

# UCSF

## UC San Francisco Previously Published Works

### Title

Improved neurocognitive test performance in both arms of the SMART study: impact of practice effect

### Permalink

<https://escholarship.org/uc/item/8dz76504>

### Journal

Journal of NeuroVirology, 19(4)

### ISSN

1355-0284

### Authors

Grund, Birgit

Wright, Edwina J

Brew, Bruce J

et al.

### Publication Date

2013-08-01

### DOI

10.1007/s13365-013-0190-x

Peer reviewed



Published in final edited form as:

*J Neurovirol.* 2013 August ; 19(4): 383–392. doi:10.1007/s13365-013-0190-x.

## Improved neurocognitive test performance in both arms of the SMART study: impact of practice effect

**Birgit Grund,**

Coordinating Center for Biometric Research, University of Minnesota, 2221 University Avenue, Suite 200, Minneapolis, MN 55414, USA

**Edwina J. Wright,**

Department of Infectious Diseases, The Alfred Hospital and Monash University, Melbourne, Australia; The Burnet Institute, Melbourne, Australia

**Bruce J. Brew,**

Department of Neurology, St Vincent's Hospital, Sydney, Australia; Department of HIV Medicine and St Vincent's Centre for Applied Medical Research, St Vincent's Hospital and University of New South Wales, Sydney, Australia

**Richard W. Price,**

Department of Neurology, University of California, San Francisco, CA, USA

**Mollie P. Roediger,**

Coordinating Center for Biometric Research, University of Minnesota, 2221 University Avenue, Suite 200, Minneapolis, MN 55414, USA

**Margaret P. Bain,**

Department of Neurology, St Vincent's Hospital, Sydney, Australia

**Jennifer F. Hoy,**

Department of Infectious Diseases, The Alfred Hospital and Monash University, Melbourne, Australia

**Judith C. Shlay,**

Denver Public Health, Denver, CO, USA

**Michael J. Vjecha, and**

Institute for Clinical Research, Veterans Affairs Medical Center, Washington, DC, USA

**Kevin R. Robertson**

Department of Neurology, University of North Carolina School of Medicine, Chapel Hill, NC, USA

**for the INSIGHT SMART Study Group**

### Abstract

We evaluated factors associated with improvement in neurocognitive performance in 258 HIV-infected adults with baseline CD4 lymphocyte counts above 350 cells/mm<sup>3</sup> randomized to intermittent, CD4-guided antiretroviral therapy (ART) (128 participants) versus continuous therapy (130) in the Neurology substudy of the Strategies for Management of Antiretroviral Therapy trial. Participants were enrolled in Australia, North America, Brazil, and Thailand, and neurocognitive performance was assessed by a five-test battery at baseline and month 6. The

primary outcome was change in the quantitative neurocognitive performance z score (QNPZ-5), the average of the z scores of the five tests. Associations of the 6-month change in test scores with ART use, CD4 cell counts, HIV RNA levels, and other factors were determined using multiple regression models. At baseline, median age was 40 years, median CD4 cell count was 513 cells/mm<sup>3</sup>, 88 % had plasma HIV RNA > 400 copies/mL, and mean QNPZ-5 was -0.68. Neurocognitive performance improved in both treatment groups by 6 months; QNPZ-5 scores increased by 0.20 and 0.13 in the intermittent and continuous ART groups, respectively (both  $P < 0.001$  for increase and  $P = 0.26$  for difference). ART was used on average for 3.6 and 5.9 out of the 6 months in the intermittent and continuous ART groups, respectively, but the increase in neurocognitive test scores could not be explained by ART use, changes in CD4, or plasma HIV RNA, which suggests a practice effect. The impact of a practice effect after 6 months emphasizes the need for a control group in HIV studies that measure intervention effects using neurocognitive tests similar to ours.

## Keywords

HIV; Test-retest; Practice effect; Learning effect; Neurocognitive performance; Neuropsychological tests

## Introduction

Untreated HIV infection may impair neurocognitive function (McArthur et al. 2010; Price et al. 1988). With the introduction of combination antiretroviral therapy (cART), the incidence of HIV-associated dementia declined sharply (Bhaskaran et al. 2008; d'Arminio Monforte et al. 2004; Dore et al. 1999; Lescure et al. 2011; McArthur 2004; McArthur et al. 2010; Sacktor et al. 2001). Nevertheless, recent studies suggest that the prevalence of neurocognitive disorders remains substantial even among individuals with high CD4 lymphocyte cell counts and suppressed viral load (Cysique and Brew 2011; Heaton et al. 2010; McArthur et al. 2010) and that up to 52 % of HIV-infected individuals exhibit some form of neurocognitive impairment (NCI). The reasons for that are uncertain but may include confounding factors, a legacy effect of irreversible injury before therapy was initiated, ongoing indolent injury, and perhaps neurotoxic effects of antiretroviral therapy (ART) (Al-Khindi et al. 2011; Heaton et al. 2011; Robertson et al. 2010). While many longitudinal studies have reported improvements in neuropsychological test scores after starting cART (Al-Khindi et al. 2011), few HIV studies have investigated the test-retest practice effect as a possible explanation.

The Strategies for Management of Antiretroviral Therapy (SMART) study randomized HIV-infected adults with CD4 cell counts above 350 cells/mm<sup>3</sup> to intermittent, CD4-guided ART or continuous ART (El-Sadr et al. 2006). In a Neurology substudy, a neurocognitive test battery was administered (Wright et al. 2010), in order to compare intermittent to continuous ART for changes in neurocognitive performance. We hypothesized that neurocognitive performance would be superior in participants receiving continuous ART because of the sustained reduction in HIV replication. The intermittent ART arm in the parent SMART study was stopped early, however, due to a higher incidence of progression of disease and death, and continuous ART was recommended to all participants; at that time, only 292 of the planned 600 Neurology substudy participants were enrolled. Consequently, the substudy was underpowered for the planned primary comparison. Surprisingly, though, neurocognitive performance had increased significantly in *both* study arms. The main goal of the current analysis was to examine the basis of this finding, including whether the improvement in neurocognitive test scores could be explained as a benefit of ART.

