks from among different normative studies, priority was given to published normative data in the manuals, unless alternate research more closely resembled the current data set in terms of subject characteristics (e.g., age, gender, HIV control status), baseline test scores, and test-retest interval. Research sample size was also considered in decision-making. When RCI building blocks were chosen from mixed sources, it was out of strategic choice or necessity. For example, means and standard deviations from large sample of HIV- controls chosen over a small sample of unspecified normal subjects tested 6 weeks apart, but test-reliability may have been unavailable from preferred source.

The following 15 neuropsychological test scores were used in the Reliable Change Index analyses: Boston Naming Test – total correct; California Verbal Learning Test (CVLT) [selected measures; List A Trial 1 words recalled (A1), List A Trial 5 words recalled (A5), Semantic Cluster Ratio (SEM), Long-Delay Cued Recall words recalled (LDCR)]; Controlled Oral Word Association (COWA) – total words; Grooved Pegboard Test – time per dominant hand and non-dominant hand; Stroop Neuropsychological Screening Test – colour-word score percentile; Symbol Digit Modalities Test – number correct; Trail Making Test A & B – completion time; Wechsler Adult Intelligence Test - Revised (WAIS-R) Digit Span - raw total forward and backward score; WAIS-R Digit Symbol Subtest – scaled score; Wisconsin Card Sorting Test – perseverative responses. Test-retest data was not located for the Figure Memory test so it was not included in the

RCI analyses. Sources for each of the RCI-PE equation building blocks are detailed in Table 5.

Table 5.
*Data Sources for Reliable Change Index Calculations*

| Test | Test-retest Means and Standard Deviations | Test-Retest Reliability Coefficients |
|---|---|---|
| Boston Naming Test | Saykin et al., (1991) | Sawrie et al., (1996) |
| COWA | Dikmen et al., (1999) | Dikmen et al., (1999) |
| WAIS-R Digit Span | Levine et al., (2004) | Levine et al., (2004) |
| WAIS R Digit Symbol | Saykin et al., (1991) | Wechsler (1981) |
| Stroop | Trennerry (1989) | Trennerry (1989) |
| CVLT LDCR | Delis et al. (1987) | Delis et al., (1987) |
| CVLT SEM | Delis et al. (1987) | Delis et al. (1987) |
| CVLT A1 | Saykin et al., (1991) | Duff et al., 2001 |
| CVLT A5 | Saykin et al., (1991) | Duff et al., 2001 |
| SDMT | Levine et al., (2004) | Levine et al., (2004) |
| Trail Making Test | Levine et al., (2004) | Levine et al., (2004) |
| WCST | Tate et al., (1998) | Tate et al., (1998) |
| Grooved Pegboard Test | Levine et al., (2004) | Levine et al., (2004) |

*Note*. COWA = Controlled Oral Word Association. WAIS-R = Wechsler Adult Intelligence Scale-Revised. CVLT = California Verbal Learning Test; LDCR = Long-Delay Cued Recall; SEM = Semantic Cluster Ratio; A1 = List A, Trial 1; A5 = List A, Trial 5; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test.

For neuropsychological test scores in which a higher score reflects better performance (i.e., BNT; COWA; Stroop; SDMT; CVLT A1, A5, SEM, LDCR; WAIS-R Digit Span and Digit Symbol), the pre-determined practice effect values were added to the positive RCI values to create a statistically significant improvement standard. Individual scores that increased beyond this standard from Time 1 to Time 2 were considered *Better*. To determine whether subjects' performance had declined significantly over time, the practice effect[5] was added to the negative RCI value and this was compared against the subject's test-retest change. If the subject's score decreased from Time 1 to Time 2 by more than this standard, it was considered W*orse*. In contrast, for neuropsychological tests in which a lower score signals a better performance (i.e., Trail Making Test, Grooved Pegboard Test, WCST perseverative responses), the practice effect values (which were negative numbers) were added to the negative RCI value to determine the standard for statistical improvement. Subject scores that decreased by more than this standard were labelled *Better*. To determine whether or not subjects' performance on these same variables had declined significantly over time, the practice effect was added to the positive RCI value. Subject change from Time 1 to Time 2 that was greater than this standard was characterized as *Worse*. Subject change scores that fell in between the two change standards were described as *No Change*.

Reliable Change Index Scores Across Domain and Profile

To determine whether a subject's neuropsychological status had changed in terms of domain specific functioning, the following data reduction method was used. Domain improvement was considered to have taken place if more than 50% of the test scores within than domain had fallen in the *Better* category, or if 50% of the tests were found to be better and the other tests were unchanged. Similarly, if more than 50% of the test scores in a given domain had declined significantly, or if 50% of the test scores had declined and the remaining tests were unchanged, the domain was considered to be *Worse*.

[5] For two tests in this category (CVLT A1 and WAIS-R Digit Symbol), the normative data indicated a negative practice effect.

To determine whether or not a subject's overall neuropsychological profile status had changed, subjects were classified based on the number of test scores that had changed. That is, an *overall profile change* score was calculated based on the number of tests that were judged to be significantly better at Time 2 (by RCI-PE method) minus the number of tests scores that were judged to be significantly worse at Time 2 (by RCI-PE method) (Heaton et al., 2001). Heaton and colleagues designated an individual's profile as changed if at least 2 out of 7 neuropsychological test scores had changed reliably, based on their finding that approximately 90% of normal subjects obtained only 1 changed test score. In other words, a subject's profile was considered changed if the change score fell between -1 and + 1. For the current research, a profile was considered unchanged in its entirety if the difference between positive and negative test scores fell between - 2 and +2. This more conservative criterion was chosen to reflect the larger number of tests in the current battery while minimizing the impact of Type II error. A profile was considered Worse if the difference score was less than -2 and considered Better if the score was greater than +2. In addition, the numerical change score (e.g., +2, -3) was used as a continuous variable in subsequent analyses.

*Statistical Analyses*

In order to address the first set of hypotheses (i.e., that subjects on HAART would show better cognitive performance compared to non-HAART at baseline, at follow-up, and over time), comparisons, by means of non-parametric tests (e.g., chi square) were made between the percentage of subjects on HAART +/+ who exhibited "Better" overall and domain ratings compared to the percentage of subjects in the HAART -/- group.

Of subjects who are classified as Better at Time 2, regardless of HAART status, it was also predicted that more subjects would show improvements in the psychomotor domain than would show improvement on other domain scores. To address this question, the percentage of subjects who showed improvement on the psychomotor domain score were compared to the percentage of Better subjects who showed improvement in other domains. Of the subjects who were classified as Better at Time 2 in the Attention & Learning domains, the number of subjects who showed improvements in the psychomotor domain was compared with improvement in other domain scores.

To address the second hypothesis about HAART's dose-response relationship,

improvement in neuropsychological functioning, as classified by Better overall RCI profile, was predicted by duration of HAART, as reflected in continuous use of HAART versus non-continuous HAART. The Pearson product-moment correlation between CNS drug penetration and RCI profile outcome was calculated.

Finally, it was predicted that participants with a history of neuromedical risk factors and those with lower educational achievement or lower estimated IQ scores would show no or minimal improvement on HAART. It was also predicted that there would be no relationship between neuropsychological improvement over time and changes in systemic immune functioning (e.g., as reflected by CD4 count and plasma viral load).

To assess whether changes in overall neuropsychological functioning were differentially affected by the aforementioned demographic and medical variables, a binary logistic regression analysis was performed. The goal was to assess prediction of RCI overall profile outcome (i.e., Worse or Better) on the basis of neuromedical comorbidity (present or absent), education and estimated IQ levels, and HIV markers (i.e., recent CD4 cell count, plasma viral load, and CDC 93 Disease stage – Asymptomatic, Symptomatic or AIDS).

# Results

*Data Screening*

Initial screening of the data indicated no significant skewness or kurtosis in any of the demographic or HIV marker variables, with the exception of viral load, which was transformed to a dichotomous variable (i.e., detectable versus undetectable) for subsequent analyses.

There were no significant differences at Time 1 between subjects taking HAART and subjects not taking HAART with respect to years of formal schooling, estimated IQ level (ANART), or recent CD4 cell count but HAART subjects were slightly older ($U$ = 13795, $p$ = .01) (see Table 6). There were more Asymptomatic subjects in the non-HAART group (18.4%) compared to the HAART group (5.8%), $\chi^2$ (2, $n$ = 328) = 15.6, which probably reflects the fact that Asymptomatic subjects usually do not start ARV therapy until their symptoms worsen or their immune system weakens (Yeni et al., 2002). In addition, among non-HAART subjects, more subjects than expected by chance (82.6%) had detectable viral loads and fewer were undetectable (17%), $\chi^2$ (1, $n$ = 351) = 72.5. This finding reflects the fact that 70% of the non-HAART group were not taking antiretrovirals and, therefore, would be expected to have detectable amounts of virus in their blood. There were no significant differences between groups in terms of neuromedical comorbidity.

Table 6.
*Demographic and HIV Marker Variables for Subjects on and off HAART at Time 1*

| | HAART | | non-HAART | |
|---|---|---|---|---|
| | *M (SD)* | *n* | *M(SD)* | *n* |
| Age (years)* | 40.2 (9.1) | 257 | 42.4 (7.9) | 127 |
| Education (years) | 14.0 (2.8) | 257 | 13.8 (2.9) | 127 |
| ANART IQ | 114.8 (7.9) | 242 | 115.5 (6.9) | 117 |
| Recent CD4 | 373.6 (262.2) | 255 | 391.1 (250.0) | 125 |
| CDC Stage** | Asymp. 5.8%<br>Symp. 39.1%<br>AIDS 55.1% | 255 | Asymp. 18.4%<br>Symp. 42.7%<br>AIDS 38.8% | 103 |
| Viral Load Undetectable** | 82.6% | 242 | 17.4% | 109 |
| Neuro-Medical Comorbidity | Present -32.3% | 220 | Present - 34.3% | 105 |

*Note.* ANART = National Adult Reading Test estimated IQ; Recent CD4 = subjects' most recent CD4 cell count in cells/μl; Viral Load Undetectable = plasma viral count below 50 or 500 copies/ml, depending on year of test.

* $p < .05$

** $p < .001$

There were also no significant differences between longitudinal antiretroviral groups' classifications in terms of baseline age, estimated IQ, education, or presence of comorbidity (see Table 7). Among those subjects who were never on HAART, there were more Asymptomatic subjects (22.2%), and fewer subjects with AIDS (30.6%), compared to those subjects who were on HAART, either continuously or non-continuously, [$\chi^2$ (4, $n$ = 164) = 11.4; $p$ = .02]. This finding most likely reflects the fact that Asymptomatic subjects usually do not start ARV therapy until their symptoms worsen or their immune system weakens to a recommended level for initiating therapy (e.g., < 200 cells/μl) (Yeni et al., 2002). Among the subjects who took HAART throughout the study, more subjects than expected (62.8%) had undetectable viral loads, whereas fewer than expected subjects in the non-HAART group (17.6%) had undetectable viral status, $\chi^2$ (2, $n$ = 167) = 28.2; $p$ < .001. Once again, this finding presumably reflects successful viral suppression among subjects taking HAART compared to those not taking HAART.

Table 7.
*Results of One-way ANOVAs on Demographic and HIV Variables*

| | **HAARTc** | | **HAARTnc** | | **non-HAART** | |
|---|---|---|---|---|---|---|
| | *M(SD)* | *n* | *M(SD)* | *n* | *M(SD)* | *n* |
| Age (years) | 42.9 (7.2) | 100 | 43.2 (9.0) | 40 | 39.9 (8.7) | 39 |
| Education (years) | 14.6 (2.6) | 100 | 14.4 (2.4) | 40 | 13.9 (3.0) | 39 |
| ANART IQ | 116.7 (7.3) | 100 | 116.0 (8.1) | 39 | 116.2 (5.6) | 36 |
| Recent CD4 | 377.5 (249.0) | 100 | 367.8 (296.3) | 39 | 459.0 (268.4) | 38 |
| CDC Stage[a] | Asymp. 6.4%<br>Symp. 38.3%<br>AIDS 55.3% | 94 | Asymp. 5.9%<br>Symp. 38.2%<br>AIDS 55.9% | 34 | Asymp. 22.2%<br>Symp. 47.2%<br>AIDS 30.6% | 36 |
| Viral Load Undetectable[a] | 62.8% | 94 | 25.6% | 39 | 17.6% | 34 |
| Neuro-Medical Comorbidity | Present -23.2% | 95 | Present - 30.6% | 36 | Present - 26.3% | 38 |

*Note.* HAARTc = continuous HAART; HAARTnc = non-continuous HAART; non-HAART = no HAART at T1 or T2.
[a] significant $p < .05$

Prior to analysis, neuropsychology data were screened separately within each of the medication status groups at Time 1 and Time 2. Outlier analysis on the neuropsychological variables at Time 1 identified six outliers within the HAART group and two outliers in the non-HAART group. At Time 2 there were four outliers identified within the HAARTc group. The contributing subjects were judged to be representative of the target population so these 10 test scores were included in subsequent analyses but adjusted to one unit higher or lower than the next closest test score.

Formal tests of normality and visual inspection of histograms and normal probability plots suggested that the following variables had potential problems with skewness and/or kurtosis: Time 1 CVLT A5, Figure Memory Test – Memory score; and Time 2 CVLT LDCR, CVLT SEM, CVLT A1, Digit Span, Figure Memory Test – Memory score, Stroop, SDMT, and WCST perseverative responses. All of the neuropsychological tests score distributions satisfied conditions for homogeneity of variance (Field, 2005; Tabachnick, & Fidell, 1996). Because ANOVA and t-tests are generally robust to violations of normality (Tabachnick, & Fidell, 1996), and it was felt important to retain the metric of these scores, the variables were not subjected to transformations.

Because the purpose of the Reliable Change Index (RCI) methodology was to identify significant change between individual subjects' scores at Time 1 and Time 2, outlier analysis was deemed inappropriate for those analyses. To illustrate, one subject showed extremely high scores on the SDMT at both Time 1 and at Time 2. Each of these scores qualified as outliers when compared to other subjects' scores and risked elimination from the sample. However, these high scores actually reflected consistent, if superior, performance on the SDMT test for that individual when viewed longitudinally. Such scores were retained in RCI analyses for that reason.

Test-retest intervals revealed less than 6 months (n = 21) or greater than 4 years (n = 6) were identified as possible outliers. However, no significant differences occurred between RCI analyses conducted with samples including and excluding these outliers. Thus, those scores were retained for further analyses.

*Reliable Change Calculations*

Table 8 shows the statistically significant improvement and worsening standards as calculated for each neuropsychological test.

Table 8.
*RCI,Practice Effect, and RCI Standards for Neuropsychological Variables*

| Neuropsychological test variable | \|RCI\| value | Practice Effect[a] | Better if (T2 - T1) change is more than | Worse if (T2 - T1) change is less than |
|---|---|---|---|---|
| Boston Naming Test (raw correct) | 1.11 | 0.6 | +2 | -1 |
| COWA (total words raw score) | 12.99 | 3.1 | +14 | -12 |
| WAIS-R Digit Span (raw total) | 4.21 | 0.3 | +4 | -4 |
| WAIS R Digit Symbol (scaled score) | 1.58 | -0.1 | +2 | -2 |
| Stroop (colour/word raw score) | 13.08 | 5.0 | +18 | -8 |
| CVLT LDCR (total words) | 2.84 | 1.0 | +4 | -2 |
| CVLT SEM (raw score) | 1.25 | 0.6 | +1.8 | -0.7 |
| CVLT A1 (raw score) | 3.26 | -0.3 | +3 | -4 |
| CVLT A5 (raw score) | 2.12 | 0.2 | +2 | -2 |
| SDMT (raw score) | 9.50 | 2.3 | +12 | -7 |
| Trail Making Test Part A (seconds) | 11.93 | -3.0 | -14 | +8 |
| Trail Making Test Part B (seconds) | 24.37 | -4.9 | -29 | +19 |
| WCST (perseverative responses) | 14.95 | -3.6 | -19 | +11 |
| Grooved Pegboard Test DH (seconds) | 11.45 | 2.5 | -14 | +9 |
| Grooved Pegboard Test NDH (seconds) | 12.08 | -2.6 | -15 | +9 |

*Note.* COWA = Controlled Oral Word Association; WAIS-R = Wechsler Adult Intelligence Scale-Revised; CVLT = California Verbal Learning Test; LDCR = Long-Delay Cued Recall; SEM = Semantic Cluster Ratio; A1 = List A, Trial 1; A5 = List A, Trial 5; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test. DH = dominant hand; NDH = non-dominant hand.

[a] a constant calculated from the mean test score change in the normative sample

*Hypothesis 1 - Group Mean Methodology Time 1*

The group means and standard deviations for Time 1 neuropsychological test variables are presented in Table 9[6]. HAART subjects, as a group, had a better performance than the non-HAART group on the Grooved Pegboard Test (dominant hand) ($p$ = .03; $d$ = .42), as expected, but also on Trail Making Test - Part B ($p$ = .05; $d$ = .39), and WCST perseverative responses ($p$ < .001; $d$ = .66), which were not predicted. In addition, there was a non-significant difference, in the expected direction, on Trails A ($p$ = .06; $d$ = .36). Consistent with the hypothesis, group mean scores on the three tests of attention (i.e., CVLT A1, Digit Span, Stroop) and two tests of learning (i.e., CVLT A5, Figure Memory – Learning) all showed (non-significant) differences in the predicted direction, with mean effect sizes of $d$ = 0.10 and $d$ = 0.21, respectively. The results did not differ substantially when subjects with neuromedical comorbidities were included or excluded from the sample.

[6] There were no significant differences between HAART subjects and non-HAART subjects on Time 1 neuropsychological tests scores when all single assessment subjects were included in the Time 1 analyses (see Appendix A). This suggests that there might be some differences between the group of subjects who had follow-up testing compared to those who had a single assessment beyond the demographic variables, HIV markers and medication status previously mentioned.

