Neurocognitive Function at the First-Line Failure and on the Second-Line Antiretroviral Therapy in Africa: Analyses From the EARNEST Trial

Andrew Kambugu, FRCP, Jennifer Thompson, MSc,† James Hakim, FRCP,‡ Dinah Tumukunde, MPH,§ Joep J. van Oosterhout, PhD,‖¶ Raymond Mwebaze, MMed,# Anne Hoppe, PhD,† James Abach, MBChB,§ Charles Kwobah, MBChB, MSc,** Alejandro Arenas-Pinto, PhD,† Sarah A. Walker, PhD,† and Nicholas I. Paton, MD,††† for the EARNEST Trial Team*

Objective: To assess neurocognitive function at the first-line antiretroviral therapy failure and change on the second-line therapy.

Design: Randomized controlled trial was conducted in 5 sub-Saharan African countries.

Methods: Patients failing the first-line therapy according to WHO criteria after >12 months on non-nucleoside reverse transcriptase inhibitors-based regimens were randomized to the second-line therapy (open-label) with lopinavir/ritonavir (400 mg/100 mg twice daily) plus either 2–3 clinician-selected nucleoside reverse transcriptase inhibitors, raltegravir, or as monotherapy after 12-week induction with raltegravir. Neurocognitive function was tested at baseline, weeks 48 and 96 using color trails tests 1 and 2, and the Grooved Pegboard test. Test results were converted to an average of the 3 individual test z-scores.

Results: A total of 1036 patients (90% of those >18 years enrolled at 13 evaluable sites) had valid baseline tests (58% women, median: 38 years, viral load: 65,000 copies per milliliter, CD4 count: 73 cells per cubic millimeter). Mean (SD) baseline z-score was −2.96 (1.74); lower baseline z-scores were independently associated with older age, lower body weight, higher viral load, lower hemoglobin, less education, fewer weekly working hours, previous central nervous system disease, and taking fluconazole ($P < 0.05$ in multivariable model). Z-score was increased by mean (SE) of +1.23 (0.04) after 96 weeks on the second-line therapy ($P < 0.001$; n = 915 evaluable), with no evidence of difference between the treatment arms ($P = 0.35$).

Conclusions: Patients in sub-Saharan Africa failing the first-line therapy had low neurocognitive function test scores, but performance improved on the second-line therapy. Regimens with more central nervous system-penetrating drugs did not enhance neurocognitive recovery indicating this need not be a primary consideration in choosing a second-line regimen.

Key Words: neurocognitive function, antiretroviral therapy, failure, second line, Africa, trial

(*J Acquir Immune Defic Syndr* 2016;71:506–513)

INTRODUCTION

Highly active antiretroviral therapy (HAART) improves survival and quality-of-life among HIV-infected individuals.[1] The remarkable increase in access to HAART in resource-limited settings (RLS) over the past decade, and also the current global efforts toward earlier HAART initiation, has amplified the benefits of HIV treatment including the impact on HIV-associated neurocognitive disorders (HAND). The introduction of HAART has been associated with strong reductions in prevalence of HIV-associated dementia—the most severe form of HAND. However, the impact of HAART on the milder forms of HAND including mild neurocognitive disorder (MND) and asymptomatic neurocognitive impairment (ANI) is less certain.[2,3] MND and ANI are still common findings in HIV cohorts in the HAART era.[4,5] It has been suggested that ART regimens including drugs with higher levels of central nervous system (CNS) penetration might have greater benefit on HAND, although evidence for this is contradictory.[6–10] The magnitude, severity, and factors associated with HAND (including MND and ANI), and also the response of HAND to antiretroviral therapy, have been fairly well characterized in resource-rich settings, but data from RLS are more limited.[11–13] RLS studies indicate that HAND is common, with significant regional differences in prevalence

Received for publication June 29, 2015; accepted October 14, 2015.

From the *Research Program, Infectious Diseases Institute, Makerere University, Kampala, Uganda; †MRC Clinical Trials Unit at UCL, London, United Kingdom; ‡University of Zimbabwe Clinical Research Centre, Harare, Zimbabwe; §Research Department, Joint Clinical Research Centre (JCRC), Kampala, Uganda; ‖Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre, Malawi; ¶Dignitas International, Zomba, Malawi; #Department of Medicine, St. Francis of Nsambya Hospital, Kampala, Uganda; **Clinical Research Centre, Moi University School of Medicine, Eldoret, Kenya; and ††Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

Supported by a grant from the European and Developing Countries Clinical Trials Partnership (IP.2007.33011.003), with funding from the Medical Research Council, United Kingdom.

The authors have no conflicts of interest to disclose.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jaids.com).

Members of the Europe Africa Research Network for Evaluation of the Second-line Therapy (EARNEST) Trial. Teams are listed in the Appendix (see Supplemental Digital Content, http://links.lww.com/QAI/A778).