## Methods

### Study design

The SMART study was an international randomized trial comparing continuous ART with CD4 cell count-guided, intermittent ART in HIV-infected adults with CD4 cell counts above 350 cells/mm<sup>3</sup>. The study design is described elsewhere (El-Sadr et al. 2006). Participants in the intermittent ART arm stopped or deferred ART until the CD4 cell count declined to below 250 cells/mm<sup>3</sup> and, thereafter, used ART with the goal of viral suppression until the CD4 count had recovered to above 350 cells/mm<sup>3</sup>. The primary outcome of the neurocognitive component of the SMART Neurology substudy was a quantitative neurocognitive performance *z* score (QNPZ-5), based on a five-test neurocognitive test battery. The study was designed to enroll 600 participants, follow them up for 5 years, and administer the test battery at baseline, month 6, and annually thereafter. The Neurology substudy commenced enrollment in July 2005. In January 2006, enrollment and the intermittent CD4-guided ART strategy of the SMART study were stopped due to increased risk of AIDS and serious non-AIDS complications, and participants were recommended to restart ART (El-Sadr et al. 2006). At the time of this protocol change, only 292 of the 600 planned substudy participants were enrolled. Data collection in this substudy continued until July 2006, when all participants had been followed up for 6 months.

Participants were enrolled at 47 sites in Australia, North America, Brazil, and Thailand. At these sites, all eligible SMART participants were offered substudy co-enrollment. Substudy eligibility criteria included age ≥ 18 years and the ability to perform the study's neurocognitive tests in the site clinician's judgment.

### Standard protocol approvals, registrations, and patient consents

Substudy approval was obtained by each site's institutional review board; the study is in accordance with the Helsinki Declaration of 1975, as revised in 2000. All participants provided written informed consent. [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT00432003) identifier: NCT00432003.

### Outcome measures

The five-test neurocognitive test battery comprised the Grooved Pegboard (GPB) (dominant hand) (Klove 1963), Color Trails 1 and 2 (CT1 and CT2) (D'Elia et al. 1996), Timed Gait (TG) (Robertson et al. 2006), and Finger Tapping (FT) (nondominant hand) (Reitan 1969) tests, described previously (Wright et al. 2010). The tests closely resemble the battery used in a large, randomized HIV treatment study (Price et al. 1999).

For the individual tests, the raw scores of each participant were standardized to *z* scores (Wright et al. 2010). A negative *z* score denotes below-average performance relative to an HIV-negative US reference population, and a *z* score of -1 corresponds to performance at 1 standard deviation below the average of the reference population. Reference distributions were provided by the licensors of the tests (GPB, CT, and FT) or obtained from published studies (TG), matched by education level for all five tests, and, where available, additionally by age (GPB, CT, and FT), sex, and race/ethnicity (GPB and FT) (Heaton et al. 2004; Morey 2006; Reitan 1969; Robertson et al. 2006).

The primary neurocognitive performance outcome measure, the QNPZ-5, was calculated as the average of the five *z* scores from the five individual tests in the battery. To further describe the study population, we defined NCI as *z* scores < -2 in at least two ability domains. Ability domains assessed were (1) speed/fine motor skills (GPB and FT), (2) attention/speed of processing (CT1), (3) executive function (CT2), and (4) gross motor skills (TG). The cutoff was chosen to reflect the Frascati criteria for abnormal neurocognitive

performance, although the test battery was too small to meet the formal Frascati criteria; also, we did not assess whether individual low test scores might be explained by comorbidities (Antinori et al. 2007).

Neuropsychological tests were administered to participants by site staff, including medical practitioners, specialists, and nurses. Site staff underwent centralized training and certification in the USA and Australia, as previously described (Wright et al. 2010). Training was led by a neuropsychologist or neurologist and an infectious disease physician.

Demography, HIV history, general medical history, ART use, and laboratory values were collected within the parent SMART study. Adherence to ART was collected by patient self-report, using a 7-day recall questionnaire at baseline and month 4; participants who took at least 80 % of their medication were considered adherent for the current analysis. The central nervous system (CNS) penetration effectiveness (CPE) rank of patients' ART regimens was calculated using the 2010 revision of the CPE score (Letendre et al. 2010).

### Statistical methods

Mean baseline neuropsychological test scores (the QNPZ-5 and individual test  $z$  scores) were compared to 0 using two-sided  $t$  tests; values below 0 denote below-average test performance compared to an HIV-negative reference population. Treatment groups were compared for month 6 changes in test scores by intent-to-treat, using unadjusted  $t$  tests, as well as using ANCOVA adjusted for change in CD4 cell counts and HIV RNA levels. Mean changes in test scores, overall and in various subgroups, were compared to 0 using  $t$  tests. Homogeneity across the locations of enrollment was assessed using analysis of variance (ANOVA; baseline  $z$  scores, changes in  $z$  scores). For six participants, one out of the five individual test scores was missing, and change in QNPZ-5 was calculated as the average of change in the individual  $z$  scores for the four available tests.

In order to determine whether the increases in QNPZ-5 and individual test  $z$  scores could be explained by changes in CD4 cell counts, plasma HIV RNA, or ART use, the associations of these factors with changes in test scores were calculated using univariate and multiple linear regression and Spearman's correlation coefficient. Subgroups by ART status at baseline and month 6 were compared for changes in QNPZ-5 by ANOVA and using  $t$  tests for contrasts with and without Scheffé's adjustment for multiple comparisons (to adjust for post hoc analyses).

In multiple linear regression models for change in neurocognitive test  $z$  scores, we always included demographics (age, years of education, sex, and location of enrollment) to account for imperfect standardization of test scores using US reference populations, as well as the baseline test scores, because of a possible association of baseline with change in test scores due to regression to the mean. In a first model, we also included change in CD4 cell counts and  $\log_{10}$  HIV RNA because the beneficial effects of ART would likely be mediated through these factors. We did not include race/ethnicity because it was strongly confounded with the location of enrollment. Coefficients were presented when the  $P$  value for association was at least 0.10. The analysis was repeated within each treatment group. In further models, all adjusted for demographics and baseline test scores, we investigated associations with single baseline factors, for example, duration and CPE rank of ART.

Analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC, USA) and R version 2.10 (R Development Core Team 2011). All tests were two-sided.  $P$  values  $\leq 0.05$  were considered significant.