Table 9.
*Group Means and Standard Deviations for Time 1 Neuropsychological Test Variables for Non-HAART and HAART groups*

| | non-HAART | | | HAART | | | Statistics |
|---|---|---|---|---|---|---|---|
| **Test Variable** | *n* | Mean | *SD* | *n* | Mean | *SD* | *t/z* value |
| BNT[a] | 42 | 49.6 | 14.8 | 85 | 47.8 | 12.9 | ns |
| COWA[a] | 41 | 46.6 | 8.9 | 85 | 48.0 | 10.9 | ns |
| Digit Span[a] | 42 | 48.8 | 11.1 | 85 | 50.3 | 10.6 | ns |
| Digit Symbol[b] | 40 | 46.6 | 9.7 | 85 | 48.9 | 10.3 | ns |
| Stroop[c] | 29 | 53.6 | 37.0 | 70 | 54.4 | 38.0 | ns |
| CVLT LDCR[b] | 42 | -1.0 | 1.5 | 85 | -0.6 | 1.3 | ns |
| CVLT SEM[b] | 42 | -0.3 | 1.3 | 85 | -0.2 | 1.1 | ns |
| CVLT A1[b] | 42 | -0.7 | 1.0 | 85 | -0.6 | 1.0 | ns |
| CVLT A5[b] | 42 | -0.9 | 1.7 | 85 | -0.4 | 1.5 | ns |
| SDMT[a] | 36 | 44.2 | 11.6 | 85 | 44.4 | 11.3 | ns |
| Trail Part A[a] | 41 | 47.3 | 10.4 | 85 | 51.4 | 12.3 | ns |
| Trail Part B[a] | 41 | 47.4 | 10.5 | 85 | 51.8 | 11.9 | -2.0* |
| WCST[a] | 40 | 43.4 | 11.6 | 83 | 50.1 | 8.6 | -3.6** |
| GPT DH[a] | 42 | 42.5 | 11.8 | 85 | 47.6 | 12.5 | -2.2* |
| GPT NDH[a] | 42 | 41.6 | 11.3 | 85 | 45.2 | 10.9 | ns |
| FM – Memory | 36 | 53.7 | 5.2 | 80 | 52.8 | 5.6 | ns |
| FM – Learning | 36 | 40.9 | 7.3 | 80 | 41.9 | 9.4 | ns |

*Note.* BNT = Boston Naming Test; COWA = Controlled Oral Word Association; Digit Span = WAIS-R Digit Span; Digit Symbol = WAIS-R Digit Symbol; CVLT = California Verbal Learning Test; LDCR = Long-Delay Cued Recall; SEM = Semantic Cluster Ratio; A1 = List A, Trial 1; A5 = List A, Trial 5; GPT = Grooved Pegboard Test; FM = Figure Memory; SDMT = Symbol Digit Modalities Test; Trail = Trail Making Test. WCST[a] = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand; FM = Figure Memory Test.
[a] T score. [b] standard score. [c] percentile
* $p<.05$
** $p<.01$

All of the mean neuropsychological test *T*-scores for both medication groups fell at the borderline impaired range or higher according to Heaton's impairment rating scale (i.e., $T > 39$). These numbers conform to the typical expectation that neurocognitive

impairment for HIV groups as a whole is mild. However, closer inspection of the data revealed that the range of individual scores subsumed by the borderline level group means showed more variation, consistent with expectation for an HIV population. In other words, the borderline range group mean belies the level of impairment shown by many individual subjects within the group.

In fact, of the all of individuals in the non-HAART group, 43% ($n$ = 18) had at least one moderately impaired rating on a neuropsychological test (i.e., $T \leq 29$; percentile $< 2$; $z$-score $\leq$ - 2.0) (see Figure 5). In addition, all but two subjects ($n$ = 40; 95%) showed a mildly impaired rating (i.e., $T \leq 39$; percentile $< 16$; $z$ score $\leq$ -1.0) on at least one neuropsychological test. Within the HAART group, 35% ($n$ = 30) had at least one moderately impaired test score and 82% ($n$= 69) had at least one mildly impaired score. This finding is consistent with expectations about the methodological discrepancies in assessing impairment by way of group means (Method A) versus individual impairment (Method B).

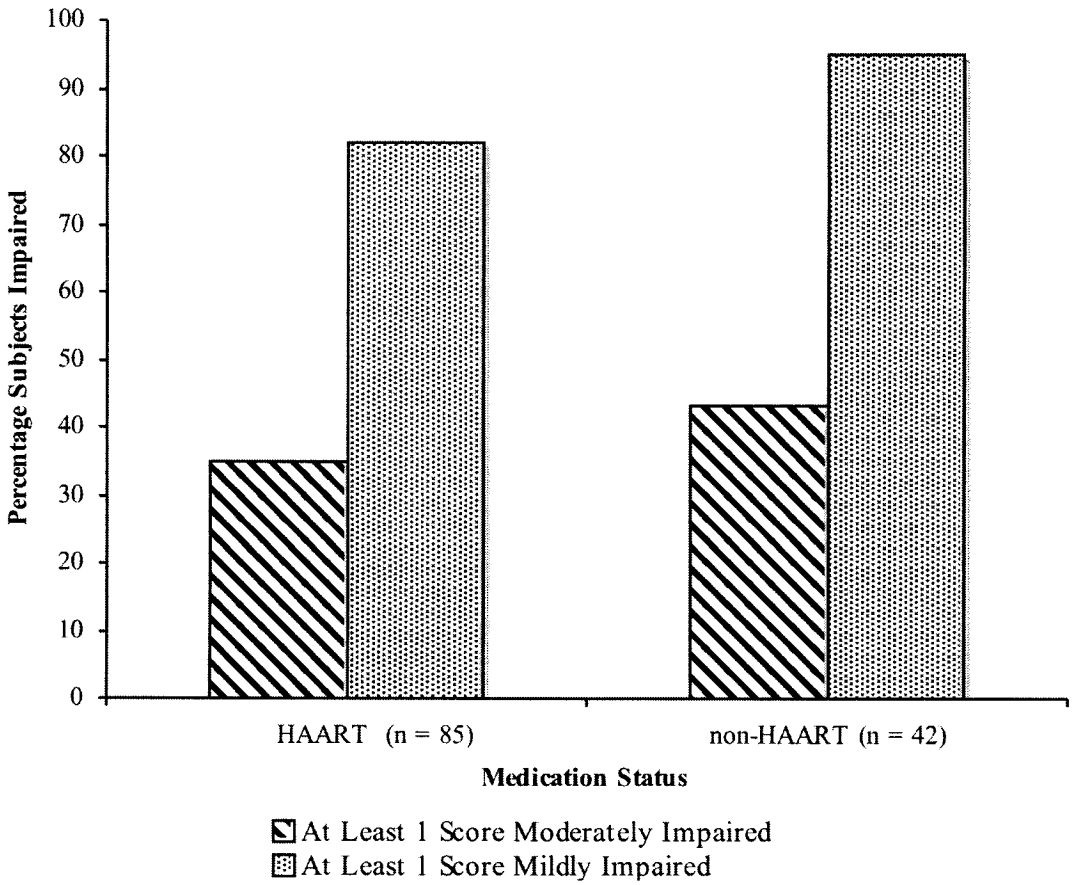


Figure 5. Percentage of Impaired Subjects by Medication Group at Time 1.

To further investigate the pattern of impaired test scores, the frequency with which HAART and non-HAART subjects showed at least mild impairment for each neuropsychological test is shown in Figure 6. For this purpose, the mildly impaired test scores were collected only from subjects who had a Global Deficit Score in the impaired range (i.e., > 0.50) in order to reflect a pattern of individual test impairment among subjects with definitive impairment. The three most frequently impaired test scores among non-HAART subjects were the fifth trial (A5; 83.3%), the long-delay cued recall trial (LDCR; 80.6%), and the first trial (A1; 77.8%) from the CVLT. The three most frequently impaired test scores among HAART subjects were the long-delay cued recall trial (LDCR; 85.8%), the first trial (A1; 84.3%), and the fifth trial (A5; 80.0%) from the CVLT. For both groups, all but one of the five the most frequently impaired tests were representative of the Attention, Learning Efficiency, and Psychomotor domains, and the other subtest reflected (verbal) Memory. These findings are largely consistent with the pattern of isolated HIV-related cognitive deficits documented for HIV subjects. Again, these impairment ratings (Method B) illustrate a varied, subtle, and highly prevalent, HIV-related impairment profile in contrast with the group mean statistics (Method A).

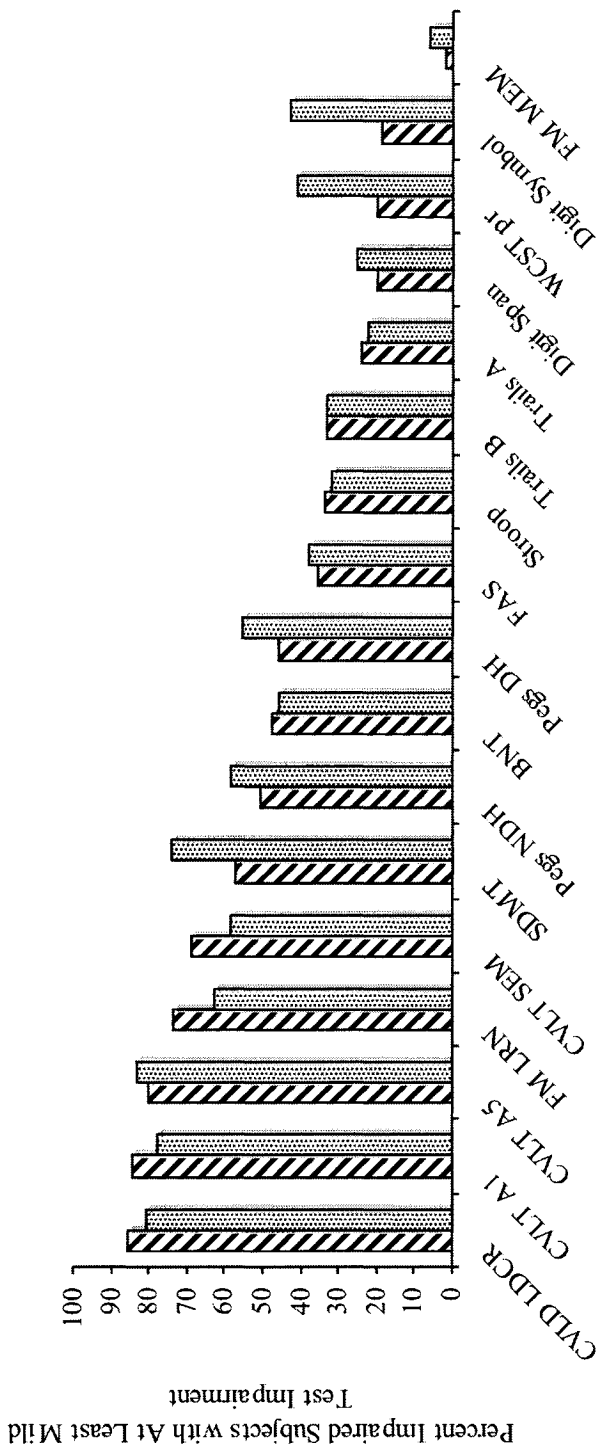


Figure 6. Percentage of Subjects with At Least Mild Impairment by Test at Time 1

When Time 1 neuropsychological test scores were combined into domain aggregates, it was revealed that subjects taking HAART had significantly better group mean scores than subjects not taking HAART in the Motor domain ($p$ = .04; $d$ = .39; see Table 10). A non-significant trend in the expected direction was also evident in the Psychomotor domain ($p$ = .20; $d$ = .37). Somewhat unexpectedly, subjects taking HAART also had significantly better group mean scores than subjects not on HAART in the area of Abstraction ($p$ = .001; $d$ = .65).

Table 10.
*Time 1 Means and Standard Deviations for Non-HAART and HAART Groups by Neuropsychological Domain*

| | non-HAART | | | HAART | | | Significance |
|---|---|---|---|---|---|---|---|
| **Domain** | *n* | Mean | *SD* | *n* | Mean | *SD* | *t/z* |
| Language | 41 | 48.3 | 9.0 | 85 | 47.9 | 9.4 | ns |
| Attention | 42 | 45.9 | 7.1 | 85 | 47.3 | 7.5 | ns |
| Learning[a] | 36 | 46.5 | 9.5 | 80 | 49.1 | 8.4 | ns |
| Memory | 36 | 46.8 | 9.2 | 80 | 48.4 | 7.4 | ns |
| Abstraction | 39 | 45.4 | 8.8 | 83 | 51.0 | 8.4 | -3.4** |
| Motor | 42 | 42.1 | 10.9 | 85 | 46.4 | 11.0 | -2.1* |
| Psychomotor | 42 | 46.0 | 8.7 | 85 | 48.2 | 9.8 | ns |
| Clinical Rating | 42 | 4.7 | 1.6 | 83 | 3.9 | 1.6 | 2.5* |
| Heaton GDS | 42 | 0.72 | 0.59 | 85 | 0.52 | 0.47 | 2.1* |

*Note.* domain scores are based on average of standardized T scores.
[a] Mann-Whitney U
[b] Means are presented in original T score metric, rather than transformed scores, for ease of interpretation
* $p < .05$
** $p < .01$

The two summary measures of overall neuropsychological performance at Time 1 also indicated a relative superiority of the HAART group mean scores compared to those of the non-HAART group. With respect to clinical ratings, the HAART group earned a mean clinical rating of 3.9 ($SD$ = 1.6), which was significantly lower than that of the non-

HAART group (4.7; *SD* = 1.6; *p* = .01; *d* = .47) but both scores fell roughly within the borderline range of impairment. The mean Global Deficit Scores for both groups met the cut-point for impairment recommended by Carey et al., (2004) but the Global Deficit Score for the non-HAART group was also significantly above that for the HAART group (*p* = .04; *d* = .39). It should be noted that 42% of the HAART subjects fell in the impaired range according to the recommended Global Deficit Score cut-point and 50% of the non-HAART subjects met the same criteria.

*Hypothesis 1 - Group Mean Methodology Time 2*

Subjects who had been taking HAART continuously (HAARTc) had better Time 2 scores than subjects who were not on HAART on all neuropsychological test measures but two (see Table 11). Only the difference between HAARTnc and non-HAART subjects on the grooved pegboard test (dominant) was statistically significant (*F* (2, 121) = 3.52; p = .03; *d* = .72). Though non-significant, the next highest effect sizes were found on Trail Making Test A (*d* = .54), Stroop (*d* = .46), and SDMT (*d* = .40). Further exploration of the hypothesis-related tests revealed that group mean scores on the three tests of attention (i.e., CVLT A1, Digit Span, Stroop) and one test of learning (i.e., CVLT A5) all showed (non-significant) differences in the predicted direction, with mean effect sizes of *d* = 0.40 and *d* = 0.36, respectively.

Table 11.

*Time 2 Means and Standard Deviations of Neuropsychological test T scores for non-HAART, HAARTnc & HAARTc groups.*

| | non-HAART | | | HAARTnc | | | HAARTc | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | *n* | Mean | *SD* | *n* | Mean | *SD* | | Mean | *SD* | ***Post-hoc*** |
| BNT | 26 | 48.8 | 13.4 | 24 | 52.8 | 14.6 | 71 | 49.3 | 13.0 | |
| COWA[a] | 26 | 46.6 | 10.9 | 25 | 50.4 | 9.3 | 72 | 48.3 | 10.3 | |
| Digit Span[a] | 26 | 47.9 | 10.7 | 24 | 48.0 | 10.9 | 72 | 51.1 | 12.0 | |
| Digit Symbol[b] | 26 | 47.7 | 10.3 | 24 | 51.2 | 9.9 | 72 | 51.8 | 10.7 | |
| Stroop[ce] | 23 | 47.4 | 38.6 | 24 | 62.9 | 35.6 | 66 | 65.1 | 37.5 | |
| CVLT LDCR[b] | 25 | -1.0 | 2.0 | 24 | -0.1 | 1.5 | 71 | -0.4 | 1.3 | |
| CVLT SEM[b] | 25 | -0.7 | 1.0 | 24 | -0.4 | 1.2 | 71 | -0.4 | 1.2 | |
| CVLT A1[b] | 25 | -1.5 | 1.0 | 24 | -1.0 | 1.2 | 71 | -1.0 | 1.1 | |
| CVLT A5[b] | 25 | -1.2 | 2.0 | 24 | -0.7 | 1.8 | 71 | -0.7 | 1.7 | |
| SDMT[ad] | 26 | 41.6 | 12.8 | 24 | 46.1 | 7.9 | 72 | 46.6 | 12.4 | |
| Trails A[a] | 26 | 47.6 | 11.8 | 24 | 53.8 | 12.7 | 72 | 54.4 | 13.3 | |
| Trails B[a] | 26 | 49.6 | 12.0 | 24 | 50.5 | 10.3 | 72 | 53.9 | 11.9 | |
| WCST[a] | 24 | 46.7 | 10.3 | 24 | 46.9 | 9.4 | 71 | 50.2 | 7.6 | |
| GPT DH[a] | 26 | 42.6 | 14.7 | 24 | 51.4 | 9.2 | 72 | 47.7 | 11.4 | Hnc > non-H* |
| GPT NDH[a] | 26 | 43.0 | 12.3 | 24 | 48.0 | 7.6 | 72 | 45.8 | 10.5 | |
| FM MEM | 18 | 54.7 | 4.5 | 11 | 55.1 | 4.5 | 43 | 53.5 | 4.9 | |
| FM LRN | 18 | 41.8 | 8.7 | 11 | 46.2 | 9.9 | 43 | 40.7 | 7.6 | |

*Note*: BNT = Boston Naming Test; COWA = Controlled Oral Word Association; CVLT = California Verbal Learning Test; LDCR = Long-Delay Cued Recall; SEM = Semantic Cluster Ratio; A1 = List A, Trial 1; A5 = List A, Trial 5; GPT = Grooved Pegboard Test; FM = Figure Memory SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand; FM MEM = Figure Memory Test Memory score; FM LRN = Figure Memory Test Learning score.

[a] T score; [b] standard score; [c] percentile; * $p < .05$

There were no statistically significant differences between subjects in terms of domains, however trends in the expected direction were evident for all domains (see Table 12). The mean domain aggregate scores for the group of subjects who never took HAART all fell in the Below Average or at the low end of the Average range, according to Heaton's criteria. In contrast, all of the mean domain aggregate scores for subjects who were continuously on HAART fell more consistently within the Average range.

Table 12.
*Time 2 Mean T Scores and Standard Deviations for Antiretroviral Groups by Neuropsychological Domain.*

| | non-HAART | | | HAARTnc | | | HAARTc | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Domain** | *n* | Mean | *SD* | *N* | Mean | *SD* | *n* | Mean | *SD* | ***Post-hoc*** |
| Language | 26 | 47.7 | 9.3 | 24 | 51.9 | 10.1 | 71 | 48.7 | 9.7 | |
| Attention | 25 | 41.7 | 7.2 | 24 | 43.8 | 8.8 | 71 | 45.4 | 8.9 | |
| Learning[a] | 17 | 43.9 | 8.6 | 11 | 48.7 | 10.8 | 42 | 47.9 | 7.2 | |
| Memory | 17 | 46.7 | 10.7 | 11 | 52.1 | 8.9 | 42 | 50.1 | 6.5 | |
| Abstraction | 24 | 48.6 | 9.4 | 24 | 48.7 | 7.0 | 71 | 52.1 | 8.4 | |
| Motor | 26 | 42.8 | 12.9 | 24 | 49.7 | 7.4 | 72 | 46.8 | 10.1 | |
| Psychomotor | 26 | 45.6 | 10.7 | 24 | 50.4 | 8.5 | 72 | 51.0 | 10.4 | |
| GDS | 26 | 0.86 | 0.72 | 25 | 0.48 | 0.40 | 72 | 0.54 | 0.52 | Hc < non-H * |

*Note.* Domain scores are based on averages of standardized T scores.
[a] Stroop not included
* $p < .05$

Once again, a trend can be seen in terms of group mean scores masking individual impairment for these domains. Of the subjects who were never on HAART, all but two subjects (93%; $n = 25$) showed mild impairment in at least one test and over 78% ($n = 21$) had at least one moderately impaired test score (see Figure 7). Despite falling in the Average range as a group, over half of the subjects who were HAART continuously (58%; $n = 42$) had at least one moderately impaired test score, and all but two subjects (97%; $n = 70$) showed mild impairment in at least one test. Within the group of subjects who were on HAART non-continuously, approximately 76% ($n = 19$) had at least one moderately impaired test score and almost 92% ($n = 23$) showed mild impairment on at least one test.