Correspondence to: Andrew Kambugu, FRCP, Research Program, Infectious Diseases Institute, Makerere University. P. O. Box, 22418 Kampala, Uganda (e-mail: akambugu@idi.co.ug).

(ranging from 25% to 61%),[14–16] and has similar risk factors as HIV-associated dementia, including low CD4 counts, older age, and male gender.[17,18] However, few studies have characterized the prevalence of neurocognitive impairment among individuals failing the first-line HAART in either RLS or resource-rich settings, and data on neurocognitive responses on the second-line therapy are even more limited.

In this study, we report prospective neurocognitive function measurements using a simple standardized battery of tests in a large multicenter trial of the second-line therapy in Africa.[19]

The aim of this study was to examine the magnitude of and factors associated with neurocognitive impairment at the time of first-line regimen failure and to assess how neurocognitive function changed over 96 weeks on 3 different protease inhibitor (PI)-based second-line regimens.

METHODS

This study was conducted within the large multicenter Europe Africa Research Network for Evaluation of Second-line Therapy (EARNEST) trial. Briefly, EARNEST was an open-label, randomized parallel-group trial (ISRCTN-37737787) performed in 14 centers in 4 sub-Saharan African countries. It enrolled HIV-infected patients >12 years who were failing the first-line HAART (according to WHO clinical, immunological, and/or virological criteria). Participants were randomly assigned 1:1:1 to receive a ritonavir-boosted PI, standardized to lopinavir/ritonavir 400 mg/100 mg twice daily, with either (1) 2–3 new or recycled nucleoside reverse transcriptase inhibitors (NRTIs) chosen without genotyping by the treating doctor (PI/NRTI), (2) raltegravir 400 mg twice daily (PI/RAL), or (3) raltegravir induction for 12 weeks only (PI-mono). Additional details, including the eligibility criteria, study design, and site settings, are described elsewhere.[14]

The study (including the neurocognitive assessments as part of the main trial protocol) was approved by ethics committees and regulatory agencies in participating countries and the United Kingdom. All participants provided written informed consent.

Neurocognitive Assessment

Neurocognitive function was assessed in all participants at baseline (the first-line treatment failure) and at week 48 and week 96 using 3 simple neurocognitive tests, chosen to reflect frontal subcortical functions, the most common neurocognitive impairments seen in HIV-infected individuals.[20] The Color trails tests are 2-part tests that assess the attention/concentration domain and the cognitive flexibility within the executive functioning domain.[21] The Grooved Pegboard test assesses psychomotor speed and fine motor function in both dominant and nondominant hands.

This simple battery of widely used tests was selected to suit the clinical environments in RLS that are often extremely busy and have no specialized neurocognitive test operators.

The tests were administered by a clinician or research nurse. Quality assurance measures were the use of a standardized testing manual across all study sites, initial and annual training of site staff who were designated to perform the tests, restriction of test performance to the designated staff, and on-site monitoring of a random selection of tests to identify systematic errors in execution.

Each neurocognitive test score was standardized using demographic-adjusted normative means of US origin (predominantly white ethnicity) to give a z-score.[22,23] This was adjusted for age, level of education for the color trail scores, and age alone for the Grooved Pegboard scores.

The z-scores for each hand on the Grooved Pegboard were averaged and then combined with the z-scores for the color trail 1 and color trail 2 tests to give an average z-score (NPZ-3 score) at each assessment.[24]

Normative means for the Grooved Pegboard data were not available for participants <18 years so they were excluded from analyses. On-site monitoring identified concerns over the procedures used during baseline testing at 1 site (1 out of the 8 sites in Uganda) so this site was excluded from primary analyses but included in a sensitivity analysis.

Statistical Analysis

We assessed the influence of the following risk factors on NPZ-3 scores at the first-line failure: age, sex, weight, body mass index, ART history, viral load, CD4, WHO stage, history of CNS disease, family history of cardiovascular disease, diabetes, alcohol exposure, smoke exposure, hemoglobin, creatinine, social economic factors (availability of food, years of education, employment status, and household monthly income), and concomitant medication.

Years on the first-line ART and creatinine were truncated at approximate 99th percentiles (to avoid undue influence of extreme outliers on the estimated associations). At baseline, the unadjusted association between NPZ-3 score and each factor was modeled using complete case univariable linear regression with continuous factors modeled using fractional polynomials to allow for nonlinear relationships with NPZ-3 score. Factors with univariable $P < 0.2$ were included in a multivariable linear regression, which used backward selection (exit criteria $P = 0.1$) to select independent risk factors using multiple FPs to allow for nonlinear relationships. In the multivariable analysis, multiple imputations using *Stata*'s mi impute command (25 imputations) was used to account for missing risk factor data and missing test times where at least 2 of the 4 test times were known. Sensitivity analyses used only complete cases or color trail norms from an African-American population, or color trail and Grooved Pegboard means from an HIV-negative Ugandan population.[25]

Mean change in NPZ-3 scores from baseline was compared between the 3 treatment arms at weeks 48 and 96 using *t* tests and analysis of variance; generalized estimating equations (independent correlation structure with robust variance, normal distribution) were used to test differences between arms across all weeks.