## Results

### Study population and baseline characteristics

The analysis includes all 258 participants in the Neurology substudy with both baseline and month 6 neurocognitive test data (Australia/USA, 98; Brazil, 17; Thailand, 143). Of these, 128 were randomized to intermittent ART and 130 to continuous ART (Fig. 1). Baseline characteristics and neurocognitive test scores are summarized in Tables 1 and 2; the treatment groups were well balanced for baseline characteristics. Median CD4 cell count was 513 cells/mm<sup>3</sup>, median nadir CD4 count was 216 cells/mm<sup>3</sup>, and 88 % had plasma HIV RNA 400 copies/mL. Only 4 % of participants were ART-naïve; 93 % were using ART at baseline, with a median of 4 years of prior ART use. ART regimens had a median CPE rank of 7, on a scale of 1–4 for each drug (Table 1).

Overall, the mean QNPZ-5 score was  $-0.68$  ( $P < 0.001$  for comparison to 0), denoting below-average performance compared to a healthy reference population. For four of the five individual tests, the average  $z$  score was negative ( $P < 0.05$  for each; Table 2). Fourteen percent of participants met the study definition of NCI ( $z$  scores  $< -2$  in two or more ability domains); 52 % had  $z$  scores  $< -1$  in two or more cognitive domains.

Mean QNPZ-5 scores differed by location of enrollment ( $-0.77$  in Australia/USA,  $-0.15$  in Brazil, and  $-0.68$  in Thailand;  $P = 0.01$  for comparison between locations). No one region performed consistently higher or lower across all tests. Baseline characteristics and test scores by location of enrollment are described elsewhere (Wright et al. 2010).

### Change in neurocognitive performance

Changes in the QNPZ-5 and in individual test  $z$  scores from baseline to month 6 are summarized in Table 2 and, by treatment group, in Fig. 2a. QNPZ-5 scores increased by 0.20 in the intermittent ART group (95 % CI, 0.12 to 0.29;  $P < 0.001$  for increase) and by 0.13 in the continuous ART group (95 % CI, 0.03 to 0.23;  $P = 0.009$ ). The estimated mean change in  $z$  scores was positive for each of the five tests in both treatment groups, and the increase was significant for GPB, CT1, and CT2 ( $P < 0.001$  each; Table 2 and Fig. 2). The increases in QNPZ-5 and  $z$  scores were homogeneous across locations (all  $P > 0.20$ ).

There was no evidence for a difference between the intermittent and continuous ART groups in either mean change in QNPZ-5 (estimated difference, 0.07; 95 % CI,  $-0.06$  to 0.20;  $P = 0.26$ ) or in any of the individual test  $z$  scores (all  $P > 0.10$ ). Moreover, the direction of the treatment difference was inconsistent across the tests (Fig. 2a).

### ART use

Participants in the intermittent ART group stopped or deferred ART at randomization, but 108 (86 %) had restarted ART by 6 months. Of those, 86 (80 %) restarted ART after January 11, 2006, when participants were recommended to resume ART. ART use is summarized in Fig. 1. Participants in the intermittent ART group used ART for an average of 3.6 of the 6 months, compared with 5.9 months in the continuous ART group. Adherence was excellent; among those who were receiving ART at month 4, 100 % in the intermittent ART arm and 98 % in the continuous ART arm self-reported adhering to their regimen; at month 6, adherence was not assessed, but among those receiving ART, 90 and 93 %, respectively, had HIV RNA 400 copies/mL.

### Factors associated with change in neurocognitive performance

In order to investigate whether initiation of ART could explain the increase in QNPZ-5 scores, we evaluated the mean change in QNPZ-5 scores within subgroups by ART status at

baseline and at month 6. The four groups are described in Fig. 1, and mean change in QNPZ-5 scores is shown in Fig. 2b. Differences between the four groups were not statistically significant by ANOVA ( $P=0.07$ ). However, when comparing participants according to their ART status at month 6, those who continued or started ART (groups 1 and 3 pooled in Fig. 2b) improved less than those who stopped or never started ART (groups 2 and 4 pooled) (estimated mean difference,  $-0.36$ ;  $P=0.01$  unadjusted,  $P=0.11$  after Scheffé's adjustment for multiple comparisons). The duration of ART use during the 6 months was not associated with change in QNPZ-5 (Spearman's  $r=-0.07$ ,  $P=0.25$ ). Similarly, the change in the CPE rank of ART was not associated with change in QNPZ-5 among the 224 participants who were using ART at both baseline and month 6 (mean change in CPE rank,  $-0.1$ ; Spearman's  $r=0.08$ ,  $P=0.22$ ); mean CPE ranks in each ART group are shown in Fig. 2b.

Table 3 summarizes the associations of baseline factors, change in CD4 cell counts, and change in plasma HIV RNA with changes in QNPZ-5 and individual test  $z$  scores, estimated in multiple regression models. Lower baseline QNPZ-5 was associated with steeper increases in QNPZ-5 from baseline to month 6 (0.23 units higher per 1 unit baseline QNPZ-5 lower); similar associations held for each individual test  $z$  score (all  $P<0.001$ ; Table 3). Mean increases in QNPZ-5 were by 0.14 lower for women compared with men.

Overall changes in plasma HIV RNA levels were small, as most participants were using ART at baseline and continued or had restarted ART by month 6; mean  $\log_{10}$  HIV RNA levels increased by 0.1  $\log_{10}$  copies/mL in the intermittent ART group and decreased by 0.1  $\log_{10}$  copies/mL in the continuous ART group. A tenfold increase in HIV RNA (1  $\log_{10}$  copies/mL) was independently associated with a 0.15 higher than average increase in QNPZ-5 (Table 3); this association was similar in both treatment groups (data not shown).

By month 6, mean CD4 cell counts had decreased by 21 cells/mm<sup>3</sup> in the intermittent ART group and increased by 33 cells/mm<sup>3</sup> in the continuous ART group. In the intermittent ART group, a 6-month increase in CD4 cell count was associated with a slightly steeper increase in QNPZ-5, by 0.03 per 50 cells/mm<sup>3</sup> CD4 increase ( $P=0.05$ ). There was no evidence for an association within the continuous ART group or in the pooled analysis of both treatment groups (Table 3).

After adjustment for change in CD4 cell counts and  $\log_{10}$  HIV RNA, the estimated mean increase in QNPZ-5 scores remained significant within each group, almost unchanged at 0.20 for the intermittent ART group, and 0.15 for the continuous ART group (Fig. 2a).