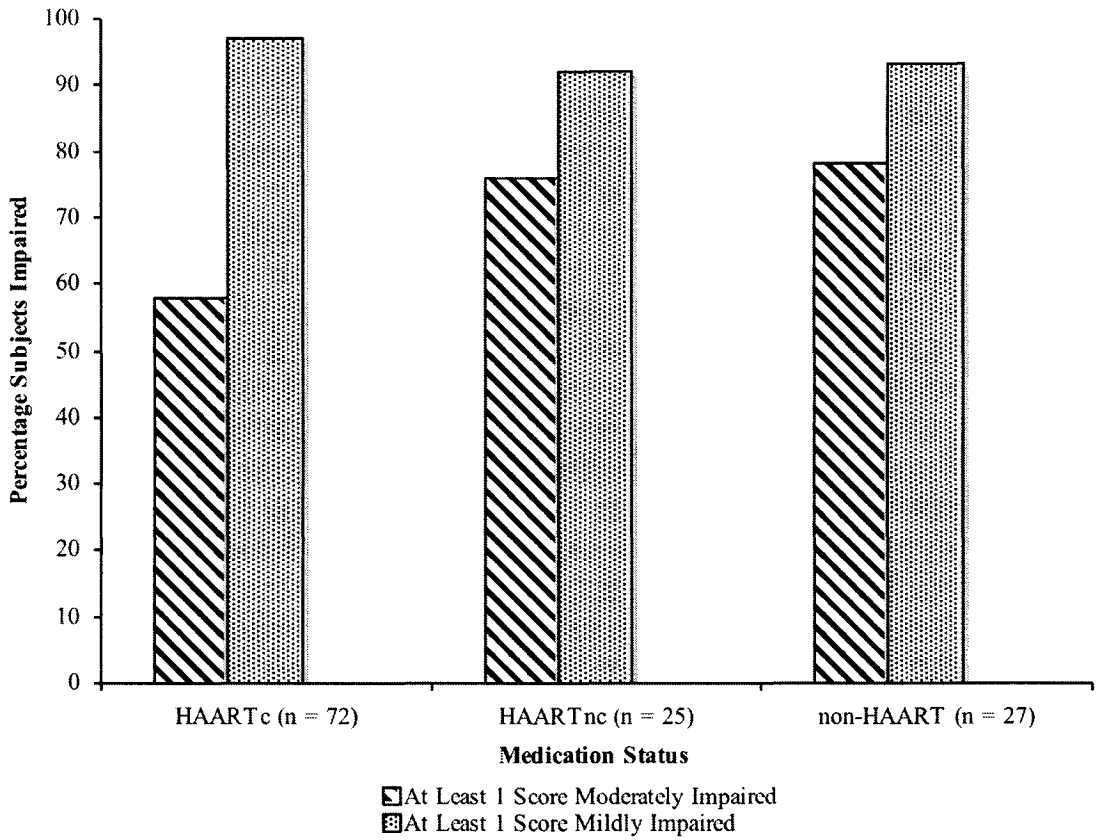


Figure 7. Percentage of Impaired Subjects by Medication Group at Time 2.

To further investigate the pattern of impaired test scores, the frequency with which HAART-any and non-HAART subjects showed at least mild impairment for each neuropsychological test is shown in Figure 8. For this purpose, the mildly impaired test scores were collected only from subjects who had a Global Deficit Score in the impaired range (i.e., > 0.50) in order to reflect a pattern of individual test impairment among subjects with definitive impairment. The four most frequently impaired test scores among non-HAART subjects were the first trial (A1; 100%), fifth trial (A; 100%), long-delay cued recall (LDCR; 100%), and semantic learning score (SEM; 85.7%), from the CVLT. The four most frequently impaired test scores among HAART-any subjects were also the first trial (A1; 97.5%), fifth trial (A; 90%), long-delay cued recall (LDCR; 80%), and semantic learning score (SEM; 82.5%), from the CVLT. All but one of five of the most frequently impaired tests, for both groups, were representative of the Attention and Learning Efficiency domains and the fifth test reflected (verbal) Memory.

These results are consistent with the hypothesis and the expected areas of impairment for HIV individuals, particularly with respect to impairment in the areas of Attention, and Learning Efficiency. It is interesting to note that 100% of impaired, non-HAART subjects showed at least mild impairment on the first, fifth, and long-delay cued recall trials of the CVLT. The impairment ratings methodology (Method B) illustrates of the variation and prevalence of individual deficit with HIV-related cognitive impairment better than the group mean score comparison, which use measures of central tendency (Method A).

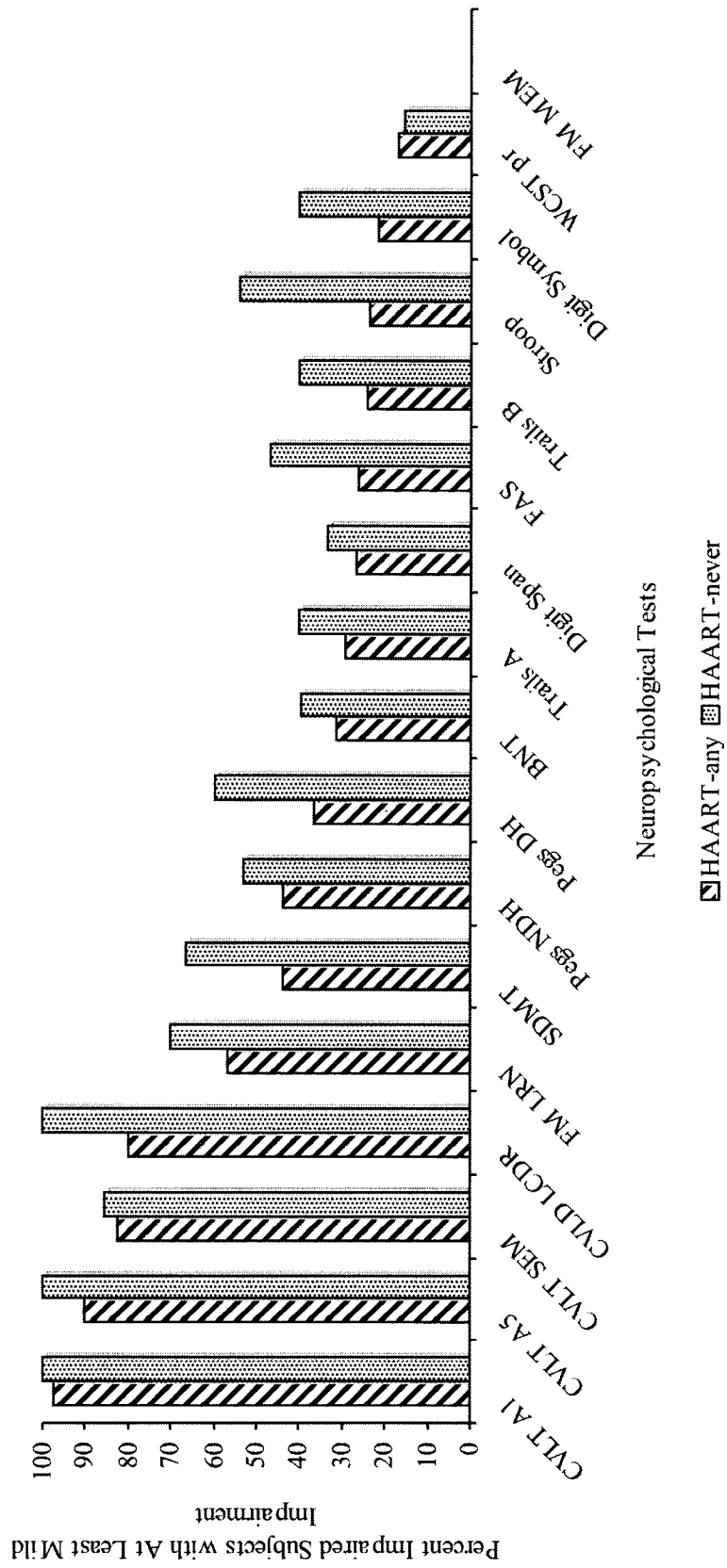


Figure 8. Percentage of Subjects with At Least Mild Impairment by Test at Time 2

The mean Global Deficit Score for subjects in the continuous HAART group fell just above the 0.50 impairment cut-point, at 0.54, whereas the mean Global Deficit Score for the subjects who had never taken HAART was significantly higher (i.e., more impaired), at 0.86, ($F$ (2, 120) = 4.06; $p$ = .02; $d$ = .52). Also, the Global Deficit Score was less impaired for subjects who were on HAART at any point (i.e., HAARTnc) compared to those who were never on HAART ($d$ = .66). Approximately 43% of the continuous HAART subjects (n = 31) fell in the impaired range according to the recommended Global Deficit Score cut-point of 0.50, whereas 56% of the subjects who never took HAART ($n$ = 15) had an impaired Deficit Score. Approximately 44% of the subjects who did not take HAART ($n$ = 11) continuously met the same criteria.

A within-subjects, repeated measures ANOVA on the Global Deficit Score at Time 1 and Time 2 showed a significant relationship between subjects' longitudinal antiretroviral classification and Global Deficit Score, $F$ (2, 120) = 4.73; $p$ = .01, but there was not a significant effect of time (see Figure 9). The Games-Howell post-hoc test showed that both HAARTc and HAARTnc subjects had less impaired Global Deficit Scores compared to subjects in the non-HAART group ($p$ = .02 and .03, respectively).

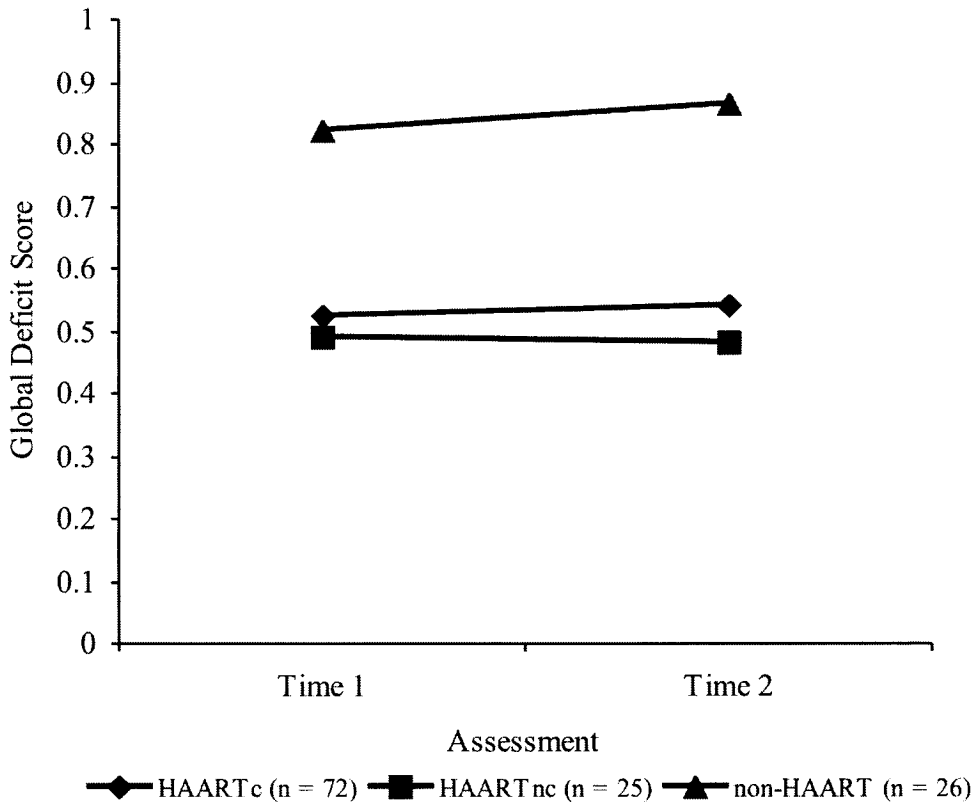


*Figure 9.* Change in Global Deficit Score from Time 1 to Time 2 by Medication Group

*Hypothesis 1 - Reliable Change Index Analyses for All HIV Subjects*

The results for RCI-PE calculations for the T1-T2 sample are shown in Table 13. Over two-thirds of the subjects did not obtain statistically significant change scores on the majority of the neuropsychological variables. The exceptions to this finding were percentage of classifications on the CVLT Long-Delay Cued Recall trial (58.9% No change), CVLT List A Trial 5 recall (46% No change), and the Boston Naming Test (36.0% No change). These findings show that, when practice effects and measurement error were taken into consideration, the majority of subjects' neuropsychological test scores were stable over time. The percentages of subjects who showed significant declines on the neuropsychological tests from Time 1 to Time 2 ranged from 5.6% on the verbal fluency test (i.e., COWA) to 39.2% on the Boston Naming Test, which had quite small critical values. The percentages of subjects who showed statistically significant improvement over time ranged from 1.6% on the first recall trial of the CVLT to 25.8% on the Digit Symbol Test.

Table 13.
*Percentage of All Follow-up Subjects Assigned to Reliable Change Index categories for Neuropsychological Test Scores*

| Domain & Test Variable | *n* | Worse | No Change | Better |
|---|---|---|---|---|
| *Language* | | | | |
| BNT | 125 | 39.2 | 36.0 | 24.8 |
| COWA | 126 | 5.6 | 91.3 | 3.2 |
| *Attention* | | | | |
| Digit Span | 126 | 15.9 | 72.2 | 11.9 |
| CVLT A1 | 124 | 12.1 | 86.3 | 1.6 |
| Stroop | 97 | 14.4 | 79.4 | 6.2 |
| *Learning* | | | | |
| CVLT A5 | 124 | 34.7 | 46.0 | 19.4 |
| CVLT SEM | 124 | 27.4 | 70.2 | 2.4 |
| *Memory* | | | | |
| CVLT LDCR | 124 | 28.2 | 58.9 | 12.9 |
| *Abstraction* | | | | |
| Trails B | 125 | 12.8 | 76.8 | 10.4 |
| WCST | 119 | 8.4 | 88.2 | 3.4 |
| *Motor* | | | | |
| GPT – DH | 126 | 13.5 | 75.4 | 11.1 |
| GPT – NDH | 126 | 18.3 | 73.0 | 8.7 |
| *Psychomotor* | | | | |
| Trails A | 125 | 14.4 | 76.8 | 8.8 |
| SDMT | 119 | 9.2 | 85.7 | 5.0 |
| Digit Symbol | 124 | 8.1 | 66.1 | 25.8 |

*Note.* BNT = Boston Naming Test; COWA = Controlled Oral Word Association; CVLT = California Verbal Learning Test; A1 = List A Trial 1; A5 = List A Trial 5; SEM = Semantic Cluster Ratio; LDCR = Long-Delay Cued Recall; GPT = Grooved Pegboard Test; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand.

The percentages of all HIV+ subjects who demonstrated domain specific neuropsychological changes are presented in Table 14. The area of cognitive functioning with the highest percentage of worsening scores was the Learning domain (42.7% Worse) whereas the highest percentage of improved scores was shared between the Language domain (15.3% Better) and the Language domain (15.3% Better).

Table 14.
*Percentage of all Follow-up Subjects Assigned to Reliable Change Index Categories for Domain Scores and Profile*

| **Domain & Test Variable** | ***n*** | **Worse** | **No Change** | **Better** |
|---|---|---|---|---|
| Language | 124 | 42.7 | 41.9 | 15.3 |
| Attention | 124 | 8.1 | 84.7 | 7.3 |
| Learning | 124 | 42.7 | 41.9 | 15.3 |
| Memory | 124 | 28.2 | 58.9 | 12.9 |
| Abstraction | 118 | 17.8 | 71.2 | 11.0 |
| Motor | 126 | 22.2 | 64.3 | 13.5 |
| Psychomotor | 125 | 7.2 | 86.4 | 6.4 |
| Profile | 127 | 26.0 | 64.6 | 9.4 |

The percentage of all HIV+ subjects who had overall profiles that worsened, improved, or had no significant change is also presented in Table 14. Approximately 65% of the subjects ($n$ = 82) showed relative stability in terms of their overall neuropsychological test profile from Time 1 to Time 2. However, approximately one quarter of the subjects ($n$ = 33) showed a reliable trend to worsening cognitive

performance based on the sum of their neuropsychological test score changes. In contrast, just fewer than 10% ($n$ = 12) showed reliable improvement on enough tests to earn a Better overall classification according to the Reliable Change Index calculations.

*Hypothesis 1 - Reliable Change Index Analyses by Medication Status*

The percentages of individuals who demonstrated significant change on neuropsychological tests for the HAARTc, HAARTnc, and non-HAART groups were compared using chi square analyses. Once again, subjects with neuromedical comorbidities were excluded from the analyses for theoretical reasons and because of the clarification of findings in the previous analyses when these subjects were removed. It should be noted that the distribution of subjects across the combination of antiretroviral status and Reliable Change Index outcome was not sufficient to meet the assumptions of the chi square test for almost two-thirds of the neuropsychological tests (i.e., there were less than 5 expected frequencies in more than 20% of the cells). The same problem was encountered for two of the domain distributions but not for the overall profile distributions. The uneven distributions stemmed predominantly from unexpectedly low frequencies of subjects in many of the Worse and Better categories compared to the No Change category.

In an effort to augment the cell sizes and improve the distribution of subjects per cell, the subjects who had been excluded on the basis of neuromedical comorbidity were temporarily reintroduced to the sample. However, it was decided that the pattern of uneven cell distributions did not change enough to justify the inclusion of subjects with known comorbidities. Furthermore, there were no substantial differences in the results obtained when the subjects were included or excluded. Ultimately, they were excluded in order be consistent with the sample used in the Method A group mean score analyses.

The antiretroviral medication groups were collapsed into two groups: HAART-any (i.e., HAARTc and HAARTnc) and non-HAART in order to increase the sample size per cell. This resulted in a reduction to 6 out of 15 tests that did not meet the chi square criteria for interpretation. The results of these analyses are presented in Table 15. Significant differences between subjects who had HAART at some point and those who never had HAART were found on the Symbol Digit Modalities test, $\chi^2(2, n = 115) = 6.3$, $p = .04$). Specifically, more non-HAART subjects fell in the Worse category than would

be expected while none achieved a performance that could be classified as Better. In contrast, fewer subjects who were on HAART at some point demonstrated Worse performance over time than would be expected by chance. Trends in the expected direction (i.e., HAART associated with Better outcome) were observed on two other tests of attention (i.e., Stroop, CVLT A1). There was a higher percentage of non-HAART subjects in the Worse category compared to HAART-any subjects for all but five tests.