Generalized estimating equations were also used to investigate the effect of the factors selected in the baseline model on NPZ-3 scores at weeks 48 and 96 (complete cases only), where possible time-updated factors were used.

Statistical tests presented are 2-sided. All analyses were performed in *Stata* version 13.1.

TABLE 1. Characteristics and Unadjusted Associations With NPZ-3 Scores at the First-Line Failure

Characteristic	Overall (N = 1036)	Difference in NPZ-3 Score at the First-Line Failure*	
		Difference (95% CI)	*P*
Demographics			
Female, n (%)	602 (58)	−0.19 (−0.41 to +0.02)	0.08
Age, mean ± SD, yrs	38 ± 10	−0.11 (−0.22 to +0.00)†	0.05
Anthropometric measures			
Weight, mean ± SD, kg	58.4 ± 11.4	+0.26 (+0.17 to +0.35)†	<0.0001
ART history			
Years on Combination ART, mean ± SD	4.3 ± 2.0	−0.00 (−0.05 to +0.05)	0.99
Previous exposure, n (%)			
Zidovudine	662 (64)	−0.40 (−0.62 to −0.18)	<0.0001
Stavudine	664 (64)	+0.21 (−0.01 to +0.43)	0.06
Tenofovir	143 (14)	−0.17 (−0.48 to +0.14)	0.27
Nevirapine	904 (87)	−0.09 (−0.41 to +0.22)	0.56
Efavirenz	315 (30)	−0.00 (−0.23 to +0.23)	0.98
Virology			
Viral load, copies/mL			
Median (IQR)	65,189 (22,151–186,004)	−0.10 (−0.15 to −0.05)‡	<0.0001
n (%) ≥100,000	412 (40)		
Immunology			
CD4, cells/mm^3			
Median (IQR)	73 (29–147)	+0.18 (+0.07 to +0.29)§	0.001
n (%) <100	629 (61)		
Medical history			
WHO stage, n (%)			
Available	638		
1/2	129 (20)	0	
3	275 (43)	−0.03 (−0.40 to +0.34)	0.06
4	234 (37)	−0.37 (−0.75 to +0.01)	
CNS disease, n (%)	88 (8)	−0.32 (−0.70 to +0.06)	0.10
CVD, n/total n (%)	69/1035 (7)	−0.12 (−0.55 to +0.30)	0.57
Diabetes, n/total n (%)	19/1033 (2)	−0.24 (−1.03 to +0.55)	0.56
Alcohol and smoking			
Alcohol, median (IQR), units/wk	0 (0–0)	+0.04 (−0.00 to +0.08)	0.06
Ever smoked, n/total n (%)	159/1033 (15)	+0.17 (−0.12 to +0.47)	0.25
Laboratory test			
Hemoglobin, mean ± SD, g/dL	12.0 ± 2.2	+0.19 (+0.15 to +0.24)	<0.0001
Creatinine, mean ± SD, mg/dL	0.78 ± 0.26	+0.09 (−0.32 to +0.50)	0.67
Socioeconomics			
Regular meals available, n/total n (%)	678/1033 (66)	+0.35 (+0.12 to +0.57)	0.002
Years of education, median (IQR)	11 (7–13)	+0.41 (+0.26 to +0.56)‡	<0.0001
Employment status, n (%)			
Available	1033		
Full time	500 (48)	0	
Part time/occasional work	205 (20)	−0.40 (−0.68 to −0.12)	
Full time student	31 (3)	+0.03 (−0.59 to +0.65)	<0.0001
Unemployed, ill health	134 (13)	−0.95 (−1.28 to −0.62)	
Unemployed, no jobs	163 (16)	−0.39 (−0.69 to −0.08)	
Hours worked per week	27.5 ± 25.5	+0.14 (+0.09 to +0.18)†	<0.0001
Household monthly income, n (%)			
Available	921		
<$50	395 (43)	0	<0.0001
$50–$200	338 (37)	+0.72 (+0.47 to +0.96)	
≥$200	188 (20)	+0.85 (+0.56 to +1.14)	

TABLE 1. (*Continued*) Characteristics and Unadjusted Associations With NPZ-3 Scores at the First-Line Failure