There was no evidence that experiencing a grade 4 or AIDS event was associated with a change in QNPZ-5; seven participants had experienced such an event by month 6. We also investigated the associations of change in QNPZ-5 in a multiple regression model that contained only baseline factors. Coefficients and  $P$  values for baseline QNPZ-5, age, sex, and location of enrollment were similar to those shown in Table 3. Higher baseline CD4 was independently associated with less increase in QNPZ-5 ( $-0.04$  less than the average increase per 100 cells/mm<sup>3</sup> higher;  $P=0.02$ ), but there was no independent association with ART use at study entry or the duration of prior ART use.

## Discussion

In this prospective study of 258 participants from Australia, North America, Brazil, and Thailand with CD4 cell counts  $>350$  cells/mm<sup>3</sup>, we found that neurocognitive performance, measured by mean change in QNPZ-5 scores, increased by month 6 among participants receiving intermittent ART as well as among those receiving continuous ART. The increase

in QNPZ-5 scores in *both* treatment groups suggests a practice effect. The presence of a practice effect emphasizes the importance of including a control group in studies that use test batteries like ours to evaluate the effect of an intervention.

We have investigated and excluded ART use as a possible explanation for the increase in QNPZ-5 in both arms:

1. The increase in QNPZ-5 scores could not be explained as a beneficial effect of ART use mediated through the suppression of plasma HIV RNA or increase in CD4 cell counts. On the contrary, an *increase* in viral load was associated with higher than average improvements in QNPZ-5 scores (Table 3), and change in CD4 was not independently associated with change in QNPZ-5.
2. In our cohort, there is no indication that starting or continuing ART contributed to the increase in QNPZ-5 in other ways. In fact, among the 16 participants who were not using ART at baseline and started ART, the average QNPZ-5 declined, compared with increases for all other participants (Fig. 2b). Duration of ART use during the 6-month follow-up in the study was not associated with changes in QNPZ-5 (Spearman's  $r=-0.07$ ,  $P=0.25$ ), nor was prior ART use, or changes in the CPE rank of ART among those who were using ART at study entry and remained on ART throughout follow-up ( $r=0.08$ ,  $P=0.22$ ). Moreover, increases in mean QNPZ-5 scores were largest for participants who never started ART or had stopped ART by month 6 (groups 2 and 4 in Fig. 2b).

The presence of a practice effect when neuropsychological tests are repeated after 6–12 months is well documented in the general population (Heilbronner et al. 2010; McCaffrey et al. 2000). The American Academy of Clinical Neuropsychology advises that “the primary element requiring careful consideration in serial testing is the practice effect ... A change in scores on repeat testing can lead to ... attributing improved scores to recovery of function or efficacy of intervention when the results are more likely to reflect the effects of test practice” (Heilbronner et al. 2010). Nevertheless, in HIV research, the practice effect has often been neglected when evaluating interventions. For example, in a recent meta-analysis of 23 longitudinal studies investigating the effect of ART on neurocognitive function (Al-Khindi et al. 2011), 11 assessed the effect of ART by comparing test scores before and after cART initiation. All 11 studies reported improvement in test scores after 3–6 months (Brouwers et al. 1997; Carvalhal et al. 2006; Chang et al. 2003; Robertson et al. 2004; Sacktor et al. 2006, 2009; Tozzi et al. 1993, 1999, 2001, 2009; Winston et al. 2010). Of the 11 studies, only 6 even mentioned the possibility of a practice effect (Brouwers et al. 1997; Chang et al. 2003; Robertson et al. 2004; Sacktor et al. 2006, 2009; Tozzi et al. 1993). Only two of the studies formally adjusted for practice effect; in both studies, little evidence remained for a benefit of ART on neurocognitive performance after the adjustment (Chang et al. 2003; Sacktor et al. 2009).

In our cohort, increases in viral load were associated with higher-than-average increases in QNPZ-5 scores. This finding is contrary to the expectation that suppressing viral load is protecting the brain and, therefore, should lead to an improvement in neurocognitive performance (Marcotte et al. 2003; Wright 2011). It is contrary also to a small, single-arm prospective study of treatment interruption ( $n=11$ ), where the global deficit score remained unchanged from baseline during the treatment interruption, but improved after restarting ART (Childers et al. 2008). Our finding is consistent, however, with results in a single-arm treatment interruption study of 167 participants with CD4 cell counts  $>350$  cells/mm<sup>3</sup> (Robertson et al. 2010), where mean neurocognitive performance improved after stopping ART for up to 1.5 years, but remained stable in the same patients following ART recommencement. The initial improvement after stopping ART in the single-arm study may



have been due to a practice effect, but the absence of further improvement after starting ART does not support the hypothesis of a beneficial effect in this population; Robertson et al. (2010) raised the possibility that ART might be neurotoxic. Nevertheless, suppressing HIV RNA may benefit neurocognitive function in individuals with lower CD4 levels; our study participants spent 91 % of their total follow-up time at CD4 cell counts above 350 cells/mm<sup>3</sup>.

The single strongest predictor for change in QNPZ-5 as well as for each individual test was the baseline test score; participants with lower baseline scores improved more (Table 3). The same was observed in a previous study of 124 HIV-infected individuals and 172 HIV-negative controls (Cysique et al. 2011). Such “regression to the mean” effect is common when the response is change in an outcome that is also included as baseline factor. Women had, on average, lower increases in QNPZ-5 scores and GPB and CT1 *z* scores (Table 3).

Limitations of our study include that we could not compare the intermittent versus continuous ART strategies as originally planned because the intermittent ART arm in the parent SMART study was stopped by the DSMB and, consequently, most substudy participants had switched to continuous ART before the first follow-up assessment at month 6. Also, the study was not powered to investigate the effect of individual drugs or drug classes. Lastly, neurocognitive function was assessed with only five tests, and *z* scores were calculated using US norms; therefore, our estimation of the prevalence of NCI needs to be interpreted with caution. Our main results refer to *changes* in *z* scores, however, and thus are not affected by the use of US versus local norms.

In our study of HIV-infected individuals with high CD4 cell counts, neurocognitive test results had improved in both study arms by 6 months. The improvement could not be explained by ART use or changes in CD4 or plasma HIV RNA levels. This suggests the presence of a practice effect, consistent with findings from other studies (Cysique et al. 2011; Levine et al. 2004; McCaffrey et al. 2000). Our finding emphasizes the need for a control group in studies that measure an intervention effect using this type of performance test battery. Where a control group is not possible for ethical reasons, adjustments for a potential practice effect are needed.