Table 15.
*Percentage of Follow-up Subjects Assigned to Reliable Change Index Categories for Test Scores for HAART-any and Non-HAART Groups*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| *Language* | | | | | |
| BNT | Non-HAART | 27 | 40.7 | 25.9 | 33.3 |
| | HAART-any | 95 | 38.9 | 38.9 | 22.1 |
| COWA | Non-HAART | 26 | 7.7 | 88.5 | 3.8 |
| | HAART-any | 97 | 5.2 | 91.8 | 3.1 |
| *Attention* | | | | | |
| Digit Span | Non-HAART | 26 | 19.2 | 65.4 | 15.4 |
| | HAART-any | 97 | 14.4 | 74.2 | 11.3 |
| CVLT A1 | Non-HAART | 26 | 15.4 | 84.6 | 0.0 |
| | HAART-any | 95 | 11.6 | 86.3 | 2.1 |
| Stroop | Non-HAART | 17 | 23.5 | 70.6 | 5.9 |
| | HAART-any | 77 | 10.4 | 83.1 | 6.5 |
| *Learning* | | | | | |
| CVLT A5 | Non-HAART | 26 | 26.9 | 50.0 | 23.1 |
| | HAART-any | 95 | 35.8 | 45.3 | 18.9 |
| CVLT SEM | Non-HAART | 26 | 30.8 | 65.4 | 3.8 |
| | HAART-any | 95 | 25.3 | 72.6 | 2.1 |
| *Memory* | | | | | |
| CVLT LDCR | Non-HAART | 26 | 30.8 | 53.8 | 15.4 |
| | HAART-any | 95 | 27.4 | 61.1 | 11.6 |
| *Abstraction* | | | | | |
| Trails B | Non-HAART | 26 | 19.2 | 65.4 | 10.4 |
| | HAART-any | 96 | 10.4 | 80.2 | 9.4 |
| WCST | Non-HAART | 23 | 17.4 | 73.9 | 8.7 |
| | HAART-any | 94 | 6.4 | 91.5 | 2.1 |

*Cont'd*

Table 15 *Continued*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| *Motor* | | | | | |
| GPT - DH | Non-HAART | 27 | 18.5 | 66.7 | 14.8 |
| | HAART-any | 96 | 11.5 | 78.1 | 10.4 |
| GPT - NDH | Non-HAART | 27 | 18.5 | 70.4 | 11.1 |
| | HAART-any | 96 | 18.8 | 72.9 | 8.3 |
| *Psychomotor* | | | | | |
| Trails A | Non-HAART | 26 | 19.2 | 69.2 | 11.5 |
| | HAART-any | 96 | 13.5 | 78.1 | 8.3 |
| SDMT * | Non-HAART | 23 | 21.7 | 78.3 | 0.0 |
| | HAART-any | 94 | 6.4 | 87.2 | 6.4 |
| Digit Symbol | Non-HAART | 25 | 8.0 | 80.0 | 12.0 |
| | HAART-any | 96 | 8.3 | 62.5 | 29.2 |

*Note.* BNT = Boston Naming Test; COWA = Controlled Oral Word Association; CVLT = California Verbal Learning Test; A1 = List A Trial 1; A5 = List A Trial 5; SEM = Semantic Cluster Ratio; LDCR = Long-Delay Cued Recall; GPT = Grooved Pegboard Test; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand.

* $p < .05$

Table 16 shows the percentage of follow-up subjects who performed Better, Worse, or had No Change over time in terms of the neuropsychological domain variables according to medication status. The differences between subjects who were on HAART at some point and those who were never on HAART were not statistically significant for any of the neuropsychological domain scores, which failed to support the hypotheses. There were trends in the expected directions for the Attention and Psychomotor domains, which was consistent with predictions. Although not statistically significant, the trend for the overall profile rating also fell in line with expectations. That is, more subjects who were on HAART at some point were assigned a Better profile classification than expected compared to subjects who were never on HAART.

Table 16.
*Percentage of Follow-up Subjects Assigned to Reliable Change Index Categories for Domain Aggregates and Profile for HAART-Any and Non-HAART Groups*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| Language | Non-HAART | 25 | 40.0 | 28.0 | 32.0 |
| | HAART-any | 91 | 36.3 | 40.7 | 23.1 |
| Attention | Non-HAART | 26 | 15.4 | 80.8 | 3.8 |
| | HAART-any | 95 | 6.3 | 85.3 | 8.4 |
| Learning | Non-HAART | 26 | 34.6 | 53.8 | 11.5 |
| | HAART-any | 95 | 43.2 | 40.0 | 16.8 |
| Memory | Non-HAART | 26 | 30.8 | 53.8 | 15.4 |
| | HAART-any | 95 | 27.4 | 61.1 | 11.6 |
| Abstraction | Non-HAART | 22 | 22.7 | 59.1 | 18.2 |
| | HAART-any | 94 | 16.0 | 74.5 | 9.6 |
| Motor | Non-HAART | 27 | 25.9 | 55.6 | 18.5 |
| | HAART-any | 96 | 20.8 | 66.7 | 12.5 |
| Psychomotor | Non-HAART | 26 | 11.5 | 84.6 | 3.8 |
| | HAART-any | 96 | 6.3 | 86.5 | 7.3 |
| Profile | Non-HAART | 27 | 22.2 | 70.4 | 7.4 |
| | HAART-any | 97 | 25.8 | 63.9 | 10.3 |

*Hypothesis 1 - Reliable Change Index Analyses with CD4 as Moderator*

The percentages of individuals who demonstrated significant change on neuropsychological tests for the HAART-any and non-HAART groups were compared again, after stratification for initial CD4 cell count. The distribution of subjects across the combination of antiretroviral and RCI outcome variable levels was not sufficient to meet the assumptions of the chi square test for many of the neuropsychological tests when considering only those subjects with LOW CD4 Cell counts (i.e., there were less than 5 expected frequencies in more than 20% of the cells). Therefore, the results of the analyses should be interpreted with caution. There were no substantial differences in the results obtained when subjects with neuromedical comorbidities were included or excluded so these subjects remained excluded for consistency with previous analyses.

There were no statistically significant relationships between Reliable Change Index outcome and antiretroviral status for subjects who had HIGH initial CD4 cell counts (see Appendix B). Among subjects who had LOW initial CD4 counts (i.e., < 200 cells/μL), subjects who were never on HAART were more likely to be assigned a Worse rating than expected by chance alone on one test of attention (i.e., Digit Span, $\chi^2$ (2, $n$ = 42) = 8.7, $p$ = .01), one test of psychomotor speed (Symbol Digit Modalities Test, $\chi^2$ (2, $n$ = 38) = 6.7, $p$ = .03), one test of learning (i.e., CVLT SEM, $\chi^2$ (2, $n$ = 41) = 4.2, $p$ = .04), and on one test of abstraction (i.e., WCST, $\chi^2$ (2, $n$ = 31) = 12.4, $p$ = .002; see Tables 17 and 18).

The relative superiority of HAART on the first three tests is consistent with the hypotheses. With respect to cognitive domains, more non-HAART subjects than expected ended up in the Worse Abstraction category compared to HAART-any subjects, $\chi^2$ (2, $n$ = 40) = 7.8, $p$ = .02. The trends for HAART superiority were also evident in the Attention, Learning, and Psychomotor domains, as predicted, as well as in the Language domain, which was not expected. The trend of having more HAART subjects than expected in the Better group and more non-HAART subjects in the Worse group was also evident in terms of overall profile score. However, all of these findings must be interpreted with caution given lack of adequate cell size.

Table 17.
*Percentage of Follow-up Subjects Assigned to Reliable Change Index Categories for Test Scores for HAART-Any and Non-HAART groups – LOW CD4*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| *Language* | | | | | |
| BNT | Non-HAART | 6 | 83.3 | 16.7 | 0.0 |
| | HAART-any | 34 | 32.4 | 47.1 | 20.6 |
| COWA | Non-HAART | 6 | 0.0 | 83.3 | 16.7 |
| | HAART-any | 36 | 8.3 | 88.9 | 2.8 |
| *Attention* | | | | | |
| Digit Span * | Non-HAART | 6 | 66.7 | 33.3 | 0.0 |
| | HAART-any | 36 | 13.9 | 72.2 | 13.9 |
| CVLT A1 | Non-HAART | 6 | 16.7 | 83.3 | 0.0 |
| | HAART-any | 35 | 17.1 | 80.0 | 2.9 |
| Stroop | Non-HAART | 3 | 0.0 | 100 | 0.0 |
| | HAART-any | 28 | 7.1 | 78.6 | 14.3 |
| *Learning* | | | | | |
| CVLT A5 | Non-HAART | 6 | 66.7 | 33.3 | 0.0 |
| | HAART-any | 35 | 25.7 | 54.3 | 20.0 |
| CVLT SEM * | Non-HAART | 6 | 50.0 | 50.0 | 0.0 |
| | HAART-any | 35 | 14.3 | 85.7 | 0.0 |
| *Memory* | | | | | |
| CVLT LDCR | Non-HAART | 6 | 33.3 | 50.0 | 16.7 |
| | HAART-any | 35 | 20.0 | 65.7 | 14.3 |
| *Abstraction* | | | | | |
| Trails B | Non-HAART | 6 | 33.3 | 50.0 | 16.7 |
| | HAART-any | 35 | 14.3 | 80.0 | 5.7 |
| WCST * | Non-HAART | 5 | 40.0 | 40.0 | 20.0 |
| | HAART-any | 35 | 5.7 | 91.4 | 2.9 |

*Cont'd*

Table 17 *Continued*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| *Motor* | | | | | |
| GPT – DH | Non-HAART | 6 | 0.0 | 66.7 | 33.3 |
| | HAART-any | 35 | 8.6 | 80.0 | 11.4 |
| GPT – NDH | Non-HAART | 6 | 0.0 | 66.7 | 33.3 |
| | HAART-any | 35 | 20.0 | 68.6 | 11.4 |
| *Psychomotor* | | | | | |
| Trails A | Non-HAART | 6 | 16.7 | 50.0 | 33.3 |
| | HAART-any | 35 | 14.3 | 80.0 | 5.7 |
| SDMT * | Non-HAART | 5 | 60.0 | 40.0 | 0.0 |
| | HAART-any | 33 | 12.1 | 78.8 | 9.1 |
| Digit Symbol | Non-HAART | 5 | 0.0 | 80.0 | 20.0 |
| | HAART-any | 35 | 11.4 | 68.6 | 20.0 |

*Note.* BNT = Boston Naming Test; COWA = Controlled Oral Word Association; CVLT = California Verbal Learning Test; A1 = List A Trial 1; A5 = List A Trial 5; SEM = Semantic Cluster Ratio; LDCR = Long-Delay Cued Recall; GPT = Grooved Pegboard Test; Trails = Trail Making Test; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand.

* $p < .05$

Table 18.
*Percentage of Follow-up Subjects Assigned to RCI Categories for Domain Aggregates and Profile for HAART-Any and Non-HAART groups – LOW CD4*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| Language | Non-HAART | 6 | 66.7 | 33.3 | 0.0 |
| | HAART-any | 33 | 30.3 | 51.5 | 18.2 |
| Attention | Non-HAART | 6 | 33.3 | 66.7 | 0.0 |
| | HAART-any | 35 | 8.6 | 80.0 | 11.4 |
| Learning | Non-HAART | 6 | 66.7 | 33.3 | 0.0 |
| | HAART-any | 35 | 34.3 | 45.7 | 20.0 |
| Memory | Non-HAART | 6 | 33.3 | 50.0 | 16.7 |
| | HAART-any | 35 | 20.0 | 65.7 | 14.3 |
| Abstraction * | Non-HAART | 5 | 40.0 | 20.0 | 40.0 |
| | HAART-any | 35 | 20.0 | 74.3 | 5.7 |
| Motor | Non-HAART | 6 | 0.0 | 50.0 | 50.0 |
| | HAART-any | 35 | 17.1 | 65.7 | 17.1 |
| Psychomotor | Non-HAART | 6 | 33.3 | 66.7 | 0.0 |
| | HAART-any | 35 | 11.4 | 85.7 | 2.9 |
| Profile | Non-HAART | 6 | 33.3 | 66.7 | 0.0 |
| | HAART-any | 36 | 19.4 | 69.4 | 11.1 |

* $p < .01$

*Hypothesis 2*

Table 19 shows the percentage of subjects who obtained Worse, No Change, and Better Reliable Change Index overall profile scores according to consistency of HAART regimen (i.e., taking HAART continuously or taking HAART non-continuously).

Table 19.
*Percentage of Subjects Demonstrating Statistically Significant Change Across the Profile According to Consistency of HAART status from Time 1 to Time 2*

| Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|
| HAARTnc | 25 | 24.0 | 60.0 | 16.0 |
| HAARTc | 72 | 26.4 | 65.3 | 8.3 |

*Note.* HAARTnc = HAART non-continuous; HAARTc = HAART continuous

There was no significant difference in Reliable Change Index profile outcome between subjects on HAART continuously and those not on HAART continuously. Similarly, there was no difference between the mean profile change scores (e.g., +2, -3, +1, etc.) for these two groups (i.e., HAARTc mean = -0.8; HAARTnc mean = - 1.0). These results do not support the hypothesis that HAART duration would be associated with cognitive outcome, although HAART duration was measured only indirectly by continuity of antiretroviral regimen.

The correlation between Reliable Change Index profile outcome and antiretroviral potency was not significant either ($r$ = .045). These findings differ from those of Letendre et al., (2004) who suggested that an improvement in Global Deficit score over time was indirectly associated with the ability of CSF-penetrating drugs to better suppress CSF viral load. On the other hand, Cysique et al. (2004) observed a positive relationship between antiretroviral potency and cognitive improvement only for a subset of HIV+ individuals (i.e., those with mild impairment) and only for one area of cognitive impairment (i.e., verbal memory) and not in terms of a composite score.

*Reliable Change Index Analyses – Post Hoc Comparisons*

Post-hoc analyses were undertaken in an attempt to examine the characteristics of the individuals on HAART who displayed improvements in cognition over time compared to those whose cognitive performance declined. For the purposes of this exploratory analysis, subjects who achieved a better rating on at least two neuropsychological tests were included in the Better group. Of the subjects who were on HAART continuously, 38.9% displayed Worse cognitive performance over time, despite the demonstrated relationship between continuous HAART and improvement in systemic HIV demonstrated earlier in this study. At the other end of the spectrum, 16.7% of the subjects on HAART were characterized through RCI analyses as Better neuropsychological performance over time. Demographic characteristics and HIV disease markers for these subjects are presented in Table 20.

There were no significant differences between the groups in terms of age, education, estimated IQ score, length of time since HIV diagnosis, or initial CD4 cell count. The HAARTc subjects who demonstrated Better cognitive performance over time had significantly worse Global Deficit Scores at Time 1 ($M$ = .90), than the subjects in the No Change ($M$ = 0.39) or Worse ($M$ = 0.52) groups, $F$ (2, 69) = 5.8; $p$ = .005; $d$ = *.77 to .80*).

Although it was not statistically significant, there was a higher representation of Asymptomatic subjects, and fewer subjects with AIDS diagnoses, among the subjects who had demonstrated improved cognitive profiles. Interestingly, of the 12 HAARTc subjects who were classified as improving over time by RCI analyses, 10 had Asymptomatic or Symptomatic diagnoses at baseline and 6 of those 10 subjects also had impaired Global Deficit Scores. It seems that, for these otherwise Asymptomatic and Symptomatic subjects, there was observable cognitive benefit to being on a HAART protocol despite having CD4 cell counts above the often recommended threshold for initiating HAART (i.e., 200 cells/μl).

Table 20.
*Demographic and HIV Characteristics of HAARTc Subjects According to Reliable Change Index Outcome Group*

| | **Worse** ($n$ = 28) | **No Change** ($n$ = 32) | **Better** ($n$ = 12) |
|---|---|---|---|
| Age | 43.5 | 42.2 | 41.6 |
| Education | 15.1 | 14.3 | 15.2 |
| ANART IQ | 116.1 | 118.6 | 118.3 |
| Time HIV + (years) | 8.5 | 8.8 | 7.6 |
| CD4 T1 | 392 | 364 | 440 |
| GDS T1 ** | 0.52 | 0.39 | 0.90 |
| % Asymp | 3.7 | 3.1 | 25.0 |
| % Symp. | 44.4 | 34.4 | 41.7 |
| % AIDS | 51.9 | 62.5 | 33.3 |
| % Undet. Viral | 61.5% | 63.3% | 70.0% |

*Note.* Asymp. = Asymptomatic; Symp. = Symptomatic; CD4 T1 = CD4 cell count at Time 1 (cells/µl); GDS T1 = Global Deficit Score at Time 1; Undet.Viral = Undetectable Viral Load.
**$p < .01$, Better > No Change, Worse

*Hypothesis 3*

In order to investigate what variables might have influenced whether a subject attained a Better or Worse profile status, a binary logistic regression analysis was performed. Predictor variables were chosen based on the theoretically driven hypotheses and from the post-hoc analyses of the Reliable Change Index outcome data. In order to

maximize the cell size in the dependent variable categories, subjects were included who attained reliably changed performances on at least two neuropsychological tests. The predictor variables chosen for the model were estimated IQ score (ANART score), CD4 cell count status over time (i.e., improving or high stable versus worsening or low stable), neuromedical comorbidity status (present/absent), and initial Global Deficit Score. The latter variable was not part of the original hypothesis but was suggested as an influential factor by way of the post-hoc analyses described above. Subjects who were on HAART at any point were included in the analyses ($n$ = 75). Evaluation of the adequacy of expected frequencies for categorical predictors revealed no need to restrict model of goodness-of-fit tests.

The model fit using the four predictors was significantly different than a model using a constant only, $\chi^2$ (4, $n$ = 75) = 18.9, $p$ = .001, indicating that the predictors, as a set, reliably distinguished between subjects with a Better outcome and subjects who showed a Worse cognitive outcome. Hosmer and Lemeshow goodness of fit data revealed a reliable fit of the four-factor model to the observed data. Table 21 shows the regression coefficients, standard errors, Wald statistics, odds ratios, and 95% confidence intervals for odds ratios, for each of the four predictors and for the model constant. Estimated IQ score, initial Global Deficit Score and CD4 cell status over time were significant predictors within the model. Analyzing a model without the presence these three variables did not reveal a significant difference between full model and constant, confirming the importance of these variables. Prediction success indicated that 96.2% of the Worse subjects and 52.2% of the Better subjects correctly predicted, for an overall success rate of 83%.

The odds ratio for the CD4 cell response variable, at 5.74, indicated that the odds of demonstrating reliable cognitive improvement were almost six times as high for a HAART-any subject who displayed improving or high stable CD4 cell counts than for a subject who had a low stable or worsening CD4 cell count. In addition, for every unit increase in initial Global Deficit Score (i.e., increasing impairment), the odds of a HAART-any subject showing improved cognitive performance increased by a factor of 5.18. In addition, there was a less substantial but statistically significant contribution to the model from the estimated IQ score. Specifically, for every unit increase in estimated

IQ, the odds of being in the Better group increased by a factor of 1.12. Although it contributed to the prediction model as a whole, there was no unique contribution of neuromedical comorbidity to cognitive outcome. The integration of these predictor variables in a significant model, as well as their unique contributions to risk, supports the hypothesis that individual variables are important factors in mediating cognitive response to HAART.