Characteristic	Overall (N = 1036)	Difference in NPZ-3 Score at the First-Line Failure*	
		Difference (95% CI)	*P*
Concomitant medication in last 10 wk, n (%)			
Dapsone	26 (3)	+0.64 (−0.04 to +1.32)	0.06
Cotrimoxazole	952 (92)	−0.57 (−0.95 to −0.18)	0.004
Fluconazole	77 (7)	−0.76 (−1.16 to −0.36)	<0.0001
Isoniazid	82 (8)	−0.46 (−0.85 to −0.06)	0.02
Ciprofloxacin	36 (3)	−0.67 (−1.25 to −0.10)	0.02
Ethambutol	71 (7)	−0.51 (−0.93 to −0.09)	0.02
Pyrazinamide	61 (6)	−0.54 (−0.98 to −0.09)	0.02
Amoxycillin	40 (4)	−0.16 (−0.71 to +0.39)	0.57

P values from univariable linear regression of factor on NPZ-3 score on complete cases with fractional polynomials used to model continuous variables.
*Difference given is difference in NPZ-3 score between groups or for a 1-unit increase unless specified.
†Difference in NPZ-3 score given for a 10-unit increase in the characteristics.
‡Difference in NPZ-3 score given for a doubling in the characteristic.
§Difference in NPZ-3 score given for a 100-unit increase in the characteristics.

RESULTS

A total of 1277 individuals were enrolled into the EARNEST trial and randomized across the 3 treatment arms. Analysis of the main trial primary outcome (good disease control at week 96) demonstrated that PI/RAL was not superior to boosted PI/NRTI ($P = 0.21$) but was noninferior. PI-mono was not noninferior to boosted PI/NRTI, and the arm was discontinued after week 96 because of markedly lower viral suppression and increased risk of the emergence of resistance mutations. Baseline characteristics and other outcomes across the 3 study arms were similar and are described elsewhere.[14]

Of the 1156 evaluable participants at the first-line failure (excluding 74 aged <18 years and 47 from the single site with implementation inconsistencies), 1036 (90%) had valid results for all 3 neurocognitive test domains (see Table S1, Supplemental Digital Content, http://links.lww.com/QAI/A770). The main reasons for invalid tests were illiteracy (n = 102 tests) and poor vision (n = 51 tests) (see Table S2, Supplemental Digital Content, http://links.lww.com/QAI/A770). The mean ± SD z-score for color trails 1 and 2 was −3.72 ± 2.37 and −2.73 ± 2.16, respectively, and for the combined pegboard z-score was −2.63 ± 2.20 (see Table S3, Supplemental Digital Content, http://links.lww.com/QAI/A770).

Factors Associated With Neurocognitive Function at Baseline

The mean ± SD NPZ-3 score at the first-line failure was −2.96 ± 1.74. Tables 1 and 2 show the unadjusted univariable and adjusted multivariable associations with NPZ-3 score at the first-line failure, respectively.

In the adjusted multivariable model (Table 2), NPZ-3 scores at the first-line failure were significantly lower in patients who were older [change in z-score per 10 years older −0.25 (95% confidence interval: −0.35 to −0.14) $P < 0.0001$] had lower body weight [per 10 kg heavier +0.12 (0.02 to 0.21) $P = 0.01$], higher viral loads [per doubling −0.07 (−0.12 to −0.03) $P = 0.002$], lower hemoglobin [per 1 mg/dL higher +0.16 (+0.11 to +0.21) $P < 0.0001$], fewer years of education [per doubling +0.39 (+0.26 to +0.52) $P < 0.0001$], worked fewer hours per week [per 10 hours longer +0.09 (+0.05 to +0.14) $P < 0.0001$], had a previous CNS disease [−0.45 (−0.82 to −0.08) $P = 0.02$], or had taken fluconazole in the last 10 weeks [−0.61 (−0.99 to −0.22) $P = 0.002$].

There was a trend toward NPZ-3 scores, also being lower in those with lower CD4 cell count [per 100 cells per cubic millimeter higher +0.10 (−0.00 to +0.21) $P = 0.06$], lower household monthly income [vs <$50: $50–$200 +0.29 (+0.03 to +0.54); >$200 +0.21 (−0.15 to +0.56); $P = 0.08$], and not taking dapsone in the last 10 weeks [+0.55 (−0.09 to +1.19) $P = 0.09$].