## Acknowledgments

The study was funded by NIH grants U01-AI4636 (National Institute of Mental Health [NIMH] and the National Institute of Neurological Disorders and Stroke [NINDS]), U01AI042170, and U01AI46362 (National Institute of Allergy and Infectious Diseases [NIAID]). We gratefully acknowledge the contributions of participants, investigators, and staff who were crucial to the success of the study. Clinical sites and investigators participating in the SMART Neurology substudy were published (Wright et al. 2010).

The following authors received research support from the NIH: BG and MPR from NIH/NIAID (U01-AI068641, U01-AI042170, and U01-AI46362); EJW (NIMH/NINDS 1U01-AI068641); BJB (R01 NS43103); RWP (NIDA P01DA026134, NIMH R01-MH62701, NIMH R21-MH083520, NIMH U01-MH083545, NIMH R01MH081772, NIMH/NIAIDU01-AI068641, and NIMH/NIAID U01-AI38858); and KRR (NIAID, supplement to 1U01AI068636-01, NIMH MH067751, NIAID U01-AI068641, and NINDS R21NS0692).

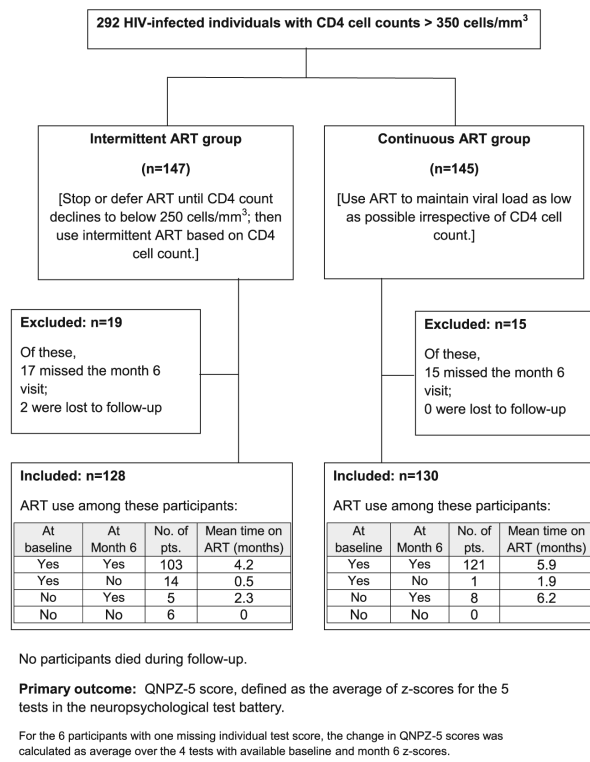
## References

- Al-Khindi T, Zakzanis KK, van Gorp WG. Does antiretroviral therapy improve HIV-associated cognitive impairment? A quantitative review of the literature. *J Int Neuropsychol Soc.* 2011; 17:956–969. [PubMed: 21813033]
- Antinori A, Arendt G, Becker JT, Brew BJ, Byrd DA, Cherner M, Clifford DB, Cinque P, Epstein LG, Goodkin K, Gisslen M, Grant I, Heaton RK, Joseph J, Marder K, Marra CM, McArthur JC, Nunn M, Price RW, Pulliam L, Robertson KR, Sacktor N, Valcour V, Wojna VE. Updated research