Table 21.
*Logistic Regression Analysis of RCI Profile Outcome as a Function of Demographic, Neuromedical and HIV Status Variables*

| Variables | B | Standard Error | Wald Test (z-ratio) | *p* | Odds Ratio | 95% Conf. Inter | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower | Upper |
| ANART IQ* | 0.12 | .049 | 5.48 | .019 | 1.12 | 1.02 | 1.24 |
| CD4 Cell Response* | 1.74 | .738 | 5.60 | .018 | 5.74 | 1.35 | 24.40 |
| Comorbidity | -0.49 | .691 | 0.51 | .477 | .61 | .158 | 2.34 |
| GDS T1** | 1.64 | .576 | 8.15 | .004 | 5.18 | 1.68 | 16.01 |
| *constant* | -16.73 | 4.97 | 6.58 | .010 | .000 | | |

*Note*. ANART IQ = National Adult Reading Test estimated IQ; GDS T1= initial Global Deficit Score.
* $p < .05$, ** $p < .01$.

# Discussion

Neurocognitive problems, ranging from mild and isolated deficits to more obvious dementia, have been demonstrated throughout the course of HIV-infection. The relationship between HAART and HIV morbidity/mortality has been well documented but the longitudinal impact of HAART on neurocognitive functioning is not as clear. There is an emerging concern that HAART may be effective in treating HIV at the systemic level, but that a concomitant improvement in neurological and neurocognitive functioning is less certain. The purpose of the current study was to investigate the impact of HAART on neurocognitive functioning over time, to identify variables that influence changes in cognition among individuals with HIV, and to highlight methodological considerations relating to the identification of cognitive impairment and change in individuals with HIV.

*Cognitive Performance and Change on HAART*

It was hypothesized that subjects on HAART would have better cognitive skills compared to those not on a HAART regimen, particularly in terms of tests of psychomotor efficiency but also on tests of attention and learning ability. This hypothesis was supported to varying degrees. At Time 1 and Time 2, cross-sectional analyses showed that HAART subjects had significantly better scores on one measure of motor speed and coordination (Grooved Pegboard Test; dominant hand).

Psychomotor slowing has been described as a hallmark feature of HIV-associated impairment (Sacktor et al., 2000) and this finding was also seen in the current study in terms of cross-sectional analyses. It could be argued that statistical significance on just one test out of a battery of 15 tests was due to chance, but the Grooved Pegboard Test superiority was evident at both assessments and the relationship between HAART and improved psychomotor speed has been well documented in terms of differentiating HIV+ from seronegative individuals (Miller et al., 1990; van Gorp, Miller, Satz, & Visscher, 1989), particularly at the Symptomatic disease stage (Reger, Welsh, Razani, Martin, & Boone, 2002). Group mean scores on other tests of psychomotor ability in the current

investigation (i.e., Symbol Digit Modalities Test and Trail Making Test –Part A) did not differ significantly in cross-sectional analyses but small to medium effect sizes in the expected direction were evident (range $d$ = *.02* to *.50)*. Therefore, the current results appear to reflect the expected effect of HAART on psychomotor and motor efficiency.

Neuroimaging studies have shown that abnormalities on the Grooved Pegboard Test (and other tests of psychomotor efficiency) correlate with subcortical atrophy in individuals with HIV (e.g., Hestad et al., 1993). The observation that psychomotor speed is superior for individuals on a HAART regimen suggests that the underlying subcortical circuitry is not only vulnerable to HIV-driven pathology but sensitive to treatment as well.

The specific mechanism by which recovery of function takes place is unknown. Sacktor et al. (2000) speculated that a HAART-related improvement in psychomotor speed could be attributed directly to viral suppression in the CNS, by way of antiretroviral medication crossing the blood-brain barrier and reducing viral burden. Alternatively, Sacktor et al. suggested that CNS improvement might be related to a reduced presence of HIV in the brain due to decreased levels elsewhere in the body.

In the both cases, the hypothesis is that the elimination or reduction of virus from the cerebral compartment would have a direct impact on the restoration of cognitive functioning. However, some researchers have observed that suppression of viral activity by HAART, as assessed in the plasma or the CSF, does not correlate directly with inactivation of HIV-related cognitive impairment (Chang et al., 2003; Sevigny et al., 2004). Given what is known, and suspected, about the toxic and deleterious effects of the neuroimmune reaction itself (Bell, 2004), it is possible that much of a HAART-induced cognitive improvement arises from a halting or reversal of the toxic immune and inflammatory processes, rather than from viral reduction or neuronal repair per se (Cohen et al., 2001).

In terms of the hypothesis about better attention and learning efficiency, cross-sectional analyses of group mean scores on the three tests of attention and two tests of learning all showed (non-significant) differences in the predicted direction. Given that the effect sizes in HIV research are often small (Reger, Welsh, Razani, Martin, & Boone,

2002), the small to medium effect sizes identified in the current study might translate to statistical significance if larger sample sizes were available.

In addition to the hypothesized changes, cross-sectional analyses also showed that individuals on HAART had better scores on Trail Making Test – Part B and WCST (perseverative responses) than individuals not on HAART, at Time 1 but not Time 2. It was not specifically predicted that subjects who were on HAART would show a higher group mean score for these two tests, which were categorized as measures of abstraction. However, combination antiretroviral superiority for individuals with HIV has been demonstrated on Trail Making Test – Part B elsewhere (Cohen et al., 2001; Sacktor et al., 1999). The worse scores for non-HAART subjects in the current study could reflect part of an executive dysfunction pattern. Executive functioning problems have been reported as part of the constellation of potential HIV problems (Reger, Welsh, Razani, Martin, & Boone, 2002), particularly in later stages of HIV, and half of the current sample had an AIDS diagnosis, indicating late stage disease.

Alternatively, the superiority of these scores for HAART subjects could reflect improvements in the psychomotor component of Trail Making Test - Part B. None of the cross-sectional or longitudinal HAART studies reviewed for the current investigation reported specifically on scores from the WCST. It is possible that individuals on HAART out-performed their non-HAART counterparts on this task due to general improvements in attention and efficiency, rather than primary improvement in executive functioning.

With respect to profile ratings, the Global Deficit Scores indicated superior cognitive standing for subjects on HAART compared to those not on HAART at Time 1 and Time 2. That is, a more consistent relationship between antiretroviral status and cognitive functioning was apparent at the global profile level. This falls in line with the idea that an aggregate or summary score offers better reliability in detecting impairment than component scores (Heaton et al., 2001) and supports the sensitivity of using the Global Deficit Score in treatment efficacy research.

The Global Deficit Scores placed the HAART subjects, as a group, just at the borderline of the Impaired range (i.e., 0.52 to 0.54). These group scores are quite similar to the group Global Deficit Score for all HIV+ subjects in the Carey et al. (2004) predictive validity study (i.e., 0.51), and thus show cross-sample consistency in HIV

research. In the current study, non-HAART subjects' Global Deficit scores fell beyond the suggested impairment cut-off with scores of 0.72 and 0.86 at Time 1 and Time 2, respectively. It is worthy of reiteration that these group mean profile scores suggested borderline and mild range deficits as a group, consistent with the expectations of using measures of central tendency to characterize individuals with a range of impairment. Using this method to characterize level of functioning can potentially mask important inter-individual variability.

The prevalence and pattern of neuropsychological impairment for all subjects groups unfolded more dramatically when the frequency of mild and moderate individual impairment ratings were calculated. By using the individual impairment method to describe impairment, it became obvious that the vast majority of people in the current study exhibited problems with attention and learning efficiency, as well as psychomotor impairment. This finding helped to support the hypothesis that HAART would be associated with better performance in these areas. The CVLT yielded the most frequently impaired test scores for all HIV individuals, regardless of HAART status, suggesting that it is a particularly sensitive instrument for capturing subtle problems in HIV.

In addition to impairment on measures of attention and learning (i.e., CVLT A1, CVLT A5, CVLT SEM), many subjects in both medication groups demonstrated at least mild impairment on the long-delay cued recall from the CVLT. The subcortical model of HIV impairment would not predict a deficit in verbal retention, unless it was reflecting more pervasive dementia among a group of individuals in late-stage HIV (Reger, Welsh, Razani, Martin, & Boone, 2002). As previously mentioned, the sample in the current study had a relatively high proportion of individuals with AIDS (i.e., 50%) so it is possible that low CVLT long-delay cued recall scores were reflecting significant and advanced memory impairment in a cohort who might have cortical as well as subcortical problems.

On the other hand, impairment on the CVLT long-delay cued recall measure has been demonstrated by HIV+ individuals within a broader range of disease stage and impairment levels (Delis et al., 1995; Murji et al., 2003). Although the CVLT long-delay cued recall measure is primarily a measure of verbal retention (Delis, Kramer, Kaplan, & Ober, 1987), a low score could also reflect problems with initial learning efficiency. In

terms of individual score analysis, it would be important to look at the long-delay cued recall score in the context of the total number of words acquired during the learning and immediate recall trials so that a 'savings' score could be calculated and interpreted as a measure of retention.

Finally, in terms of verbal memory scores and HIV, it is also worthwhile to note that Brew (2004) recently asserted that the presentation of HIV-associated dementia is changing in the post-HAART era, such that more cortically based dysfunction is becoming more common. Cysique et al. (2004) recently identified a HAART-based superiority on the short- and long-delay recall measures from the CVLT. These researchers asserted that there is an increasing involvement of the hippocampus in otherwise mild, HIV-associated cognitive dysfunction in the post-HAART era. Brew speculated that, with increasing survival rates, HIV may lead to more Alzheimer-like changes in cognition. Therefore, the findings in the current study related to impaired CVLT long-delay cued recall could reflect a trend to more cortical-like cognitive deficits in post-HAART era patient samples.

With respect to methodology comparisons, several salient points are illustrated by the current data. When between-group mean score comparisons are used to assess the heterogeneous nature of cognitive deficits in a disease such as HIV, there is a substantial risk that individuals with above average and below average scores will cancel each other out on measures of central tendency. This worry is compounded in HIV research in which impairment can be not only varied or "spotty", but also subtle, at any stage of disease. The assessment of impairment prevalence becomes imperative if one is to understand the impact of HIV disease on an individual level.

As expected, group mean scores in the current study failed to identify areas of compelling deficit compared to normative standards, whereas prevalence ratings allowed a more accurate portrayal of the breadth and depth of cognitive impairment. Even within the continuous HAART group, over half of the sample had at least one moderately impaired test score at Time 2. These findings are consistent with those of Burgess et al. (1994) and provide further evidence for the importance of using individual impairment ratings to expand upon group mean neuropsychological test findings when drawing conclusions about HIV cognitive impairment patterns.

To summarize, the cross-sectional comparison of antiretroviral regimens in the current study provided some support for the hypothesis of HAART being associated with better neuropsychological performance compared to non-HAART regimens in the areas of psychomotor speed, attention, learning and certain aspects of memory. Not unexpectedly, cross-sectional analyses did not yield many significant results in terms of group comparisons, with the exception of global ratings. Furthermore, the measures of central tendency concealed meaningful variations in individual impairment patterns. Nonetheless, these results are generally consistent with existing cross-sectional HAART research (Ferrando et al., 1998; Sacktor et al., 1999; Starace et al., 2002).

*Reliable Change Index Methodology*

One of the original contributions of the current investigation was using Reliable Change Index (with practice effect) methodology to establish critical values for determining change on 15 commonly used neuropsychological test measures for use in an HIV population. Establishing statistically sound parameters for detecting reliable change, while incorporating corrections for practice effects, is an important methodological advance in the longitudinal investigation of HIV. The specific critical values were generated from published normative test-retest data and, where available, from studies using HIV research control subjects with clinically relevant test-retest intervals. The normative data presented here can thus assist clinicians and researchers when determining whether longitudinal cognitive change is reliably different from chance, particularly for individuals and groups who are demographically similar to the current sample.

It is worthwhile recalling that one advantage of the Reliable Change Index methodology is that significant change can be documented relative to individuals' initial performance. In other words, by this methodology one can theoretically identify improvement or deterioration in individuals who have average or better functioning at the outset. In contrast, group mean methodology allows identification only of individuals who are impaired relative to normative data.

One of the first areas of neuropsychology in which Reliable Change Index techniques were employed was in epilepsy surgery research (Chelune, Naugle, Lueders, Sedlak, & Awad, 1993; Martin et al., 2002). The techniques, though, are becoming more

commonly used with other neuropsychological populations including traumatic brain injury (Millis et al., 2001), coronary artery bypass surgery (Kneebone, Andrew, Baker, & Knight, 1998), schizophrenia (Moritz, Iverson, & Woodward, 2003), and aneurysm surgery treatment (Towgood, Ogden, & Mee, 2005). In a recent article, the HIV Neurobehavioral Research Center group described preliminary HIV research techniques using modified Reliable Change Index methodology (Woods et al., in press). At the time of writing, however, the current study is the first of its kind to use Reliable Change Index and practice effects methodology to investigate the association between antiretroviral medication and cognitive change in HIV. This methodology should be replicated with other HIV samples, using a broad range of tests when possible, to further extend the applicability of critical values for evaluating change.

*Changes in Cognition Over Time for HIV Subjects*

Applying the Reliable Change Index methodology to the current sample indicated that the majority of HIV+ subjects demonstrated stable cognitive performance over time on all but two neuropsychological tests and on all but one domain. This is an important statistic in terms of conceptualizing the natural history of HIV in the current treatment era. In terms of overall profile change, nearly 65% of the current sample did not exhibit either significant improvement or decline over the test-retest period, as reflected by scores on neuropsychological tests. The 65% statistic is consistent with a general clinical impression that the majority of subjects at this particular HIV research and clinical program did not show clinically meaningful shifts in cognitive status from one visit to the next.

In addition to identifying cognitive stability, Reliable Change Index analysis is able to identify individuals who exhibit unusual cognitive deterioration or improvement. When assessing the impact of an antiretroviral regimen on a group of individuals with cognitive deficits, one is particularly interested in identifying true positive change in order to ascribe improvement to the given intervention. That is, there can be a danger in mistaking improved test scores for improvements in cognition when the change could be largely due to measurement error and/or practice effects. Similarly, it is important to be able to identify how much of a loss in test score from Time 1 to Time 2 is statistically noteworthy, so that treatment modifications or other clinical intervention can be initiated.

In terms of HIV research in particular, deterioration in neurocognitive performance has been associated with increased risk of dementia, AIDS and death (Sacktor et al., 1996).

When all HIV subjects were considered together in the current sample, approximately 25% of the subjects showed overall cognitive decline and approximately 10% of subjects were said to have improved. The prevalence of Worse and Better change scores varied from 6% to 43% across neuropsychological tests, reflecting a variability similar to that observed in static HIV-related impairment.

It was hypothesized that individuals on HAART would be over-represented in the Better outcome groups while non-HAART subjects would more frequently show a Worse performance over time. This hypothesis was minimally supported. There was a significant relationship between HAART and cognitive outcome on one of the psychomotor tests (i.e., Symbol Digit Modalities Test). This finding supported the hypothesis and was consistent with previous research in terms of HAART-associated changes and psychomotor speed (Cohen et al., 2001; Ferrando, Rabkin, van Gorp, Lin, & McElhiney, 2003; Ferrando et al., 1998; Honn, 2003; Robertson et al., 2004; Sacktor et al., 1999; Sacktor et al., 2000; Suarez et al., 2001; Tozzi et al., 1999; Tozzi et al., 2001). There were no significant relationships between antiretroviral status and cognitive change on any other neuropsychological tests. Similarly, although the Attention and Psychomotor domains showed trends in the expected direction, there were no statistically significant associations between being on HAART and showing improved cognitive status.

The trends toward improved attention and psychomotor performance on HAART are consistent with significant improvements in these domains identified in previous longitudinal research (Cohen et al., 2001; Ferrando, Rabkin, van Gorp, Lin, & McElhiney, 2003; Ferrando et al., 1998; Robertson et al., 2004; Sacktor et al., 1999; Sacktor et al., 2000; Suarez et al., 2001; Tozzi et al., 1999). However, researchers have also identified significant relationships between HAART and other tests of cognitive functioning including measures of language and visual-spatial ability (Cohen et al., 2001; Robertson et al., 2004; Tozzi et al., 1999), learning and memory (Cysique, Maruff, & Brew, 2004; Ferrando et al., 1998; Honn, 2003; Robertson et al., 2004; Suarez et al., 2001; Tozzi et al., 1999), fine-motor speed (Cohen et al., 2001; Honn, 2003), abstraction

(Cohen et al., 2001; Ferrando et al., 1998; Robertson et al., 2004; Tozzi et al., 1999), and global measures of overall performance (Ferrando et al., 1998; Tozzi et al., 1999). These relationships were not supported by the current research. However, the results of the current study are similar to those of Starace et al. (2002) who found no cognitive advantage for individuals on HAART compared to non-HAART.

The minimal association seen between HAART status and cognitive change in the current study might have occurred for a variety of reasons. In terms of antiretroviral experience, the subjects were not necessarily assessed at the beginning of a HAART regimen. Impaired subjects who might experience a HAART-associated cognitive improvement might have already done so prior to the initial testing such that their initial test score could represent the beginning of a 'plateau' phase. Tozzi et al. (2001) identified sustained cognitive improvement after HAART initiation for 4 of 17 test scores after 15 months, but gains continued for only two of the tests (Grooved Pegboard Test and Stroop Colour-Word) after two and one-half years.

Brew (2004) raised a similar idea by noting that subjects with AIDS dementia complex were able to achieve viral suppression on a changed HAART regimen, but they did not exhibit an associated improvement in cognitive symptoms. Brew hypothesized that although these individuals responded to the new drug regimen with renewed viral suppression, they had achieved maximum benefit from HAART in terms of cognitive improvement prior to the trial. With the publication of more longitudinal HAART studies such this investigation, and the increasing survival of individuals with HIV, researchers will gradually be able to identify and better appreciate the long-term durability, or lack thereof, of HAART.

Another possible factor in explaining the weak relationship between HAART and cognitive change in this study is that, although the antiretroviral groups were collapsed in order to boost statistical power, the resulting sample sizes could have still been inadequate for detecting subtle impairments, as well as subtle changes over time in cognitive status. Because the effect sizes associated with changes in neuropsychological functioning are known to be small, it stands to reason that changes in neuropsychological status would be similarly difficult to detect. In addition, the lack of association between HAART and cognitive outcome might be explained by the fact that the Reliable Change

Index (with practice effect) methodology is, by design, a conservative procedure and could be insensitive to detecting more subtle changes in cognitive ability.

In support of the latter argument, Heaton et al., (2001) noted that using a categorical (i.e., Better/No Change/Worse) outcome approach potentially obscures more subtle individual changes in ability. That is, there is no information about the magnitude of cognitive change when subjects are assigned to an outcome category on the basis of a given confidence interval. An attempt was made in the current study to investigate the relationship impact of antiretroviral status by using a continuous change score, but the hypothesized relationship was still not observed. Nevertheless, future Reliable Change Index investigations within an HIV population might benefit from using a continuous change score (e.g.,Woods et al., in press).

It is also possible that the chosen levels of the independent variable, and/or the collapsing of antiretroviral groups to increase statistical power, obscured a potential relationship between specific subtypes of HAART regimens and changes in cognitive functioning. Perhaps a more definitive relationship between HAART and cognitive change would have arisen if the non-HAART comparison group had excluded individuals who were not taking any antiretrovirals or if the non-continuous and continuous HAART groups had not been combined. Similarly, a more definitive relationship might have arisen if different subtypes of HAART regimens were considered as separate independent variables (i.e., HAART with a protease inhibitor versus HAART without a protease inhibitor). Although an insufficient sample size precluded an accurate interpretation of the aforementioned strategies, applying them did not yield significant results or trends when attempted in an exploratory manner.

