Neurocognitive function at first-line failure and on second-line antiretroviral therapy in Africa: Analyses from the EARNEST Trial

Authors:

Andrew Kambugu[1]

Jennifer Thompson[2]

James Hakim[3]

Dinah Tumukunde[4]

Joep J van Oosterhout[5,6]

Raymond Mwebaze[7]

Anne Hoppe[2]

James Abach[4]

Charles Kwobah[8]

Alejandro Arenas-Pinto[2]

A. Sarah Walker[2]

Nicholas I. Paton[2,9]

For the EARNEST Trial Team

[1]Infectious Diseases Institute, Makerere University, Kampala, Uganda

[2] MRC Clinical Trials Unit at UCL, London, United Kingdom; [3]University of Zimbabwe Clinical Research Centre, Harare, Zimbabwe; [4]Joint Clinical Research Centre (JCRC), Kampala, Uganda; [5]Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre, Malawi; [6]Dignitas International, Zomba; [7]St Francis of Nsambya Hospital, Kampala, Uganda; [8]Moi University School of Medicine, Eldoret, Kenya; [9]National University of Singapore, Singapore

Members of the Europe Africa Research Network for Evaluation of Second-line Therapy (EARNEST) Trial Team are listed in the Appendix

Target journal: AIDS

Word count: 2814 (limit 3500)

Abstract: 241 (limit 250 words)

Tables 3, Figures 2 (maximum 5 tables or figures)

References: 34 (limit not stated)

Correspondence to: Andrew Kambugu, Infectious Diseases Institute, Makerere University. P. O. Box 22418 Kampala Uganda Tel: +256312307227
E-mail: akambugu@idi.co.ug

ABSTRACT

Objective

To assess neurocognitive function at first-line antiretroviral therapy failure and change on second–line therapy

Design

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

Methods

Patients failing first-line therapy according to WHO criteria after >12 months on NNRTI-based regimens were randomised to second-line therapy (open-label) with lopinavir/ritonavir (400mg/100mg twice daily) plus either 2-3 clinician-selected NRTIs, raltegravir, or as monotherapy after 12 weeks induction with raltegravir. Neurocognitive function was tested at baseline, weeks 48 and 96 using colour 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

1036 patients (90% of those >18y enrolled at 13 evaluable sites) had valid baseline tests (58% female, median 38 years, viral load 65,000 c/ml, CD4 count 73 cells/mm^3). 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 haemoglobin, less education, fewer weekly working hours, previous CNS disease, and taking fluconazole ($P<0.05$ in multivariable model). Z-score increased by mean (SE) of +1.23 (0.04) after 96 weeks on 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 first-line therapy had low neurocognitive function test scores, but performance improved on second-line therapy. Regimens with more CNS-penetrating drugs did not enhance neurocognitive recovery indicating this need not be a primary consideration in choosing a second-line regimen.

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, as well as the current global efforts towards earlier HAART initiation, have 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 (HAD)-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 CNS penetration might have greater benefit on HAND, although evidence for this is contradictory.[6 7 8 9 10] The magnitude, severity and factors associated with HAND (including MND and ANI), as well as the response of HAND to antiretroviral therapy, have been fairly well characterized in resource-rich settings, but data from resource-limited settings (RLS) are more limited.[11 12 13] RLS studies indicate that HAND is common, with significant regional differences in prevalence (ranging from 25%-61%),[14 15 16] and has similar risk factors as HAD, including low CD4 counts, older age, and male gender.[17 18] However, few studies have characterized the prevalence of neurocognitive impairment among individuals failing first-line HAART in either RLS or resource-rich settings, and data on neurocognitive responses on second-line therapy are even more limited.

Here, we report prospective neurocognitive function measurements using a simple standardised battery of tests in a large multicentre trial of 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 assess how neurocognitive function changed over 96 weeks on 3 different PI based second-line regimens.

Methods

The current study was conducted within the large multi-centre 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 centres in 4 sub-Saharan African countries. It enrolled HIV-infected patients >12 years who were failing first-line HAART (according to WHO clinical, immunological and/or virological criteria). Participants were randomly assigned 1:1:1 to receive a ritonavir-boosted protease inhibitor (PI), standardised to lopinavir/ritonavir 400mg/100mg twice daily, with either (i) 2-3 new or recycled nucleoside-reverse-transcriptase-inhibitors (NRTIs) chosen without genotyping by the treating doctor (PI/NRTI); (ii) raltegravir 400mg twice daily (PI/RAL); or (iii) 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 UK. All participants provided written informed consent.