Significant unadjusted effects of previous ART exposure, availability of regular meals, employment status, and

TABLE 2. Multivariable Associations With NPZ-3 Score at the First-Line Failure

Characteristic	Difference in NPZ-3 Score (95% CI)* (N = 1137)	*P*
Age per 10 yr older	−0.25 (−0.35 to −0.14)	<0.0001
Weight per 10 kg heavier	+0.12 (+0.02 to +0.21)	0.01
Viral load at failure per doubling	−0.07 (−0.12 to −0.03)	0.002
CD4 at failure per 100 cell higher	+0.10 (−0.00 to +0.21)	0.06
Hemoglobin per 1 g/dL higher	+0.16 (+0.11 to +0.21)	<0.0001
Years of education per doubling	+0.39 (+0.26 to +0.52)	<0.0001
Hours worked per week per 10 h longer	+0.09 (+0.05 to +0.14)	<0.0001
Household income		
≤$50	0	0.08
$50–$200	+0.29 (+0.03 to +0.54)	
>$200	+0.21 (−0.15 to +0.56)	
Previous CNS disease	−0.45 (−0.82 to −0.08)	0.02
Fluconazole in the last 10 wk	−0.61 (−0.99 to −0.22)	0.002
Dapsone in the last 10 wk	+0.55 (−0.09 to +1.19)	0.09

*Also adjusted for center ($P < 0.0001$). Multivariable linear regression based on multiple imputations and allowing nonlinearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with $P < 0.2$ (exit $P = 0.1$).

other concomitant medication were no longer independent predictors after adjusting for the characteristics above.

All sensitivity analyses gave broadly comparable results (see Tables S4–S8, Supplemental Digital Content, http://links.lww.com/QAI/A770).

Neurocognitive Response to Treatment

Overall, the NPZ-3 score increased on the second-line therapy with a mean ± SE change across all 3 study arms of +0.91 ± 0.04 and +1.23 ± 0.04 at week 48 and week 96, respectively ($P < 0.001$). There was no statistically significant difference between the second-line regimens (Table 3 and Figs. 1A, B) ($P > 0.2$).

At week 48 and 96, NPZ-3 scores were no longer associated with viral load (current viral load $P = 0.69$, viral load at failure $P = 0.38$), years of education at failure ($P = 0.19$), current hours worked per week ($P = 0.20$), CNS disease before current time ($P = 0.70$), fluconazole use before current time ($P = 0.70$), or taking dapsone in 10 weeks before failure ($P = 0.55$) but remained associated ($P < 0.05$) with all other factors that were significantly related to baseline function as listed above (see Table S9, Supplemental Digital Content, http://links.lww.com/QAI/A770).

DISCUSSION

In this analysis of a large second-line ART trial in Africa, we report reduced neurocognitive function scores among individuals failing the first-line therapy. The scores were significantly lower in patients who were older, had lower body weight, higher viral load, lower hemoglobin, fewer years of education, fewer working hours, previous CNS disease, and who were taking fluconazole. Neurocognitive function improved after starting the second-line ART with no significant difference observed between the 3 study arms.

The very low z-scores we observed in our patients may in part be a function of the norms used for adjustment that were derived from a healthy, mostly white, American population. The same American normative data sets have been shown to produce inadequate adjustment of neurocognitive function in African HIV-positive patients living in the United Kingdom, and the limitations may be even greater for our trial population.[26] In a sensitivity analysis, we normalized results using a small data set of HIV-negative individuals from Uganda (see Table S7, Supplemental Digital Content, http://links.lww.com/QAI/A770) and found that evidence of neurocognitive impairment was persisted, but the magnitude of this effect was reduced markedly.[27] Although different normative data sets will generate different relative levels of impairment, the comparison with Ugandan norms together with the independent associations between scores at the first-line failure and multiple HIV disease-related factors regardless of normative data used suggests that much of this impairment is likely to be genuine.

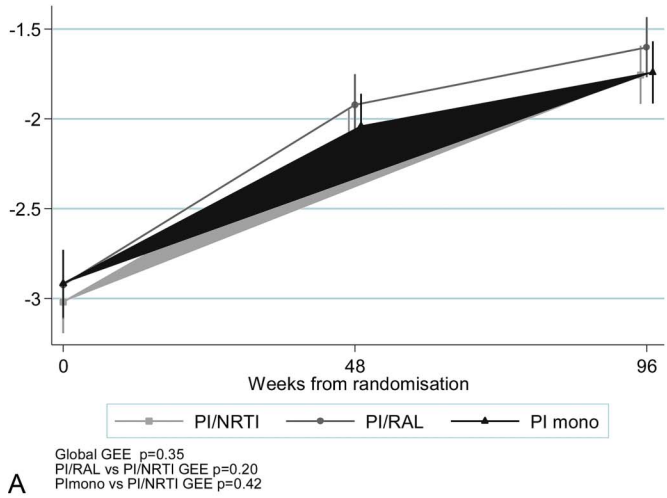
Similar to most other studies, we observed that lower NPZ-3 scores were associated with higher viral loads at the first-line failure after adjusting for other factors.[28,29] HIV is a neurotropic virus that has both direct and indirect pathogenic effects on the CNS, and patients failing the first-line ART in Africa often have very high viral loads (not only in the peripheral circulation but also possibly in the CNS) due to late detection of treatment failure because monitoring is largely clinical and immunological with no routine HIV viral load monitoring. We also found a weak association at the first-line failure between CD4 count and NPZ-3 score independent of viral load. It is noteworthy that patients with a previous CNS disease had lower NPZ-3 scores at the first-line failure. CNS diseases are a very common manifestation of HIV disease in Africa. Infections like *cryptococcal meningitis* not only cause considerable mortality in these settings but can also leave critical damage to the CNS. We observed that taking fluconazole was an independent predictor of lower neurocognitive function even after adjusting for previous CNS disease. It could be that patients taking fluconazole were generally sicker in a variety of ways than those who were not taking this medication. These multiple disease-related associations indicate that the cause of severe neurocognitive impairment is likely multifactorial, in keeping with the heterogeneity of patients' clinical condition at the time of the first-line failure.