*HAART Potency*

It was not possible to obtain accurate information about all subjects' duration on HAART prior to enrolment in the study. There was no support for a hypothesis about potential dose-response relationships when alternate analyses were undertaken. Given the possibility of a 'plateau' effect in reaching maximum cognitive benefit after some continuous time period on HAART, it would be worthwhile to track this information for future research. With respect to HAART potency, regimens that contained more CSF-penetrating antiretrovirals did not have a significant relationship with cognitive outcome.

This finding is consistent with existing literature that also failed to find a beneficial long-term impact of HAART drugs that were thought to have a higher CNS penetrance potential (Cysique, Maruff, & Brew, 2004; Dougherty, Skolasky, & McArthur, 2002; Sacktor et al., 2001b), although differing cognitive outcomes have been identified for subsets of HIV-positive individuals (Cysique, Maruff, & Brew, 2004).

The lack of CSF-penetrating antiretroviral influence in the current study might relate to the complexity with which CNS penetrance is evaluated (Strazielle & Ghersi-Egea, 2005) and to how the drugs were coded, or to the fact that the potency relationship might be evident in just one area of cognitive functioning rather than across the profile (Ferrando, Rabkin, van Gorp, Lin, & McElhiney, 2003). Alternatively, a relationship between increasing antiretroviral potency and cognitive change might not be demonstrable because factors other than cerebral viral level (e.g., immune response) are influencing cognitive change, as discussed below. The number of issues raised here underscores the need for further research related to drug potency and efficacy.

*Influence of Mediating Variables on Cognitive Change*

The dissociation observed in this study between HAART effectiveness at the systemic level and lack of relationship at the cognitive level supports concerns in the HIV community that combination antiretroviral therapy is capable of reducing and preventing HIV morbidity and mortality but is not necessarily as successful at preventing or reversing cognitive decline (Dore et al., 1999). There have been surveillance reports of HAART leading to reduced incidence of HIV-associated dementia and improvement in dementia-related symptoms (e.g., Dougherty, Skolasky, & McArthur, 2002; Maschke et al., 2000; Sacktor et al., 2001b). These studies have also linked HAART with cognitive improvement in terms of specific cognitive abilities. However, given the failure to find a direct benefit in the current study, it is possible that the improvements are temporary and/or the benefits may be evident only for a subset of individuals.

Support for the theory that HAART is associated with cognitive changes only for certain subsets of individuals with HIV was provided by the results of the Reliable Change Index analyses after CD4 stratification. Among HIV+ individuals with low initial CD4 cell counts at Time 1 (i.e., <200 cells/μl), there was a significant relationship between HAART status and cognitive outcome on four individual neuropsychological

tests (i.e., Digit Span, Symbol Digit Modalities Test, CVLT Semantic Learning, and WCST – perseverative responses). These tests represent skills in attention, learning, abstraction, and psychomotor speed and support the hypothesis of an association between HAART and cognitive change in tests reflective of subcortical functioning.

For all of these tests, non-HAART subjects were over-represented in the Worse cognitive outcome category, confirming subcortical pattern deficits for individuals with HIV, and identifying a deterioration for subjects who were not on HAART. In four out of the six significant results, HAART subjects were over-represented in the Better outcome group as well. This suggests that the underlying subcortical processes are both vulnerable to HIV-mediated deterioration and amenable to HAART-driven recovery. Potential mechanisms of recovery as discussed earlier include reduction in viral load, either directly at the CNS level or indirectly at the systemic level, and reversal of toxic immune response.

When interpreting the weak relationship in the current study between HAART and cognitive change, it is imperative to note that there were clear and statistically significant relationships between HAART and changes in markers of HIV status. That is, HIV+ individuals who were on HAART, either continuously or non-continuously, had clinically meaningful changes in immune functioning (i.e., achieved or maintained a CD4 cell count of 200 or higher, below which the risk for opportunistic infection rises significantly) and disease burden (i.e., achieved or maintained undetectable plasma viral load status). In other words, HAART was shown to have the expected effect in combating HIV at the systemic level but was not associated with a concomitant improvement in cognitive functioning, except for a few measures, and only in a subsample of individuals with low initial CD4 cell counts.

*Predictors of Cognitive Change in HIV Subjects on HAART*

With respect to subgroups in the current study, approximately one quarter (26.4%) of the HIV+ individuals who were on HAART continuously showed significant deterioration in cognition during the course of this investigation. This finding runs contrary to the bulk of post-HAART era research, which generally shows improved cognitive functioning on HAART. The findings highlight a potentially important

dissociation between systemic and cognitive functioning in HIV for a subset of HIV+ individuals.

Continuous HAART subjects whose cognitive profiles deteriorated were not differentiated from HAART subjects whose cognitive performance improved on the basis of demographic variables (i.e., age, education, estimated IQ) or HIV marker variables (i.e., initial CD4 cell count, plasma viral load status, length of time HIV+, disease stage). However, there was a significant difference based on initial level of performance in that subjects in the Better outcome group tended to have worse initial performance. This relationship between initial level of performance and cognitive change was confirmed through the Logistic Regression analysis discussed next.

For individuals on HAART, Better cognitive outcome was predicted by a combination of initial level of impairment, immune response (as measured by changes to CD4 cell count status), estimated IQ and neuromedical comorbidity. Differences in initial impairment level and immune system response were each associated with an approximately 5-fold increase in the odds of demonstrating cognitive improvement.

The relationship between initial impairment level and cognitive outcome indicated that cognitive improvement was not universally achieved for all individuals on HAART. Rather, HAART subjects who were impaired at Time 1 were more likely to experience reliable cognitive improvement. This finding supports pre-HAART era research that showed relative improvements in psychomotor speed for individuals with poor initial functioning (Sacktor et al., 1999). It suggests that being on HAART has the potential to reverse significant cognitive deficits, particularly when other factors such as immune response and estimated IQ are added to the equation. Based on the current results, it appears that HAART is less likely to improve cognitive functioning for HIV+ individuals with less severe impairment. These findings could indicate that some sort of impairment threshold needs to be reached before the benefits of HAART on cognition are evident.

HAART subjects' estimated IQ score was also found to have a small but significant contribution in the model predicting cognitive outcome, which provided support for the hypothesis that certain demographic variables would be related to cognitive change. As part of the overall model, higher estimated IQ scores increased the

chances of achieving a Better cognitive outcome on HAART. One explanation for this finding could be that IQ is a surrogate marker for 'cognitive reserve' (Stern, 2002) and that individuals with lower IQ scores have less cerebral capacity to overcome the onset of cognitive deficits associated with HIV.

An alternate explanation for the observed relationship could be that individuals with lower IQ scores might have more difficulty following the antiretroviral medication regimens, which are notoriously complicated and demanding, and thus would not experience maximum treatment efficacy. This theory is speculative but is supported by the findings of Kleeberger et al., (2004) who found that 'less than college' education was one significant predictor of decreasing medication adherence. It must be kept in mind, though, that in the current study the difference between estimated IQs for the Worse versus Better groups was not significant in isolation. Furthermore, mean scores for the groups were above 115, which reflected in part the demographics of this particular sample. In order to better understand any potential association between IQ scores and cognitive outcome, it would be important to include subjects with a broader range of IQ scores in future investigations.

It is particularly noteworthy that the logistic regression analyses showed that HIV+ individuals' chances of experiencing cognitive improvement while on HAART could be predicted by their measured immune response in terms of CD4 cell response. That is, individuals whose CD4 cell counts either achieved or maintained a level greater than 200 cells/μl were over 5 times as likely to achieve improved cognitive status overall. This result runs contrary to what was originally hypothesized and adds important information to our growing knowledge about the complicated relationships observed between antiretroviral medication, immune functioning, and cognitive change. It implies that the effectiveness of HAART in reconstituting immune functioning, at an individual level, is a prognostic indicator of improvement in cognitive functioning.

The fact that systemic immune response was indirectly related to cognitive outcome in the current sample diverges from previous HAART research that failed to show a relationship between systemic markers of immune response and cognitive change (Ferrando, Rabkin, van Gorp, Lin, & McElhiney, 2003; Honn, 2003). However, the current findings complement and extend research from the pre-HAART era that

identified a predictive relationship between static measures of the body's early immune status and incidence of cognitive impairment (Marcotte et al., 2003; Sacktor et al., 2003).

It is important to note that the relationship between immune response and cognitive outcome found in the current study was not based on initial CD4 cell count alone or on raw change in CD4 cell count over time, but on individuals' functional immune status over time (i.e., maintain or meet a threshold level). In other words, in the subset of HAART subjects who showed a demonstrable improvement in immunologic functioning, a concomitant improvement in cognitive status was also observed. It is possible that the current relationship was identified because of the way that systemic immune system change was categorized.

These data suggest that knowing an individual's static viral load, or one-time CD4 count, does not seem to indicate whether or not cognitive change will ensue. Instead, an assessment of the individual's *response* to HAART, in terms of immune system reconstitution or viral suppression seems to be necessary to predict whether an individual will show cognitive gains associated with the antiretroviral therapy. This dynamic classification of immune response should be considered essential for future research in order to identify and characterize the subgroup of HIV+ individuals for whom HAART is having the intended immunologic effect.

*Theoretical Implications*

The current finding of a reduced, or indirect, association between the systemic effects of HAART and the cognitive effects of HAART provides further evidence toward the theory that the HAART is not managing neurobehavioural sequelae as successfully as it is managing systemic recovery. That is, the body's successful response to HAART in terms of decreased viral burden and improved immune functioning was not matched directly by improvement in cognitive status. This discrepancy gives rise to at least two possible theories to account for dissociation between systemic and cognitive response to HAART.

One possibility is that despite HAART's ability to suppress plasma viral load and restore immune function, viral load in the CNS (not measured in the current study) could be proliferating unchecked and continuing to cause neurocognitive impairment. This premise is based on a theory of HIV compartmentalization, in which some systemic cells,

and cells in the CNS itself, provide a sanctuary for HIV (Bell, 2004; Pialoux et al., 1997). In this scenario, the virus is thought to escape blood-based antiretroviral attack and remains inactive for an indefinite period of time, undetectable with common viral assessment mechanisms. These inactive HIV cells are thought to lie dormant and are never completely eradicated by HAART, particularly not in the CNS where there is a reduced ability of certain antiretroviral drugs to cross the blood-brain barrier (Strazielle & Ghersi-Egea, 2005).

Therefore, it is possible that an inactive concentration of HIV in the CNS could lead to neurological and neurocognitive impairment in an individual who has otherwise achieved viral suppression and reconstituted immune system functioning. Support for this theory is generated by the fact that HIV virus has been detected in the brain of individuals with HIV even from the earliest stages of disease (Bell, 2004), and most combination antiretroviral drugs do not cross the blood-brain barrier well (Strazielle & Ghersi-Egea, 2005). Letendre et al. (2004) identified a relatively greater cognitive improvement only for HIV+ individuals in whom CSF-based viral suppression had been achieved. Letendre's group noted that the differences in cognitive improvement did not depend on initial viral load or on plasma viral suppression, but on presumed suppression of viral activity in the CNS.

However, systemic markers of HIV disease have been associated with neuropsychological performance elsewhere (Marcotte et al., 2003; Marra et al., 2003; Tozzi et al., 2005), which suggests that the dissociation between systemic HAART response and cognitive change in the current study might be better explained by alternate theories.

*Primary versus Secondary Effects of HIV*

One of the mechanisms by which HIV is thought to affect neurological functioning is not through direct attack on neurons but by way of an HIV-triggered immune reaction that sets up a cascade of neurotoxic events. Specifically, Bell (2004) has suggested that HIV-associated cognitive impairment correlates with immune system activation in terms of activated, if not necessarily infected, macrophages. In this proposed model, successful suppression of virus, at either the systemic or the CNS level, does not halt the toxic neuroimmunologic reaction already underway. Therefore, change in

cognitive status is not related directly to viral burden, or by extension to factors affecting viral burden such as antiretroviral medication. In contrast, cognitive change would be linked to changes in immune system functioning.

Although they used different markers of immune response and viral activity than those employed in the current study, Sevigny et al. (2004) observed a dissociation between viral status and cognitive impairment (in terms of progression to HIV-associated dementia). Consistent with the current research finding of CD4 response being predictive of cognitive change, Sevigny et al. noted a relationship between markers of immune functioning (i.e., CSF monocyte chemoattractant protein-1 and plasma tumor necrosis factor alpha) and time to dementia onset.

Taken together, these findings provide support for the theory that HIV-associated cognitive impairment is not uniquely driven by viral-dependent neuropathology but may be related more to the effects of HIV-induced immune system activation. Specifically, Sevigny et al. proposed that immune system responders (i.e., microglia, monocytes, macrophages) are primed by the presence of HIV infection. Despite antiretroviral suppression of viral activity, the prolonged immune response is eventually toxic to the CNS. In the end, the risk of developing dementia remains, despite successful viral suppression. Further work in this area, particularly in terms of collaborations between neuropsychologists and cell biologists could contribute greatly to our understanding of the neuropathology of HIV and the associated neurobehavioural changes. In addition, increasing our understanding of the possible mechanisms for cognitive deterioration and recovery in HIV could augment our knowledge of these processes in other neuropsychological disorders with neuroimmune components, such as multiple sclerosis and even Alzheimer's disease (Minagar et al., 2002).

*Research Implications*

The results of this study offer new longitudinal data on the natural history of HIV, and on the relationship between combination antiretroviral treatment and cognition in the post-HAART era. Application of the Reliable Change Index model (with practice effects) provides HIV researchers with a tool for investigating cognitive changes that better accounts for potential problems with measurement error and practice effects than the methodology used in previous investigations. The importance of using impairment

prevalence measurements (rather than group mean scores) to identify cognitive impairment at the individual level was highlighted and supports previous HIV research recommendations by Heaton et al. (1995).

This investigation also supports and extends the idea that there are significant individual differences that mediate cognitive outcome when individuals with HIV are on HAART. Based on the current results, it appears that individual differences such as pre-existing demographic variables (i.e., IQ), acute HIV-related deficit (i.e., initial impairment level, initial immune system compromise), and response to treatment (i.e., immune system reconstitution) all play a role in determining an individual's course of HIV disease progression or recovery. In contrast, the potential influence of non-HIV health problems (i.e., neuromedical comorbidity) and sub-characteristics of HAART itself (i.e., continuity, potency) on cognitive change were not supported by the current results.

Given the importance of individual characteristics in mediating cognitive outcome in the current study, and the lack of consensus with respect to the relationship between longitudinal HAART and cognitive change in the existing literature, it would be wise to continue to identify HIV subgroups based on HIV-related and HAART-related characteristics. There still exists a worrisome percentage of individuals who show cognitive deterioration despite being on HAART. It will be important to determine what unique characteristics distinguish that group from those who improve or remain stable over time.

*Clinical Implications*

With respect to more direct clinical implications, the Reliable Change Index analyses and critical values presented here also allow neuropsychologists to identify more accurately which HIV+ subjects have demonstrated significant decline or significant improvement over time on an individual basis. Recognizing true change in an individual's cognitive profile arms neuropsychologists with a better framework from which to communicate information about cognitive improvement and decline since a previous testing or after beginning of a new antiretroviral treatment regimen.

Furthermore, the base rates of cognitive improvement and deterioration presented in this study, as well as the odds associated with the various predictor variables, will

provide neuropsychologists and other treating clinicians with valuable information about cognitive prognosis for patients with similar demographic characteristics to those described here. One positive clinical message suggested from these results is that some HIV+ individuals who exhibit significant cognitive impairment can be given the encouraging information that if they follow a HAART protocol and their immune system seems to react by staying healthy or improving, there are good chances that they may experience improved cognitive functioning over time.

Also, knowing that potentially subtle changes in cognition can precede more obvious markers of medical distress (Cherner et al., 2002), primary health care providers can have another piece of information at their disposal for making decisions about adjusting HIV treatment protocols. This can be crucial for maintaining viral suppression once treatment has been initiated.

In terms of being aware of HIV subtypes, it is worthwhile for the treatment provider to note that there was a select group of subjects in the current study that had Asymptomatic or Symptomatic HIV and also demonstrated significant cognitive impairment at Time 1. Some of these individuals, who were already on HAART despite having a CD4 cell count below 200, did achieve Better cognitive functioning by Time 2. This finding shows a potential value of prescribing HAART in terms of cognitive improvement for cognitively impaired individuals whose CD4 cell count is higher than 200, which might otherwise disqualify them for initiating the HAART protocol.

*Comments and Limitations*

In addition to the specific caveats noted in preceding sections, there are some other limitations to the interpretation and generalizability of the results related to antiretroviral assignment, sample size, subject stratification, and external validity. Because of the nature of HIV research and the ethical obligations associated with treatment, the assignment of HIV+ subjects to antiretroviral treatment groups was not by way of randomized allocation but rather based on clinical need. This medical necessity makes using a within-subject methodology such as the Reliable Change Index an important tool for research with HIV and similar populations but limits causal inferences.

Also with respect to medication regimen, no specific measure beyond self-report was used to assess antiretroviral adherence, which is an important factor in maintaining

adequate viral suppression. There are indications that individuals with HIV risk increased mortality if they are unable to maintain adherence levels above 75%, and there are additional concerns about possible drug resistance when less than perfect adherence is maintained (Hogg et al., 2002). It would be worthwhile to incorporate more specific measures of adherence into future longitudinal research. However, Hogg et al., note that the accuracy of assessing true adherence is controversial.

Efforts were made, statistically, to ensure that follow-up participants did not differ in important ways from subjects who only attended one session. However, the trends for follow-up subjects having higher education, higher estimated IQs and being slightly older potentially limit the generalizability of these findings to individuals with vastly dissimilar demographic information. Similarly, in order to derive information about the natural history of HIV from as homogenous a group as possible, the analyses were conducted without subjects who acknowledged some form of neuromedical comorbidity. In reality, many HIV patients do have another medical or influential condition that would be considered a confound in the current study. Therefore, the results might not apply directly to those individuals.

Because the cognitive impairments that accompany (some) individuals with HIV can be quite subtle, particularly large sample sizes may be needed to have sufficient power to uncover the subtle impairment or small differences between groups. The data from the current study was collected from consecutive referrals to a major metropolitan HIV clinic over a period of eight years. Arguably, this sample taps into the vast majority of individuals with HIV in the city. However, the current era of HIV research, and of antiretroviral treatment in particular, is sufficiently dynamic that a cohort of this size can become undesirably heterogeneous as a result of continuing changes in drug regimens and medication treatment guidelines. In order to increase sample size, while maintaining a reasonably cohesive and homogenous cohort, it could be advantageous to initiate multi-centre studies for future research so that data could be gathered simultaneously.