Neurocognitive assessment

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

This simple battery of widely-used tests was selected to suit the clinical environments in resource-limited settings 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 standardised using demographic-adjusted normative means of US origin (predominantly Caucasian ethnicity) to give a z-score.[22 23] This was adjusted for age, and level of education for the colour 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 colour trail 1 and colour 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 one site (one 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 first-line failure: age, sex, weight, BMI, ART history, viral load, CD4, WHO stage, history of CNS disease, family history of cardiovascular disease, diabetes, alcohol exposure, smoke exposure, haemoglobin, creatinine, social economic factors (availability of food, years of education, employment status, household monthly income), and concomitant medication.

Years on 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 modelled using complete case univariable linear regression with continuous factors modelled using fractional polynomials (FP) to allow for non-linear 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 non-linear relationships. In the multivariable analysis multiple imputation 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 colour trail norms from an African-American population, or colour trail and grooved pegboard means from an HIV-negative Ugandan population.[25]

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

GEE (independent correlation structure with robust variance, normal distribution) 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 two-sided. All analyses were carried out in *Stata®* version 13.1.

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 non-inferior. PI-mono was not non-inferior to boosted PI/NRTI, and the arm was discontinued after week 96 due to 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 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 (supplementary table 1). The main reasons for invalid tests were illiteracy (n=102 tests) and poor vision (n=51 tests) (supplementary table 2). The mean±sd z-score for colour trails 1 and 2 were -3.72±2.37, and -2.73±2.16 respectively, and for the combined pegboard z-score was -2.63±2.20 (supplementary table 3).

Factors Associated with neurocognitive function at Baseline

The mean±sd NPZ-3 score at 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 first-line failure respectively.

In the adjusted multivariable model (Table 2), NPZ-3 scores at first-line failure were significantly lower in patients who were older (change in Z-score per 10 years older -0.25 (95% CI -0.35,-0.14) p<0.0001), had lower body weight (per 10 kg heavier +0.12 (0.02,0.21) p=0.01), higher viral loads (per doubling -0.07 (-0.12,-0.03) p=0.002), lower hemoglobin (per 1mg/dl higher +0.16 (+0.11,+0.21) p<0.0001), fewer years of education (per doubling +0.39 (+0.26,+0.52) p<0.0001), worked fewer hours per week (per 10 hours

longer +0.09 (+0.05,+0.14) p<0.0001), had a previous CNS disease (-0.45 (-0.82,-0.08) p=0.02), or had taken fluconazole in the last 10 weeks (-0.61 (-0.99,-0.22) p=0.002).

There was a trend towards NPZ-3 scores also being lower in those with lower CD4 cell count (per 100 cells/mm3 higher +0.10 (-0.00, +0.21) p=0.06), lower household monthly income (vs <$50: $50-$200 +0.29 (+0.03, +0.54); >$200 +0.21(-0.15, +0.56); p=0.08), and not taking dapsone in the last 10 weeks (+0.55 (-0.09, +1.19) p=0.09).

Significant unadjusted effects of prior ART exposure, availability of regular meals, employment status, and other concomitant medication were no longer independent predictors after adjusting for the characteristics above.

All sensitivity analyses gave broadly comparable results (supplementary tables 4-8).

Neurocognitive Response to Treatment

Overall, the NPZ-3 score increased on second-line therapy with a mean±SE change across all three 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 Figures 1a and 1b) (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 prior to current time (p=0.70), fluconazole use prior to current time (p=0.70), or taking dapsone in the 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 (supplementary table 9).

Discussion

In this analysis of a large second-line ART trial in Africa, we report reduced neurocognitive function scores among individuals failing first-line therapy. The scores were significantly lower in patients who were older, had lower body weight, higher viral load, lower haemoglobin, fewer years of education, fewer working hours, previous CNS disease, and who were taking fluconazole. Neurocognitive function improved after starting 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 Caucasian, American population. The same American normative datasets have been shown to produce inadequate adjustment of neurocognitive function in African HIV-positive patients living in the UK and the limitations may be even greater for our trial population.[26] In a sensitivity analysis, we normalized results using a small dataset of HIV negative individuals from Uganda (Supplementary Table 7), and found that evidence of neurocognitive impairment persisted but the magnitude of this effect was reduced markedly.[27] Although different normative datasets will generate different relative levels of impairment, the comparison with Ugandan norms together with the independent associations between scores at first-line failure and multiple HIV disease-related factors regardless of normative data used suggest 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 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 first-line ART in Africa often have very high viral loads (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 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 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 first-line failure.