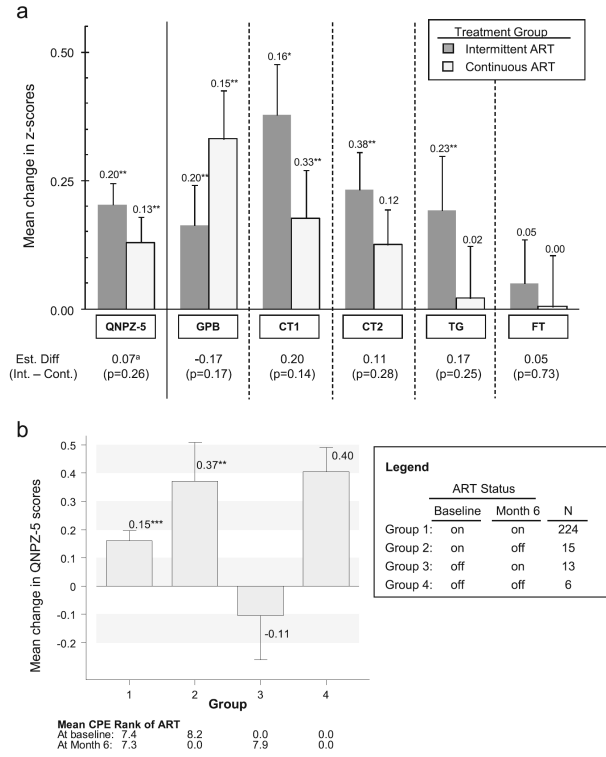
- nosology for HIV-associated neurocognitive disorders. *Neurology*. 2007; 69:1789–1799. [PubMed: 17914061]
- Bhaskaran K, Mussini C, Antinori A, Walker AS, Dorrucchi M, Sabin C, Phillips A, Porter K, Collaboration C. Changes in the incidence and predictors of human immunodeficiency virus-associated dementia in the era of highly active antiretroviral therapy. *Ann Neurol*. 2008; 63:213–221.
- Brouwers P, Hendricks M, Lietzau JA, Pluda JM, Mitsuya H, Broder S, Yarchoan R. Effect of combination therapy with zidovudine and didanosine on neuropsychological functioning in patients with symptomatic HIV disease: a comparison of simultaneous and alternating regimens. *AIDS*. 1997; 11:59–66. [PubMed: 9110076]
- Carvalho AS, Rourke SB, Belmonte-Abreu P, Correa J, Goldani LZ. Evaluation of neuropsychological performance of HIV-infected patients with minor motor cognitive dysfunction treated with highly active antiretroviral therapy. *Infection*. 2006; 34:357–360. [PubMed: 17180595]
- Chang L, Ernst T, Witt MD, Ames N, Walot I, Jovicich J, DeSilva M, Trivedi N, Speck O, Miller EN. Persistent brain abnormalities in antiretroviral-naive HIV patients 3 months after HAART. *Antivir Ther*. 2003; 8:17–26. [PubMed: 12713060]
- Childers ME, Woods SP, Letendre S, McCutchan JA, Rosario D, Grant I, Mindt MR, Ellis RJ, San Diego HIV Neurobehavioral Research Center Group. Cognitive functioning during highly active antiretroviral therapy interruption in human immunodeficiency virus type 1 infection. *J Neurovirol*. 2008; 14:550–557. [PubMed: 19016380]
- Cysique LA, Franklin D, Abramson I, Ellis RJ, Letendre S, Collier A, Clifford D, Gelman B, McArthur J, Morgello S, Simpson D, McCutchan JA, Grant I, Heaton RK, Grp C, Grp H. Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. *J Clin Exp Neuropsychol*. 2011; 33:505–522. [PubMed: 21391011]
- Cysique LA, Brew BJ. Prevalence of non-confounded HIV-associated neurocognitive impairment in the context of plasma HIV RNA suppression. *J Neurovirol*. 2011; 17:176–183. [PubMed: 21416169]
- d'Arminio Monforte A, Cinque P, Mocroft A, Goebel FD, Antunes F, Katlama C, Justesen US, Vella S, Kirk O, Lundgren J, EuroSIDA Study Group. Changing incidence of central nervous system diseases in the EuroSIDA cohort. *Ann Neurol*. 2004; 55:320–328. [PubMed: 14991809]
- D'Elia, LF.; Satz, P.; Uchiyama, CL.; White, T. Professional manual. Psychological Assessment Resources; Odessa: 1996. Color Trail test.
- Dore GJ, Correll PK, Li Y, Kaldor JM, Cooper DA, Brew BJ. Changes to AIDS dementia complex in the era of highly active antiretroviral therapy. *AIDS*. 1999; 13:1249–1253. [PubMed: 10416530]
- El-Sadr W, Lundgren J, Neaton J, Gordin F, Abrams D, Arduino R, Babiker A, Burman W, Clumeck N, Cohen C, Cohn D, Cooper D, Darbyshire J, Emery S, Fatkenheuer G, Gazzard B, Grund B, Hoy J, Klingman K, Losso M, Mejia J, Markowitz N, Neuhaus J, Phillips A, Rappoport C, Antiretrovir SM. CD4+ count-guided interruption of antiretroviral treatment. *N Eng J Med*. 2006; 355:2283–2296.
- Heaton RK, Clifford DB, Franklin DR Jr, Woods SP, Ake C, Vaida F, Ellis RJ, Letendre SL, Marcotte TD, Atkinson JH, Rivera-Mindt M, Vigil OR, Taylor MJ, Collier AC, Marra CM, Gelman BB, McArthur JC, Morgello S, Simpson DM, McCutchan JA, Abramson I, Gamst A, Fennema-Notestine C, Jernigan TL, Wong J, Grant I, CHARTER Group. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. *Neurology*. 2010; 75:2087–2096. [PubMed: 21135382]
- Heaton RK, Franklin DR, Ellis RJ, McCutchan JA, Letendre SL, LeBlanc S, Corkran SH, Duarte NA, Clifford DB, Woods SP, Collier AC, Marra CM, Morgello S, Mindt MR, Taylor MJ, Marcotte TD, Atkinson JH, Wolfson T, Gelman BB, McArthur JC, Simpson DM, Abramson I, Gamst A, Fennema-Notestine C, Jernigan TL, Wong J, Grant I, CHARTER Group; HNRC Group. HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: differences in rates, nature, and predictors. *J Neurovirol*. 2011; 17:3–16. [PubMed: 21174240]
- Heaton, RK.; Miller, SW.; Taylor, JJ.; Grant, I. Revised comprehensive norms for an expanded Halstead-Reitan battery: demographically adjusted neuropsychological norms for African Americans and Caucasian adults. PAR Psychological Assessment Resources, Inc.; Lutz: 2004.

- Heilbronner RL, Sweet JJ, Attix DK, Krull KR, Henry GK, Hart RP. Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges of repeat test administrations in clinical and forensic contexts. *Clin Neuropsychol*. 2010; 24:1267–1278. [PubMed: 21108148]
- Klove, H. *Clinical neuropsychology*. Saunders; New York: 1963.
- Lescure FX, Omland LH, Engsig FN, Roed C, Gerstoft J, Pialoux G, Kronborg G, Larsen CS, Obel N. Incidence and impact on mortality of severe neurocognitive disorders in persons with and without HIV infection: a Danish nationwide cohort study. *Clin Infect Dis*. 2011; 52:235–243. [PubMed: 21288850]
- Letendre, SL.; FitzSimons, C.; Ellis, RJ.; Clifford, D.; Collier, AC.; Gelman, B.; McArthur, J.; Vaida, F.; Heaton, R.; Grant, I.; the CHARTER Group. Correlates of CSF viral loads in 1221 volunteers of the CHARTER cohort. *Proceedings of the 17th Conference on Retrovirus and Opportunistic Infections*; San Francisco, CA. 2010.
- Levine AJ, Miller EN, Becker JT, Selnes OA, Cohen BA. Normative data for determining significance of test–retest differences on eight common neuropsychological instruments. *Clin Neuropsychol*. 2004; 18:373–384. [PubMed: 15739809]
- Marcotte TD, Deutsch R, McCutchan JA, Moore DJ, Letendre S, Ellis RJ, Wallace MR, Heaton RK, Grant I. Prediction of incident neurocognitive impairment by plasma HIV RNA and CD4 levels early after HIV seroconversion. *Arch Neurol*. 2003; 60:1406–1412. [PubMed: 14568811]
- McArthur JC. HIV dementia: an evolving disease. *J Neuroimmunol*. 2004; 157:3–10. [PubMed: 15579274]
- McArthur JC, Steiner J, Sacktor N, Nath A. Human immunodeficiency virus-associated neurocognitive disorders: mind the gap. *Ann Neurol*. 2010; 67:699–714. [PubMed: 20517932]
- McCaffrey, RJ.; Duff, K.; Westervelt, HJ. *Practitioner’s guide to evaluating change with neuropsychological assessment instruments*. 1st edn. Springer; New York: 2000.
- Morey, LC. PAI® software portfolio. PAR Psychological Assessment Resources, Inc.; Lutz: 2006.
- Price RW, Yiannoutsos CT, Clifford DB, et al. Neurological outcomes in late HIV infection: adverse impact of neurological impairment on survival and protective effect of antiviral therapy. *AIDS*. 1999; 13:1677–1685. [PubMed: 10509569]
- Price RW, Brew B, Sidtis J, Rosenblum M, Scheck AC, Cleary P. The brain in AIDS: central nervous system HIV-1 infection and AIDS dementia complex. *Science*. 1988; 239:586–592. [PubMed: 3277272]
- R Development Core Team. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing; Vienna: 2011.
- Reitan, RM. *Manual for administration of neuropsychological test batteries for adults and children*. Reitan; Indianapolis: 1969.
- Robertson KR, Parsons TD, Sidtis JJ, Inman TH, Robertson WT, Hall CD, Price RW. Timed Gait test: normative data for the assessment of the AIDS dementia complex. *J Clin Exp Neuropsychol*. 2006; 28:1053–1064. [PubMed: 16840235]
- Robertson KR, Robertson WT, Ford S, Watson D, Fiscus S, Harp AG, Hall CD. Highly active antiretroviral therapy improves neurocognitive functioning. *J Acq Imm Def Synd: JAIDS*. 2004; 36:562–566.
- Robertson KR, Su Z, Margolis DM, Krambrink A, Havlir DV, Evans S, Skiest DJ, Team AS. Neurocognitive effects of treatment interruption in stable HIV-positive patients in an observational cohort. *Neurology*. 2010; 74:1260–1266. [PubMed: 20237308]
- Sacktor N, Lyles RH, Skolasky R, Kleeberger C, Selnes OA, Miller EN, Becker JT, Cohen B, McArthur JC, Study MAC. HIV-associated neurologic disease incidence changes: Multicenter AIDS Cohort Study, 1990–1998. *Neurology*. 2001; 56:257–260. [PubMed: 11160967]
- Sacktor N, Nakasujja N, Skolasky R, Robertson K, Wong M, Musisi S, Ronald A, Katabira E. Antiretroviral therapy improves cognitive impairment in HIV+ individuals in sub-Saharan Africa. *Neurology*. 2006; 67:311–314. [PubMed: 16864825]
- Sacktor N, Nakasujja N, Skolasky RL, Robertson K, Musisi S, Ronald A, Katabira E, Clifford DB. Benefits and risks of stavudine therapy for HIV-associated neurologic complications in Uganda. *Neurology*. 2009; 72:165–170. [PubMed: 19139369]