The study also found a strong independent association between age and also years of education and NPZ-3 scores among patients failing the first-line ART. These factors are well

TABLE 3. Changes in NPZ-3 Score by the Second-Line Regimen

	PI/NRTI, N = 390	PI/RAL, N = 389	PI Mono, N = 377	Global *P*	PI/RAL vs PI/NRTI		PI Mono vs PI/NRTI	
					Difference (95% CI)	*P*	Difference (95% CI)	*P*
Week 0								
Available	359	345	332					
Mean score ± SD	−3.02 ± 1.7	−2.92 ± 1.8	−2.92 ± 1.8					
Week 48								
Available	324	315	304					
Mean change ± SE	+0.86 ± 0.07	+0.95 ± 0.07	+0.91 ± 0.07	0.65	+0.09 (−0.10 to +0.28)	0.34	+0.05 (−0.15 to +0.24)	0.65
Week 96								
Available	311	306	298					
Mean change ± SE	+1.23 ± 0.07	+1.28 ± 0.07	+1.18 ± 0.08	0.66	+0.05 (−0.16 to +0.26)	0.65	−0.04 (−0.26 to +0.16)	0.64

Under 18 years excluded from all analyses. *P* values from analysis of variance (ANOVA) and *t* tests.



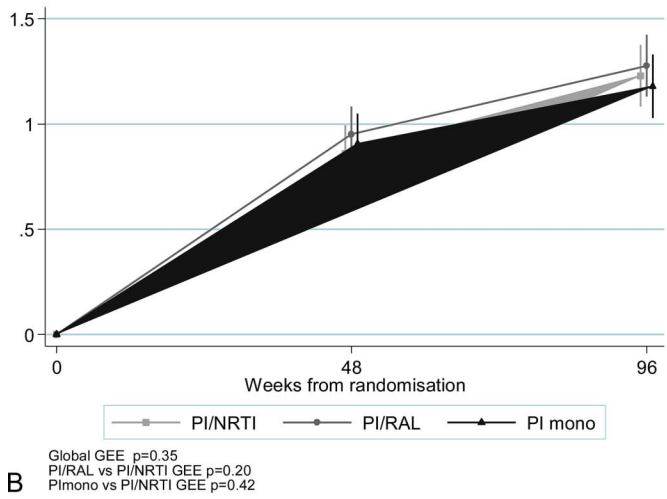


FIGURE 1. A, Mean absolute NPZ-3 score over time on the second-line therapy. B, Mean change in NPZ-3 score over time on the second-line therapy.

known to influence neurocognitive function, which is why neurocognitive data are usually presented as z-scores that attempt to adjust for these factors. The residual associations we have observed are likely to represent incomplete adjustment. Although the color trail tests were adjusted for age and education level, the pegboard scores were adjusted for age only.

Our study additionally provides the first substantive data on the changes in neurocognitive function on the second-line therapy in a large population. We found evidence of improvement in neurocognitive function 48 weeks after starting the second-line therapy, which continued to week 96. This indicates that at least some of the excess impairment associated with the first-line failure is likely to be reversible and is a further illustration of the clinical benefits (aside from avoidance of death and opportunistic infections) that may accrue from starting patients with ART failure on the second-line therapy.

The similarity of the improvement of neurocognitive function across the 3 study arms is surprising for several reasons. First, the PI-mono arm had markedly worse systemic virological suppression rates, which has been associated with progression of CNS disease.[30] Second, the 2 combination arms had greater CNS 4 (CPE) score than the PI-mono arm (PI/NRTI combined score of 6, based on TDF/3 TC as the commonest NRTI selection; PI/RAL combined score of 6; and PI-mono score of 3), often considered to be related to neurocognitive outcomes.[31] Although superior neurocognitive recovery might have been expected in the NRTI-containing arm given that CNS penetration of this class is well established, most of the patients in this arm were taking lamivudine with tenofovir, which has the lowest CPE in this class. Raltegravir and lopinavir have similarly good CPE scores, and we would therefore have expected an improved neurocognitive response in the arm in which they were combined.