A potential limitation of the Reliable Change Index methodology used in the current study is that confidence intervals were applied to all subjects regardless of initial level of performance. In a comparison of various statistical methods to detect clinical change, Heaton et al. (2001) noted that low initial test performances showed more

variability in the Time 2 test score, whereas individuals with better initial performance showed reduced score variability at Time 2. The general concern in applying an averaged confidence interval to all subjects is that lack of change might be mislabeled in individuals who have extremely impaired test scores, and true change might be missed among those subjects with extremely high initial test scores. Because cognitive deficits are mild for the majority of individuals with HIV, it is reasonable to use the averaged confidence interval approach. Nevertheless, it would be interesting to examine any potential changes in assignment of cognitive change, if future Reliable Change Index critical values were calculated using normative data that had been stratified according to initial impairment level. To this end, current research at the St. Michael's Neurobehavioural Research Unit is focusing on assessing a broad range of 'control' subjects who are Asymptomatic and do not show cognitive impairment.

Further consideration of impairment level could also be given to the large number of subjects in the No Change group. Although these individuals did not show reliable change over the course of the test-retest interval, this is not to say that individuals in the No Change group were exhibiting normal cognitive functioning. Rather, this group would theoretically consist of a mix of impaired, average functioning, and above average individuals who more or less maintained their initial cognitive status. It would be interesting, sample size permitting, to consider stratifying the No Change group by impairment level to see if further differentiation in this outcome category helps to discern subtypes of response to HAART.

With any change in psychometric test score, whether statistically significant or not, it is important to assess the clinical meaningfulness of that change. Reger et al. (2002) noted that HIV-associated cognitive deficits had the most potential for interfering with everyday functioning only at later disease stages. There was no effort made in the current investigation to assess the external validity of the identified changes in cognition in terms of changes in daily functioning. However, neuropsychological impairment, particularly the kinds of subcortical, or frontostriatal impairment seen in HIV, has been associated with problems on psychometric indicators of everyday functioning (Heaton et al., 2004). Because the effect sizes associated with HAART-mediated in neurocognitive change are typically small, it is predicted that identifiable and meaningful changes in

daily functioning would not be obvious unless multiple areas of cognition were involved in this sample. It will be an important task for future researchers to examine whether or not the cognitive decline and improvement identified here are associated with meaningful decline or improvement in everyday functioning, and to investigate the relationship, if any, with mood and quality of life.

*Future Directions*

In addition to the suggestions for future research made in the previous sections, it is important to reiterate that longitudinal research with HAART is still a relatively new field of study. The current investigation was essentially the first of its kind to pair Reliable Change Index methodology with a longitudinal investigation of HAART and cognition in HIV. Therefore, the findings should be replicated with a different sample in order to examine generalizability to the broader HIV population.

The Reliable Change Index calculations generated here were calculated from samples that most closely matched the demographic characteristics of the HIV subjects under investigation. Unfortunately, there is a notable dearth of long-term normative data for some of the neuropsychological tests employed in the battery. To better assist longitudinal neuropsychological research in general, future researchers could attempt to fill these gaps in normative data by providing more test-retest data that is suitable in length for clinical comparisons. HIV research in particular would benefit from the large-scale, longitudinal assessment of unimpaired, Asymptomatic HIV+ individuals who have not begun antiretroviral treatment. In the meantime, researchers who are investigating the effects of HAART on cognition in individuals who differ substantially from the current subject pool should use normative data that better matches their particular sample in terms of gender, race, HIV-infection mode and so forth.

The compartmentalized relationships between antiretroviral status and cognitive change that are detailed here provide some confirmation of previous research findings, but also underscore new areas for investigation. Potentially relevant individual and group characteristics that were not explored in the current study include CSF markers of infection, immune response, and treatment response, sub-classification of neuromedical comorbidity; and age. With respect to age, as the HIV population survives longer on HAART and enters an older age bracket, it will be important to identify any potential risk

factors for cognitive impairment associated with aging itself, as well as any potential interactions with metabolic, neuropsychiatric, and cardiovascular morbidities (Casau, 2005; Valcour et al., 2004). Given this emerging importance of individualized response to HIV disease and to HIV treatment, it would be useful for future researchers to focus on identifying HIV subtypes in an effort to better understand potential group patterns of cognitive recovery and cognitive deterioration.

As the use of HAART approaches the 10-year mark, questions of increased dementia prevalence and ineffective neurological potency are raised in parallel with the findings of reduced or postponed morbidity and mortality. It is apparent that longitudinal research in this area is crucial. Regardless of specific methodology, it is important to continue to work towards the most complete understanding possible of the relationships between HIV, HAART, and cognitive performance, so that theory, research, and practice can all benefit.

# References

Abramason, I. S. (2000). Reliable change formula query: A statistician's comments. *Journal of the International Neuropsychological Society, 6*, 365.

American Academy of Neurology (AAN). (1991). Nomenclature and research case definitions for neurologic manifestations of human immunodeficiency virus-type 1 (HIV-1) infection: Report of a working group of the Amercian Academy of Neurology AIDS Task Force. *Neurology, 41*, 778-785.

Bell, J. E. (2004). An update on the neuropathology of HIV in the HAART era. *Histopathology, 45*, 549-559.

Bornstein, R. A. (1994). Methodological and conceptual issues in the study of cognitive change in HIV infection. In I. Grant, & A. Martin (Eds.), *Neuropsychology of HIV Infection* (pp. 146-160). New York: Oxford University Press.

Brew, B. J. (2004). Evidence for a change in AIDS dementia complex in the era of highly active antiretroviral therapy and the possibility of new forms of AIDS dementia complex. *AIDS, 18*, S75-S78.

Burgess, A. P., Riccio, M., Jadresic, D., Pugh, K., Catalan, J., Hawkins, D. A., et al. (1994). A longitudinal study of the neuropsychiatric consequences of HIV-1 infection in gay men. I. Neuropsychological performance and neurological status at baseline and at 12-month follow-up. *Psychological Medicine, 24*, 885-895.

Butters, N., Grant, I., Haxby, J., Judd, L. L., Martin, A., McClelland, J., et al. (1990). Assessment of AIDS-related cognitive changes: Recommendations of the NIHM workshop on neuropsychological assessment approaches. *Journal of Clinical and Experimental Neuropsychology, 12*, 963-978.

Carey, C. L., Woods, S. P., Gonzalez, R., Conover, E., Marcotte, T. D., Grant, I., et al. (2004). Predictive validity of Global Deficit Scores in detecting neuropsychological impairment in HIV infection. *Journal of Clinical & Experimental Neuropsychology, 26*, 307-319.

Carpenter, C. C., Cooper, D. A., & Fischl, M. A. (2000). Antiretroviral therapy in adults: Updated recommendations of the international AIDS Society-USA panel. *Journal of the American Medical Association, 283*, 381-390.

Casau, N. C. (2005). Perspective on HIV infection and aging: Emerging research on the horizon. *Clinical Infectious Diseases, 41*, 855-863.

Centers for Disease Control. (1982). Current trends update on acquired immune deficiency syndrome (AIDS) - United States. *Morbidity and Mortality Weekly Report, 31*, 507-508.

Centers for Disease Control. (1986). Current trends classification system for Human T-Lymphotorpic virus Type III/Lymphadenopathy-associated virus infections. *Morbidity and Mortality Weekly Report*, *35*, 334-339.

Centers for Disease Control. (1987). Revision of the CDC surveillance case definition for acquired immune deficiency syndrome. *Morbidity and Mortality Weekly Report*, *36*, 1S-16S.

Centers for Disease Control. (1992). 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *Morbidity and Mortality Weekly Report*, *41*, RR-17.

Centers for Disease Control. (1998). Report of the NIH Panel to define principles of therapy of HIV Infection and guidelines for use of antiretroviral agents in HIV-infected adults and adolescents. *Morbidity and Mortality Weekly Report*, *47*, RR-5.

Chang, L., Ernst, T., Witt, M. D., Ames, N., Walot, I., Jovicich, J., et al. (2003). Persistent brain abnormalities in antiretroviral-naive HIV patients 3 months after HAART. *Antiretroviral Therapy*, *8*, 17-26.

Chelune, G. J. (2003). Assessing reliable neuropsychological change. In R. D. Franklin (Ed.), *Prediction in Forensic and Neuropsychology: Sound Statistical Practices* (pp. 123-147). New Jersey: Lawrence Erlbaum Associates.

Chelune, G. J., Naugle, R. I., Lueders, H., Sedlak, J., & Awad, I. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, *7*, 41-52.

Cherner, M., Masliah, E., Ellis, R. J., Marcotte, T. D., Moore, D. J., Grant, I., et al. (2002). Neurocognitive dysfunction predicts postmortem findings of HIV encephalitis. *Neurology*, *59*, 1563-1567.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Cohen, R. A., Boland, R., Paul, R., Tashima, K. T., Schoenbaum, E. E., Celentano, D. D., et al. (2001). Neurocognitive performance enhanced by highly active antiretroviral therapy in HIV-infected women. *AIDS*, *15*, 341-345.

Cysique, L. A., Maruff, P., & Brew, B. J. (2004). Antiretroviral therapy in HIV infection: Are neurologically active drugs important? *Archives of Neurology*, *61*, 1699-1704.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test: Research Edition Manual*. San Antonio: Harcourt Brace Jovanovich, Inc.

Delis, D. C., Peavy, G., Heaton, R. K., Butters, N., Salmon, D. P., Taylor, M. J., et al. (1995). Do patients with HIV-associated minor cognitive/motor disorder exhibit a "subcortical" memory profile? Evidence using the California Verbal Learning Test. *Assessment, 2*, 151-165.

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test-retest reliability and practice effects of Expanded Halstead-Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society, 5*, 346-356.

Dodrill, C. B., & Troupin, A. S. (1975). Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. *Journal of Nervous & Mental Disease, 161*, 185-190.

Dore, G. J., Correll, P. K., Li, Y., Kaldor, J. M., Cooper, D. A., & Brew, B. J. (1999). Changes to AIDS dementia complex in the era of highly active antiretroviral therapy. *AIDS, 13*, 1249-1253.

Dougherty, R. H., Skolasky, R. L., & McArthur, J. C. (2002). Progression of HIV-associated dementia treated with HAART. *AIDS Reader, 12*, 69-74.

Duff, K., Westervelt, H. J., McCaffrey, R. J., & Haase, R. F. (2001). Practice effects, test-retest stability, and dual baseline assessments with the California Verbal Learning Test in an HIV sample. *Archives of Clinical Neuropsychology, 16*, 461-476.

Eggers, C., Hertogs, K., Sturenburg, H., van Lunzen, J., & Stellbrink, H. (2003). Delayed central nervous system action virus suppression during highly active antiretroviral therapy is associated with HIV encephalopathy, but not with viral load drug resistance or poor central nervous system drug penetration. *AIDS, 17*, 1897-1906.

Ellis, R. J., Deutsch, R., Heaton, R. K., Marcotte, T. D., McCutchan, J. A., Nelson, J. A., et al. (1997). Neurocognitive impairment is an independent risk factor for death in HIV infection. *Archives of Neurology, 54*, 416-424.

Evers, S., Rahmann, A., Schwaag, S., Frese, A., Reichelt, D., & Husstedt, I. W. (2004). Prevention of AIDS dementia by HAART does not depend on cerebrospinal fluid drug penetrance. *AIDS Research and Human Retroviruses, 20*, 483-491.

Ferrando, S., Rabkin, J., van Gorp, W., Lin, S.-H., & McElhiney, M. (2003). Longitudinal improvement in psychomotor processing speed is associated with potent combination antiretroviral therapy in HIV-1 infection. *The Journal of Neuropsychiatry and Clinical Neurosciences, 15*, 208-214.

Ferrando, S., van Gorp, W., McElhiney, M., Goggin, K., Sewell, M., & Rabkin, J. (1998). Highly active antiretroviral treatment in HIV infection: Benefits for neuropsychological function. *AIDS, 12*, F65-F70.

Field, A. (2005). *Discovering statistics uisng SPSS*. London: Sage Publications.

Goodwin, G. M., Pretsell, D. O., Chiswick, A., Egan, V., & Brettle, R. (1996). The Edinburgh cohort of HIV-positive injecting drug users at 10 years after infection: A case-control study of the evolution of dementia. *AIDS, 10*, 431-440.

Grant, I., Marcotte, T. D., Heaton, R. K., & the HNRC group. (1999). Neurocognitive complications of HIV disease. *Americal Psychological Society, 10*, 191-195.

Hall, M., Whaley, R., Robertson, K., Hamby, S., Wilkins, J., & Hall, C. (1996). The correlation between neuropsychological and neuroanatomic changes over time in asymptomatic and symptomatic HIV-1-infected individuals. *Neurology, 46*, 1697-1702.

Hamby, S., Bardi, C. A., & Wilkins, J. (1997). Neuropsychological assessment of relatively intact individuals: Psychometric lessons from an HIV+ sample. *Archives of Clinical Neuropsychology, 12*, 545-556.

Health Canada. (2001). *HIV and AIDS in Canada. Surveillance Report to December 31, 2000*. Ottawa, ON: Centre for Infectious Disease Prevention and Control, Health Canada.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). *Wisconsin Card Sorting Test (WCST) manual revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Grant, I., Anthony, W. Z., & Lehman, R. A. W. (1981). A comparison of clinical and automated interpretation of the Halstead-Reitan battery. *Journal of Clinical Neuropsychology, 3*, 121-141.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive Norms for an Expanded Halstead-Reitan Battery: Demographic Corrections, Research Findings, and Clinical Applications*. Odessa, FL.: Psychological Assessment Resources.

Heaton, R. K., Grant, I., Butters, N., White, D. A., Kirson, D., Atkinson, J. H., et al. (1995). The HNRC 500: Neuropsychology of HIV infection at different disease stages. *Journal of the International Neuropsychological Society, 1*, 231-251.

Heaton, R. K., Kirson, D., Velin, R. A., Grant, I., & the HNRC Group. (1994). The utility of clinical ratings for detecting cognitive change in HIV infection. In I. Grant, & A. Martin (Eds.), *Neuropsychology of HIV Infection* (pp. 188-206). New York: Oxford University Press.

Heaton, R. K., Marcotte, T. D., Rivera Mindt, M., Sadek, J., Moore, D. J., Bentley, H., et al. (2004). The impact of HIV-associated neuropsychological impairment on

everyday functioning. *Journal of the International Neuropsychological Society, 10*, 317-331.

Heaton, R. K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M. J., Marcotte, T. D., et al. (2001). Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology, 16*, 75-91.

Hestad, K., McArthur, J. H., Dal Pan, G. J., Selnes, O. A., Nance-Sproson, T. E., Aylward, E., et al. (1993). Regional brain atrophy in HIV-1 infection: Association with specific neuropsychological test performance. *Acta Neurologica Scandinavica, 88*, 112-118.

Hinton-Bayre, A. (2000). Reliable change formula query. *Journal of the International Neuropsychological Society, 6*, 362-363.

Hogg, R. S., Heath, K., Bangsberger, D., Yip, B., Press, N., O'Shaughnessy, M. V., et al. (2002). Intermittent use of triple-combination therapy is predictive of mortality at baseline and after 1 year of follow-up. *AIDS, 16*, 1051-1058.

Honn, V. J. (2003). Medical treatments for HIV infection: Psychological outcomes and moderators (Doctoral dissertation, The Ohio State University, 2003).

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test (2nd ed.)*. Philadephia: Lea & Febiger.

Karlsen, N. R., Reinvang, I., & Froland, S. S. (1993). A follow-up study of neuropsychological function in asymptomatic HIV-infected patients. *Acta Neurologica Scandinavica, 87*, 83-87.

Kleeberger, C., Buechner, J., Palella, F. J., Detels, R., Riddler, S., Godfrey, R., et al. (2004). Changes in adherence to highly active antiretroviral therapy medications in the Multicenter AIDS Cohort Study. *AIDS, 18*, 683-688.

Kneebone, A. C., Andrew, M. J., Baker, R. A., & Knight, J. L. (1998). Neuropsychologic changes after coronary artery bypass grafting: use of reliable change indices. *Annals of Thoracic Surgery, 65*, 1320-1325.

Letendre, S. L., McCutchan, J. A., Childers, M., Woods, S. P., Lazzaretto, D., Heaton, R. K., et al. (2004). Enhancing antiretroviral therapy for human immunodeficiency virus cognitive disorders. *Annals of Neurology, 56*, 416-423.

Levine, A. J., Miller, E. N., Becker, J. T., Selnes, O. A., & Cohen, B. A. (2004). Normative data for determining significance of test-retest differences on eight common neuropsychological instruments. *The Clinical Neuropsychologist, 18*, 373-384.

Lezak, M. D. (1995). *Neuropsychological Assessment*. New York: Oxford University Press.

Marcotte, T. D., Deutsch, R., McCutchan, J. A., Moore, D., Letendre, S. L., Ellis, R. J., et al. (2003). Prediction of incident neurocognitive impairment by plasma HIV RNA and CD4 levels early after HIV seroconversion. *Archives of Neurology, 60*, 1406-1412.

Marra, C., M., Lockhart, D., Zunt, J. R., Perrin, M., Coombs, R. W., & Collier, A. C. (2003). Changes in CSF and plasma HIV-1 RNA and cognition after starting potent antiretroviral therapy. *Neurology, 60*, 1388-1390.

Marra, C., M., Longstreth, W. T., Handsfield, H. H., Townes, B. D., Coombs, R. W., Murphy, V. L., et al. (1996). Neurological manifestations of HIV infection without AIDS: Follow-up of a cohort of homosexual and bisexual men. *Journal of Neuro-AIDS, 1*, 41-65.

Martin, A., & Grant, I. (1994). The challenge of HIV. In I. Grant, & A. Martin (Eds.), *Neuropsychology of HIV Infection* (pp. 359-365). New York: Oxford University Press.

Martin, R., Sawrie, S. M., Gilliam, F., Mackey, M., Faught, E., Knowlton, R., et al. (2002). Determining reliable cognitive change after epilepsy surgery: Development of Reliable Change Indices and standardized regression based change norms of the WMS-III and WAIS-III. *Epilepsia, 43*, 1551-1558.

Maschke, M., Kastrup, O., Esser, S., Ross, B., Hengg, U., & Hufnagel, A. (2000). Incidence and prevalence of neurological disorders associated with HIV since the introduction of highly active antiretroviral therapy (HAART). *Journal of Neurology, Neurosurgery & Psychiatry, 69*, 376-380.

Mauri, M., Sinforiani, E., Muratori, S., Zerboni, R., & Bono, G. (1993). Three-year neuropsychological follow-up in a selected group of HIV-1 infected homosexual/bisexual men. *AIDS, 7*, 241-245.

McArthur, J. C. (1994). Neurological and neuropathological manifestations of HIV infection. In I. Grant, & A. Martin (Eds.), *Neuropsychology of HIV Infection* (pp. 56-107). New York: Oxford University Press.

McArthur, J. C., McDermott, M., McClernon, D. R., St. Hillaire, C., Conant, K., Marder, K., et al. (2004). Attenuated central nervous system infection in advanced HIV/AIDS with combination antiretroviral therapy. *Archives of Neurology, 61*, 1687-1696.

McCaffrey, R. J., Cousins, J. P., Westervelt, H. J., & Martynowicz, M. (1995). Practice effects with the NIMH AIDS abbreviated neuropsychological battery. *Archives of Clinical Neuropsychology, 10*, 241-250.