- Tozzi V, Balestra P, Galgani S, Narciso P, Ferri F, Sebastiani G, D'Amato C, Affricano C, Pigorini F, Pau FM, De Felici A, Benedetto A. Positive and sustained effects of highly active antiretroviral therapy on HIV-1-associated neurocognitive impairment. *AIDS*. 1999; 13:1889–1897. [PubMed: 10513647]
- Tozzi V, Balestra P, Galgani S, Narciso P, Sampaolesi A, Antinori A, Giulianelli M, Serraino D, Ippolito G. Changes in neurocognitive performance in a cohort of patients treated with HAART for 3 years. *J Acq Imm Def Synd: JAIDS*. 2001; 28:19–27.
- Tozzi V, Balestra P, Salvatori MF, Vlassi C, Liuzzi G, Giancola ML, Giulianelli M, Narciso P, Antinori A. Changes in cognition during antiretroviral therapy: comparison of 2 different ranking systems to measure antiretroviral drug efficacy on HIV-associated neurocognitive disorders. *J Acq Imm Def Synd: JAIDS*. 2009; 52:56–63. *J Acquir Immune Defic Syndr*. Dec 1.2009 52(4): 529. Erratum appears in.
- Tozzi V, Narciso P, Galgani S, Sette P, Balestra P, Gerace C, Pau FM, Pigorini F, Volpini V, Camporiondo MP, et al. Effects of zidovudine in 30 patients with mild to end-stage AIDS dementia complex. *AIDS*. 1993; 7:683–692. [PubMed: 8318176]
- Winston A, Duncombe C, Li PCK, Gill JM, Kerr SJ, Puls R, Petoumenos K, Taylor-Robinson SD, Emery S, Cooper DA, Altair Study Group. Does choice of combination antiretroviral therapy (cART) alter changes in cerebral function testing after 48 weeks in treatment-naive, HIV-1-infected individuals commencing cART? A randomized, controlled study. *Clin Infect Dis*. 2010; 50:920–929. [PubMed: 20146627] *Clin Infect Dis*. Sep 1.2010 51(5):638. Erratum appears in.
- Wright E. Neurocognitive impairment and neuroCART. *Curr Opin HIVAIDS*. 2011; 6:303–308.
- Wright E, Grund B, Robertson K, Brew B, Roediger M, Bain M, Drummond F, Vjecha M, Hoy J, Miller C, de Oliveira A, Pumpradit W, Shlay J, El-Sadr W, Price R, Grp ISS. Cardiovascular risk factors associated with lower baseline cognitive performance in HIV-positive persons. *Neurology*. 2010; 75:864–873. [PubMed: 20702792]



**Fig. 1.** Study design, CONSORT flow diagram, and use of antiretroviral treatment (ART)



**Fig. 2.** **a** Change in QNPZ-5 and z scores from baseline to month 6, with 1 SE indicated by error bars. **b** Change in QNPZ-5 scores within subgroups by ART status, and mean CPE rank of ART at baseline and month 6. \* $P < 0.05$ , increase in z score is statistically significant (for two-sided  $t$  test); \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . <sup>a</sup>Superscript letter  $a$  estimated difference 0.05 ( $P=0.40$ ) after adjustment for change in CD4 and HIV RNA. CPE Rank CNS penetration rank of ART (Letendre et al. 2010), CT Color Trails tests, FT Finger Tapping test (nondominant hand), GPB Grooved Pegboard (dominant hand), QNPZ-5 quantitative neurocognitive performance z score, TG Timed Gait test

Table 1

Baseline characteristics (*N*=258)