The similar response in the 3 arms suggests that the general response to ART (including recovery in general health, recovery from opportunistic infections, and improvement in mental status and nutritional status) rather than CNS drug penetration is the key determinant of neurocognitive function among patients on ART. The longitudinal changes in neurocognitive function and comparisons across study arms are likely to be reliable,

less dependent on the validity of normative data described above.

Additional possible limitations of this study are that we used a smaller test battery (3 domains), and it is possible that a more comprehensive battery might have given a different picture. Because key function domains such as learning and memory were not explored, we cannot tell whether the observed recovery with second line is limited to the motor domains with possible persistence or even progression of poor performance on other cognitive function domains. However, pragmatic considerations made use of a more comprehensive neurocognitive test battery impossible, given the scale of the study with over 1000 patients tested on repeat occasions, located across a diversity of sites and challenging settings.

We have shown that this short battery of well-established tests can detect changes in response to therapy.

Moreover, this test battery was performed by non-specialists and has the potential to be rolled out in real-world settings to document prospective neurocognitive changes on ART. As with all such studies, we cannot exclude the possibility that practice effects are contributed to the some of the observed improvements in neurocognitive function over time. However, an HIV clinical trial in clinically stable patients that applied a similar brief battery of tests at annual intervals found an increase in NPZ-5 score of 0.53 after 3–5 years of follow-up,[32] and a similarly modest change (NPZ-5 increase of 0.13) was observed in a trial that retested stable patients with a similar battery after 6 months.[33] Thus, it is unlikely that practice effects alone would explain the magnitude of change in neurocognitive function (increase in NPZ-3 score of 1.2 over 96 weeks) that we observed. Finally, we did not systematically evaluate participants for depression and therefore did not determine its influence on neurocognitive function test results.[34]

In summary, our study suggests that neurocognitive function is reduced among individuals failing the first-line HAART. We documented improvements in neurocognitive function that occur on the second-line ART irrespective of the antiretroviral regimens used in the study, suggesting that the penetration of drugs into the CNS may not be a primary consideration in selecting a second-line regimen. These findings may provide an additional justification for timely identification of the first-line failure and switch to the second-line therapy.