McSweeny, A. J., Naugle, R. I., Chelune, G. J., & Luders, H. (1993). "T Scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist, 7*, 300-312.

Miller, E. N., Selnes, O. A., McArthur, J. C., Satz, P., Becker, J. T., Cohen, B. A., et al. (1990). Neuropsychological performance in HIV-1 infected homosexual men: The Multicenter AIDS Cohort Study (MACS). *Neurology, 40*, 197-203.

Millis, S. R., Rosenthal, M., Novack, T. A., Sherer, M., Nick, T. G., Kreutzer, J. S., et al. (2001). Long-term neuropsychological outcome after traumatic brain injury. *The Journal of Head Trauma Rehabilitation, 16*, 343-355.

Minagar, A., Shapshak, P., Fujimura, R., Ownby, R., Heyes, M., & Eisdorfer, C. (2002). The role of macrophage/microglia and astrocytes in the pathogenesis of three neurologic disorders: HIV-associated dementia, Alzheimer disease, and multiple sclerosis. *Journal of the Neurological Sciences, 202*, 13-23.

Moritz, S., Iverson, G. L., & Woodward, T. S. (2003). Reliable change indexes for memory performance in Schizophrenia as a means to determine drug-induced cognitive decline. *Applied Neuropsychology, 10*, 115-120.

Murji, S., Rourke, S. B., Donders, J., Carter, S. L., Shore, D., & Rourke, B. P. (2003). Theoretically derived CVLT subtypes in HIV-1 infection: Internal and external validation. *Journal of the International Neuropsychological Society, 9*, 1-16.

Navia, B. A., Cho, E. S., Petito, C. K., & Price, R. W. (1986). The AIDS dementia complex: II. Neuropathology. *Annals of Neurology, 19*, 517-524.

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

Navia, B. A., & Price, R. W. (1987). The acquired immunodeficiency syndrome dementia complex as the presenting or sole manifestation of human immunodeficiency virus infection. *Archives of Neurology, 44*, 65-69.

Pialoux, G., Fournier, S., Moulignier, A., Proveda, J.-D., Clavel, F., & Dupont, B. (1997). Central nervous system as a sanctuary for HIV-1 infection despite treatment with zidovudine, lamivudine, and indinavir. *AIDS, 11*, 1302-1303.

Price, R. W., & Brew, B. J. (1988). The AIDS dementia complex. *The Journal of Infectious Diseases, 158*, 1079-1083.

Public Health Agency of Canada. (2004). *HIV and AIDS in Canada. Surveillance report to December 31, 2004.* Ottawa, ON: Surveillance and Risk Assessment Division, Centre for Infectious Disease Prevention and Control, Public Health Agency of Canada.

Reger, M., Welsh, R., Razani, J., Martin, D., & Boone, K. B. (2002). A meta-analysis of the neuropsychological sequelae of HIV infection. *Journal of the International Neuropsychological Society, 8*, 410-424.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery.* Tucson, AZ: Neuropsychology Press.

Robertson, K., Robertson, W. T., Ford, S., Watson, D., Fiscus, S., Harp, A. G., et al. (2004). Highly active antiretroviral therapy improves neurocognitive functioning. *Journal of Acquired Immune Deficiency Syndromes, 36*, 562-566.

Sacktor, N., Skolasky, R. L., Tarwater, P. M., McArthur, J. C., Selnes, O. A., Becker, J., et al. (2003). Response to systemic HIV viral load suppression correlates with psychomotor speed performance. *Neurology, 61*, 567-569.

Sacktor, N. C., Bacellar, H., Hoover, D. R., Nance-Sproson, T. E., Selnes, O. A., & Miller, E. N. (1996). Psychomotor slowing in HIV infection: a predictor of dementia, AIDS and death. *Journal of NeuroVirology, 2*, 404-410.

Sacktor, N. C., Lyles, R. H., Skolasky, R. L., Anderson, D. E., McArthur, J. C., McFarlane, G., et al. (1999). Combination antiretroviral therapy improves psychomotor speed performance in HIV-seropositive homosexual men. *Neurology, 52*, 1640-1647.

Sacktor, N. C., Lyles, R. H., Skolasky, R. L., Kleeberger, C., Selnes, O. A., Miller, E. N., et al. (2001a). HIV-associated neurologic disease incidence changes: Multicenter AIDS Cohort study, 1990-1998. *Neurology, 56*, 257-260.

Sacktor, N. C., Skolasky, R. L., Lyles, R. H., Esposito, D., Selnes, O. A., & McArthur, J. C. (2000). Improvement in HIV-associated motor slowing after antiretroviral therapy including protease inhibitors. *Journal of NeuroVirology, 6*, 84-88.

Sacktor, N. C., Tarwater, P. M., Skolasky, R. L., McArthur, J. C., Selnes, O. A., Becker, J., et al. (2001b). CSF antiretroviral drug penetrance and the treatment of HIV-associated psychomotor slowing. *Neurology, 57*, 542-544.

Saykin, A. J., Janssen, R., Sprehn, G. C., Kaplan, J. E., Spira, T. J., & O'Connor, B. (1991). Longitudinal evaluation of neuropsychological function in homosexual men with HIV infection: 18-month follow-up. *Journal of Neuropsychiatry, 3*, 286-298.

Schmitt, F. A., Dickson, L., & Brouwers, P. (1994). Neuropsychological response to antiretroviral therapy in HIV infection. In I. Grant, & A. Martin (Eds.), *Neuropsychology of HIV Infection* (pp. 276-294). New York: Oxford University Press.

Selnes, O. A., Galai, N., Bacellar, H., Miller, E. N., Becker, J. T., Wesch, J., et al. (1995). Cognitive performance after progression to AIDS: A longitudinal study from the Multicenter AIDS cohort study. *Neurology, 45*, 267-275.

Selnes, O. A., McArthur, J. C., Royal, W., Updike, M. L., Nance-Sproson, T. E., Concha, M., et al. (1992). HIV-1 infection and intravenous drug use: Longitudinal neuropsychological evaluation of asymptomatic subjects. *Neurology, 42*, 1924-1930.

Selnes, O. A., Miller, E. N., McArthur, J. C., Gordon, B., Munoz, A., Sheridan, K., et al. (1990). HIV-1 infection: No evidence of cognitive decline during the asymptomatic stages. *Neurology, 40*, 204-208.

Sevigny, J. J., Albert, S. M., McDermott, M. P., McArthur, J. C., Sacktor, N., Conant, K., et al. (2004). Evaluation of HIV RNA and markers of immune activation as predictors of HIV-associated dementia. *Neurology, 63*, 2084-2090.

Silberstein, C. H., O'Dowd, M., Chartock, P., Schoenbaum, E. E., Friedland, G., Hartel, D., et al. (1993). A prospective four-year follow-up of neuropsychological function in HIV seropositive and seronegative methadone-maintained patients. *General Hospital Psychiatry, 15*, 351-359.

Smith, A. (1991). *Symbol Digit Modalities Test*. Los Angeles: Western Psychological Services.

Spreen, O., & Strauss, E. (1991). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York: Oxford University Press.

Starace, F., Bartoli, L., Aloisi, M. S., Antinori, A., Narciso, P., Ippolito, G., et al. (2002). Cognitive and affective disorders associated to HIV infection in the HAART era: findings from the NeuroICONA study. *Acta Psychiatrica Scandanvia, 106*, 20-26.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society, 8*, 448-460.

Stern, Y., Liu, X., Marder, K., Todak, G., Sano, M., Ehrhardt, A., et al. (1995). Neuropsychological changes in a prospectively followed cohort of homosexual and bisexual men with and without HIV infection. *Neurology, 45*, 467-472.

Strazielle, N., & Ghersi-Egea, J.-F. (2005). Factors affecting delivery of antiviral drugs to the brain. *Reviews in Medical Virology, 15*, 105-133.

Suarez, S., Baril, L., Stankoff, B., Khellar, M., Dubois, B., Lubetzki, C., et al. (2001). Outcome of patients with HIV-1-related cogntiive impairment on highly active antiretroviral therapy. *AIDS, 15*, 195-200.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics*. New York: Harper Collins Publishers Inc.

Tate, R. L., Rollings, J., Collins, M., Pickering, A., & Jacobson, R. R. (1998). Stability of the Wisconsin Card Sorting Test and the determination of reliability of change in scores. *The Clinical Neuropsychologist, 12*, 348-357.

Temkin, N. R. (2004). Standard error in the Jacobson and Truax Reliable Change Index: The 'classical approach' leads to poor estimates. *Journal of the International Neuropsychological Society, 10*, 899-901.

Temkin, N. R., Heaton, R. K., Grant, I., & Dikmen, S. S. (2000). Reliable change formula query: Temkin et al., reply. *Journal of the International Neuropsychological Society, 6*, 364.

Temkin, N. R., Heaton, R. K., Grant, I., & Dikmen, S. S. (1999). Detecting significant change in neuropsychological test performance: A comparison of four models. *Journal of the International Neuropsychological Society, 5*, 357-369.

Towgood, K., Ogden, J. A., & Mee, E. (2005). Neurological, neuropsychological, and funtional outcome following treatment for unruptured intracranial aneurysms. *Journal of the International Neuropsychological Society, 11*, 522-534.

Tozzi, V., Balestra, P., Galgani, S., Narciso, P., Ferri, F., Sebastiani, G., et al. (1999). Positive and sustained effects of highly active antiretroviral therapy on HIV-1-associated neurocognitive impairment. *AIDS, 13*, 1889-1897.

Tozzi, V., Balestra, P., Galgani, S., Narciso, P., Sampaolesi, A., Antinori, A., et al. (2001). Changes in neurocognitive performance in a cohort of patients treated with HAART for 3 years. *Journal of Acquired Immune Deficiency Syndromes, 28*, 19-27.

Tozzi, V., Balestra, P., Serraino, D., Bellagamba, R., Corpolongo, A., Piselli, P., et al. (2005). Neurocognitive impairment and survival in a cohort of HIV-infected patients treated with HAART. *AIDS Research & Human Retroviruses, 21*, 706-713.

Trennerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *Stroop Neuropsychological Screening Test (SNST)*. Odessa, FL.: Psychological Assessment Resources.

Valcour, V., Shikuma, C., Shiramizu, B., Watters, M., Poff, P., Selnes, O. A., et al. (2004). Higher frequency of dementia in older HIV-1 individuals: The Hawaii aging with HIV-1 cohort. *Neurology*, *63*, 822-827.

van Gorp, W., Lamb, D. G., & Schmitt, F. A. (1993). Methodological issues in neuropsychological research with HIV-spectrum disease. *Archives of Clinical Neuropsychology*, *8*, 17-33.

van Gorp, W., Miller, E. N., Satz, P., & Visscher, B. (1989). Neuropsychological performance in HIV-1-immunocompromised patients: A Preliminary report. *Journal of Clinical & Experimental Neuropsychology*, *11*, 763-773.

Villa, G., Solida, A., Moro, E., Tavolozza, M., Antinori, A., De Luca, A., et al. (1996). Cognitive impairment in asymptomatic stages of HIV infection: A longitudinal study. *European Neurology*, *36*,

Wechsler, D. (1981). *WAIS-R manual.* New York: The Psychological Corporation.

White, D. A., Heaton, R. K., Monsch, A. U., & the HNRC Group. (1995). Neuropsychological studies of asymptomatic human immunodeficiency virus-type 1 infected individuals. *Journal of the International Neuropsychological Society*, *1*, 304-315.

Williams, K. C., & Hickey, W. F. (2002). Central nervous system damage, monocytes and macrophages, and neurological disoders in AIDS. *Annual Review of Neuroscience*, *25*, 537-562.

Woods, S. P., Childers, M., Ellis, R. J., Guaman, S., Grant, I., Heaton, R. K., et al. (in press). A battery approach for measuring neuropsychological change. *Archives of Clinical Neuropsychology*,

Yarchoan, R., Brouwers, P., Spitzer, A. R., Grafman, J., Safai, B., Perno, C. F., et al. (1987). Response of human-immunodeficiency-virus-associated neurological disease to 3'-AZIDO-3'-DEOXYTHYMIDINE. *The Lancet, January 17*, 132-135.

Yeni, P. G., Hammer, S. M., Carpenter, C. C. J., Cooper, D. A., Fischl, M. A., Gatell, J. M., et al. (2002). Antiretroviral treatment for adult HIV infection in 2002: Updated recommendations of the International AIDS Society-USA Panel. *Journal of the American Medical Association*, *288*, 222-235.

Appendix A

Table A1.
*Time 1 Means and Standard Deviations for Neuropsychological Variables for Non-HAART and HAART Groups for All Baseline Subjects (n = 406).*

| | non-HAART | | | HAART | | | Statistics |
|---|---|---|---|---|---|---|---|
| Variable | *n* | Mean | *SD* | *n* | Mean | *SD* | *t/z* value |
| BNT[a] | 132 | 45.5 | 14.6 | 265 | 43.9 | 14.0 | ns |
| COWA[a] | 131 | 43.9 | 9.2 | 269 | 43.9 | 10.6 | ns |
| Digit Span[a] | 133 | 48.9 | 10.2 | 270 | 49.4 | 10.4 | ns |
| Digit Symbol[b] | 131 | 47.1 | 10.5 | 267 | 47.4 | 10.2 | ns |
| Stroop[c] | 106 | 48.5 | 38.4 | 238 | 46.0 | 38.0 | ns |
| CVLT LDCR[b] | 133 | -1.1 | 1.6 | 270 | -1.1 | 1.6 | ns |
| CVLT SEM[b] | 133 | -.41 | 1.2 | 270 | -.52 | 1.1 | ns |
| CVLT A1[b] | 133 | -.74 | 1.0 | 270 | -.92 | 1.0 | ns |
| CVLT A5[b] | 133 | -1.4 | 1.9 | 270 | -1.2 | 1.8 | ns |
| SDMT[a] | 120 | 41.6 | 12.0 | 264 | 41.0 | 12.6 | ns |
| Trail Part A[a] | 132 | 49.1 | 11.8 | 267 | 48.7 | 12.3 | ns |
| Trail Part B[a] | 132 | 48.2 | 11.3 | 268 | 48.1 | 12.3 | ns |
| WCST | 128 | 45.3 | 12.3 | 262 | 47.0 | 11.5 | ns |
| GPT DH[a] | 133 | 44.3 | 13.4 | 267 | 44.1 | 12.2 | ns |
| GPT NDH[a] | 130 | 43.4 | 11.9 | 267 | 43.4 | 11.6 | ns |
| FM – Memory | 116 | 52.8 | 8.6 | 250 | 53.2 | 5.5 | ns |
| FM – Learning | 116 | 38.8 | 8.6 | 251 | 38.3 | 9.1 | ns |

*Note.* BNT = Boston Naming Test; COWA = Controlled Oral Word Association; Digit Span = WAIS-R Digit Span; Digit Symbol = WAIS-R Digit Symbol; CVLT = California Verbal Learning Test; LDCR = Long-Delay Cued Recall; SEM = Semantic Cluster Ratio; A1 = List A, Trial 1; A5 = List A, Trial 5; Trail = Trail Making Test GPT = Grooved Pegboard Test; FM = Figure Memory; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand; FM = Figure Memory Test.
[a] T score
[b] standard score
[c] percentile

Appendix B

Table B1.
*Percentage of Follow-up Subjects Assigned to Reliable Change Index Categories – HIGH CD4.*

| **Domain & Test Variable** | **Medication Status** | ***n*** | **Worse** | **No Change** | **Better** |
|---|---|---|---|---|---|
| *Language* | | | | | |
| BNT | Non-HAART | 31 | 32.3 | 32.3 | 35.5 |
| | HAART-any | 96 | 34.4 | 37.5 | 28.1 |
| COWA | Non-HAART | 30 | 6.7 | 90.0 | 3.3 |
| | HAART-any | 98 | 4.1 | 86.7 | 9.2 |
| *Attention* | | | | | |
| Digit Span | Non-HAART | 30 | 10.0 | 76.7 | 13.3 |
| | HAART-any | 97 | 16.5 | 72.2 | 11.3 |
| CVLT A1 | Non-HAART | 30 | 13.3 | 86.7 | 0.0 |
| | HAART-any | 97 | 11.3 | 85.6 | 3.1 |
| Stroop | Non-HAART | 23 | 26.1 | 69.6 | 4.3 |
| | HAART-any | 81 | 9.9 | 82.7 | 7.4 |
| *Learning* | | | | | |
| CVLT A5 | Non-HAART | 30 | 16.7 | 56.7 | 26.7 |
| | HAART-any | 97 | 35.1 | 41.2 | 23.7 |
| CVLT SEM | Non-HAART | 30 | 26.7 | 70.0 | 3.3 |
| | HAART-any | 97 | 23.7 | 74.2 | 2.1 |
| *Memory* | | | | | |
| CVLT LDCR | Non-HAART | 30 | 20.0 | 66.7 | 13.3 |
| | HAART-any | 97 | 28.9 | 60.8 | 10.3 |
| *Abstraction* | | | | | |
| Trails B | Non-HAART | 30 | 10.0 | 76.7 | 13.3 |
| | HAART-any | 98 | 13.3 | 73.5 | 13.3 |
| WCST | Non-HAART | 28 | 7.1 | 82.1 | 10.7 |
| | HAART-any | 96 | 8.3 | 85.4 | 6.3 |

*Cont'd*

Table B1 *Continued*

| Domain & Test Variable | Medication Status | *n* | Worse | No Change | Better |
|---|---|---|---|---|---|
| *Motor* | | | | | |
| GPT – DH | Non-HAART | 31 | 22.6 | 64.5 | 12.9 |
| | HAART-any | 98 | 12.2 | 71.4 | 16.3 |
| GPT – NDH | Non-HAART | 31 | 25.8 | 67.7 | 6.5 |
| | HAART-any | 98 | 16.3 | 70.4 | 13.3 |
| *Psychomotor* | | | | | |
| Trails A | Non-HAART | 30 | 16.7 | 70.0 | 13.3 |
| | HAART-any | 98 | 13.3 | 75.5 | 11.2 |
| SDMT | Non-HAART | 28 | 10.7 | 85.7 | 3.6 |
| | HAART-any | 97 | 8.2 | 87.6 | 4.1 |
| Digit Symbol* | Non-HAART | 30 | 6.7 | 86.7 | 6.7 |
| | HAART-any | 98 | 68.2 | 62.2 | 29.6 |

*Note.* BNT = Boston Naming Test; COWA = Controlled Oral Word Association; CVLT = California Verbal Learning Test; A1 = List A Trial 1; A5 = List A Trial 5; SEM = Semantic Cluster Ratio; LDCR = Long-Delay Cued Recall; GPT = Grooved Pegboard Test; SDMT = Symbol Digit Modalities Test; WCST = Wisconsin Card Sorting Test perseverative responses; DH = dominant hand; NDH = non-dominant hand.
* $p < .05$

# Vita Auctoris

Susan Hayman-Abello was born in 1968 in London, Ontario. She completed a B. Sc. (Honours) degree in Psychology at Queen's University in Kingston (1990) and a Diploma in Child Study (Assessment and Counselling) at the University of Toronto (1994). She completed a Master's degree in Clinical Neuropsychology at the University of Windsor (2000) and is currently completing her doctoral degree at the University of Windsor.