Characteristic	Percent or median (IQR)
Demographics	
Age (years)	40 (35, 45)
Female (%)	42.2
Race/ethnicity (%)	
Asian	56.6
Black	18.6
White, other, or unknown	24.8
Education (years)	12 (10, 15)
Location of enrollment (%)	
Australia	5.0
Brazil	6.6
North America	32.9
Thailand	55.4
HIV transmission modes (%)	
Male–male sexual	32.2
Heterosexual	66.7
Intravenous drug use	3.1
Other or unknown	8.9
HIV-related factors	
Duration of HIV infection (years)	5 (4, 8)
Prior AIDS diagnosis (%)	19.0
CD4 count (cells/mm <sup>3</sup> )	513 (430, 667)
Nadir CD4 count (cells/mm <sup>3</sup> )	216 (161, 290)
HIV RNA >400 copies/mL (%)	87.5
Preexisting conditions	
Prior CVD <sup>a</sup> (%)	3.9
Baseline ECG; major abnormality <sup>b</sup> (%)	11.3
Diabetes (%)	3.5
Hepatitis B (%)	2.3
Hepatitis C (%)	7.0
ART use at baseline	
ART naive (%)	3.9
Duration of prior ART use (years)	4 (3, 6)
Current ART use (%)	
NNRTIs	28.3
PIs	65.1
NRTIs	92.6
CPE rank <sup>c</sup>	7 (7, 8)

*ART* antiretroviral therapy, *CPE* central nervous system penetration effectiveness of ART, *CVD* cardiovascular disease, *PI* protease inhibitor, *NNRTI* non-nucleoside reverse transcriptase inhibitor, *NRTI* nucleoside reverse transcriptase inhibitor

<sup>a</sup> Previous myocardial infarction (four), stroke (zero), coronary heart disease (three), congestive heart failure (two), or peripheral vascular disease (five)

<sup>b</sup> Major ischaemic abnormalities, atrioventricular blocks, bundle branch blocks, arrhythmias, or left ventricular hypertrophy

<sup>c</sup> CNS penetration effectiveness rank of the ART regimen; each drug is scored as 1–4, with higher scores for higher CNS penetration (Letendre et al. 2010)



**Table 2**

Neurocognitive tests, baseline and change to month 6

	<u>Baseline</u>		<u>Change to month 6</u>		
	Mean	±1 SD	Mean	SE	P value
QNPZ-5	-0.68*	±0.78	0.17	0.05	<0.001
Test z scores <sup>a</sup>					
GPB	-0.26*	±1.20	0.25	0.09	<0.001
CT1	-0.52*	±1.24	0.28	0.10	<0.001
CT2	-0.00	±0.99	0.18	0.07	<0.001
TG	-1.64*	±1.75	0.11	0.10	0.15
FT	-0.98*	±1.10	0.03	0.09	0.68
Raw tests scores					
GPB (s)	72.5	±26.3	-3.8	1.5	<0.001
CT1 (s)	47.5	±19.0	-3.8	1.3	<0.001
CT2 (s)	93.3	±32.6	-5.1	2.0	<0.001
TG (s)	11.9	±2.1	-0.1	0.1	0.19
FT (taps)	39.3	±7.9	0.1	0.6	0.90

The QNPZ-5 score, individual test z scores, and raw test scores at baseline, and their change to month 6 are shown for 258 participants, pooled across the intermittent ART and continuous ART treatment groups. At baseline, the mean QNPZ-5 score and mean z scores for four of the five tests were below 0, denoting below-average performance compared with a healthy reference population. The QNPZ-5 score and individual test z scores increased from baseline to month 6; see also Fig. 1. There was no evidence for a treatment difference in mean change in QNPZ-5 or any of the z scores

*CT* Color Trails test, *FT* Finger Tapping test (nondominant hand), *GPB* Grooved Pegboard test (dominant hand), *QNPZ-5* quantitative neurocognitive performance z score, *TG* Timed Gait test

\*  $P < 0.001$ , mean z scores  $< 0$  (two-sided *t* test)

<sup>a</sup> z scores were calculated with reference distributions (matched for education [all tests], age [GPB, CT, and FT], gender [GPB and FT], and race/ethnicity [GPB and FT]). Thirty-five participants (13.6 %) had NCI at baseline, defined as z scores  $< -2$  in at least two cognitive ability domains, and the same number had NCI at month 6. Of those who were impaired at baseline, 16 participants (46 %) were not impaired at month 6, and vice versa

**Table 3**

Factors associated with change in test  $z$  scores from baseline to month 6 in multiple regression models, in 258 participants

Factor <sup>a</sup>	QNPZ-5			Regression coefficients for individual test $z$ scores				
	Coefficient	(SE)	<i>P</i> value	GPB	CT1	CT2	FT	TG
Baseline $z$ score <sup>b</sup>	-0.23***	(0.04)	<0.001	-0.36***	-0.48***	-0.32***	-0.49***	-0.26***
Age (years)	n.s.			n.s.	n.s.	-0.01*	n.s.	n.s.
Sex (female)	-0.14*	(0.07)	0.05	-0.28*	-0.27*	n.s.	n.s.	n.s.
Location								
Australia/North America	Ref.							
Brazil	0.25	(0.13)	0.06	0.55*	n.s.	n.s.	0.66**	0.59*
Thailand	0.15*	(0.07)	0.03	0.39**	n.s.	n.s.	n.s.	0.27
Change in CD4 (per 50 cells/mm <sup>3</sup> higher)	n.s.			n.s.	n.s.	n.s.	0.04*	n.s.
Change in log HIV RNA (per 1 log <sub>10</sub> copies/mL higher)	0.15***	(0.06)	0.001	0.19*	n.s.	n.s.	n.s.	n.s.

Coefficients were provided when *P* values for associations were 0.10 or less. Analyses by treatment group: Within the intermittent ART group, an increase in CD4 was associated with an increase in QNPZ-5 (by 0.03 per 50 cells/mm<sup>3</sup> increase; *P*=0.05), but not with any of the individual tests, and there was no evidence for an association within the continuous ART group. Associations of an increase in log<sub>10</sub> HIV RNA with increased QNPZ-5 were similar across both treatment groups. Sensitivity analyses: (1) In a model that included baseline factors only, associations were similar. (2) Among the 202 participants with HIV RNA < 400 copies/mL at baseline and at month 6, change in CD4 was associated with FT only, with a similar coefficient

CT Color Trails test, FT Finger Tapping test (nondominant hand), GPB Grooved Pegboard test (dominant hand), QNPZ-5 quantitative neurocognitive performance  $z$  score, TG Timed Gait test

\* *P* 0.05,

\*\* *P* 0.01,

\*\*\* *P* 0.001, association is significant in multiple regression (for individual tests)

<sup>a</sup> Education (in years) was included as explanatory factor in all regression models, but associations were not significant

<sup>b</sup> QNPZ-5 or individual test  $z$  score