REFERENCES

1. Palombi L, Marazzi MC, Guidotti G, et al; DREAM Program. Incidence and predictors of death, retention, and switch to second-line regimens in antiretroviral- treated patients in sub-Saharan African Sites with comprehensive monitoring availability. *Clin Infect Dis.* 2009;48:115–122.
2. Mothobi NZ, Brew BJ. Neurocognitive dysfunction in the highly active antiretroviral therapy era. *Curr Opin Infect Dis.* 2012;25:4–9.
3. Joska JA, Gouse H, Paul RH, et al. Does highly active antiretroviral therapy improve neurocognitive function? A systematic review. *J Neurovirol.* 2010;16:101–114.
4. Xia C, Luo D, Yu X, et al. HIV-associated dementia in the era of highly active antiretroviral therapy (HAART). *Microbes Infect.* 2011;13:419–425.
5. Nabha L, Duong L, Timpone J. HIV-associated neurocognitive disorders: perspective on management strategies. *Drugs*. 2013;73:893–905.
6. Smurzynski M, Wu K, Letendre S, et al Effects of central nervous system antiretroviral penetration on cognitive functioning in the ALLRT cohort. *AIDS*. 2011;25:357–365.
7. Casado JL, Marín A, Moreno A, et al. Central nervous system antiretroviral penetration and cognitive functioning in largely pretreated HIV-infected patients. *J Neurovirol.* 2014;20:54–61.
8. Fabbiani M, Grima P, Milanini B, et al. Antiretroviral neuropenetration scores better correlate with cognitive performance of HIV-infected patients after accounting for drug susceptibility. *Antivir Ther.* 2015;20:441–447.
9. Caniglia EC, Cain LE, Justice A, et al; HIV-CAUSAL Collaboration. Antiretroviral penetration into the CNS and incidence of AIDS-defining neurologic conditions. *Neurology*. 2014;83:134–141.
10. Cross HM, Combrinck MI, Joska JA. HIV-associated neurocognitive disorders: antiretroviral regimen, central nervous system penetration effectiveness, and cognitive outcomes. *S Afr Med J.* 2013;103:758–762.
11. Cross S, Önen N, Gase A, et al. Identifying risk factors for HIV-associated neurocognitive disorders using the international HIV dementia scale. *J Neuroimmune Pharmacol.* 2013;8:1114–1122.
12. McCombe JA, Vivithanaporn P, Gill MJ, et al. Predictors of symptomatic HIV-associated neurocognitive disorders in universal health care. *HIV Med.* 2013;14:99–107.
13. Pozniak A, Rackstraw S, Deayton J, et al. HIV-associated neurocognitive disease: case studies and suggestions for diagnosis and management in different patient subgroups. *Antivir Ther.* 2014;19:1–13.
14. Atashili J, Gaynes BN, Pence BW, et al. Prevalence, characteristics and correlates of a positive-dementia screen in patients on antiretroviral therapy in Bamenda, Cameroon: a cross-sectional study. *BMC Neurol.* 2013;13:86.
15. Kelly CM, van Oosterhout JJ, Ngwalo C, et al. HIV associated neurocognitive disorders (HAND) in Malawian adults and effect on adherence to combination anti-retroviral therapy: a cross sectional study. *PLoS One*. 2014;9:e98962.
16. Nakku J, Kinyanda E, Hoskins S. Prevalence and factors associated with probable HIV dementia in an African population: a cross-sectional study of an HIV/AIDS clinic population. *BMC Psychiatry.* 2013;13:126.
17. Njamnshi AK, Bissek AC, Ongolo-Zogo P, et al. Risk factors for HIV-associated neurocognitive disorders (HAND) in sub-Saharan Africa: the case of Yaoundé-Cameroon. *J Neurol Sci.* 2009;285:149–153.
18. Lawler K, Mosepele M, Ratcliffe S, et al. Neurocognitive impairment among HIV-positive individuals in Botswana: a pilot study. *J Int AIDS Soc.* 2010;13:15.
19. Paton NI, Kityo C, Hoppe A, et al; EARNEST Trial Team. Assessment of second-line antiretroviral regimens for HIV therapy in Africa. *N Engl J Med.* 2014;371:234–247.
20. Carey C, Woods SP, Rippeth JD, et al. Initial Validation of a Screening Battery for the detection of HIV-associated Cognitive Impairment. *Clin Neuropsychol.* 2004;18:234–248.
21. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. *Nat Protoc.* 2006;1:2277–2281.
22. *Grooved Pegboard User's Manual*. Europe LIC, ed. Loughborough, United Kingdom: Lafayette Instrument Co. Europe; 2003.
23. D'Elia L, ed. *Color Trails Test Professional Manual*. 2nd ed. Lutz, FL: PAR Psychological Assessmant Resources, Inc; 1996:82.
24. Arenas-Pinto A, Winston A, Stöhr W, et al; PIVOT Trial Team. Neurocognitive function in HIV-infected patients: comparison of two methods to define impairment. *PLoS One*. 2014;9:e103498.
25. Robertson KR, Nakasujja N, Wong M, et al. Pattern of neuropsychological performance among HIV positive patients in Uganda. *BMC Neurol.* 2007;7:8.
26. Winston A, Arenas-Pinto A, Stöhr W, et al; PIVOT Trial Team. Neurocognitive function in HIV infected patients on antiretroviral therapy. *PLoS One*. 2013;8:e61949.
27. Sacktor N, Nakasujja N, Skolasky R, et al. Antiretroviral therapy improves cognitive impairment in HIV+ individuals in sub-Saharan Africa. *Neurology*. 2006;67:311–314.
28. Giesbrecht CJ, Thornton AE, Hall-Patch C, et al. Select neurocognitive impairment in HIV-infected women: associations with HIV viral load, hepatitis C virus, and depression, but not leukocyte telomere length. *PLoS One*. 2014;9:e89556.

29. Tozzi V, Balestra P, Serraino D, et al. Neurocognitive impairment and survival in a cohort of HIV-infected patients treated with HAART. *AIDS Res Hum Retroviruses.* 2005;21:706–713.
30. Childs EA, Lyles RH, Selnes OA, et al. Plasma viral load and CD4 lymphocytes predict HIV-associated dementia and sensory neuropathy. *Neurology*. 1999;52:607–613.
31. Eisfeld C, Reichelt D, Evers S, et al. CSF penetration by antiretroviral drugs. *CNS Drugs.* 2013;27:31–55.
32. Paton NI, Stöhr W, Arenas-Pinto A, et al; for the Protease Inhibitor Monotherapy Versus Ongoing Triple Therapy (PIVOT) Trial Team. Protease inhibitor monotherapy for long-term management of HIV infection: a randomised, controlled, open-label, non-inferiority trial. *Lancet HIV.* 2015;2:e417–26.
33. Grund B, Wright EJ, Brew BJ, et al; INSIGHT SMART Study Group. Improved neurocognitive test performance in both arms of the SMART study: impact of practice effect. *J Neurovirol.* 2013;19: 383–392.
34. Evans VC, Iverson GL, Yatham LN, et al. The relationship between neurocognitive and psychosocial functioning in major depressive disorder: a systematic review. *J Clin Psychiatry.* 2014;75:1359–1370.