The study also found a strong independent association between age as well as years of education and NPZ-3 scores among patients failing first-line ART. These factors are well 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 colour 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 second-line therapy in a large population. We found evidence of improvement in neurocognitive function 48 weeks after starting second-line therapy which continued to week 96. This indicates that at least some of the excess impairment associated with 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 second-line therapy.

The similarity of the improvement of neurocognitive function across the three study arms is surprising for several reasons. Firstly, the PI-mono arm had markedly worse systemic virologic suppression rates, which has been associated with progression of CNS disease.[30] Secondly, the two combination arms had greater CNS penetration effectiveness (CPE) score than the PI-mono arm (PI/NRTI combined score of 6, based on TDF/3TC as the commonest NRTI selection; PI/RAL combined score of 6 ; 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 NC 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 NCI changes on ART. As with all such studies, we cannot exclude the possibility that practice effects 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 follow-up,[32] and a similarly modest change (NPZ-5 increase of 0.13) was observed in a trial that re-tested 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. Lastly, 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 first-line HAART. We documented improvements in neurocognitive function that occur on 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 first-line failure and switch to second-line therapy.

References

[1]Palombi L1, Marazzi MC, Guidotti G, Germano P, Buonomo E, Scarcella P, Doro Altan A, Zimba Ida V, San Lio MM, De Luca A; 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 Jan 1;48(1):115-122

[2] Mothobi NZ, Brew BJ.Neurocognitive dysfunction in the highly active antiretroviral therapy era. Curr Opin Infect Dis. 2012 Feb;25(1):4-9

[3] Joska JA, Gouse H, Paul RH, Stein DJ, Flisher AJ Does highly active antiretroviral therapy improve neurocognitive function? A systematic review. J Neurovirol. 2010 Mar;16(2):101-14

[4] Xia C, Luo D, Yu X, Jiang S, Liu S. HIV-associated dementia in the era of highly active antiretroviral therapy (HAART). Microbes Infect. 2011 May;13(5):419-25.

[5] Nabha L, Duong L, Timpone J. HIV-associated neurocognitive disorders: perspective on management strategies. Drugs. 2013 Jun;73(9):893-905

[6] Smurzynski M1, Wu K, Letendre S, Robertson K, Bosch RJ, Clifford DB, Evans S, Collier AC, Taylor M, Ellis R. Effects of central nervous system antiretroviral penetration on cognitive functioning in the ALLRT cohort. AIDS. 2011 Jan 28;25(3):357-65.

[7] Casado JL, Marín A, Moreno A, Iglesias V, Perez-Elías MJ, Moreno S, Corral I.Central nervous system antiretroviral penetration and cognitive functioning in largely pretreated HIV-infected patients. J Neurovirol. 2014 Feb;20(1):54-61.

[8] Fabbiani M, Grima P, Milanini B, Mondi A, Baldonero E, Ciccarelli N, Cauda R, Silveri MC, De Luca A, Di Giambenedetto S.Antiretroviral neuropenetration scores better correlate with cognitive performance of HIV-infected patients after accounting for drug susceptibility. Antivir Ther. 2015;20(4):441-7.

[9] Caniglia EC, Cain LE, Justice A, Tate J, Logan R, Sabin C, Winston A, van Sighem A, Miro JM, Podzamczer D, Olson A, Arribas JR, Moreno S, Meyer L2, del Romero J2 Dabis F, Bucher HC, Wandeler G, Vourli G, Skoutelis A, Lanoy E, Gasnault J, Costagliola D, Hernán MA; HIV-CAUSAL Collaboration.Antiretroviral penetration into the CNS and incidence of AIDS-defining neurologic conditions. Neurology. 2014 Jul 8;83(2):134-41

[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 Sep 18;103(10):758-62

[11] Cross S, Önen N, Gase A, Overton ET, Ances BM.Identifying risk factors for HIV-associated neurocognitive disorders using the international HIV dementia scale.J Neuroimmune Pharmacol. 2013 Dec;8(5):1114-22.

[12] McCombe JA, Vivithanaporn P, Gill MJ, Power C. Predictors of symptomatic HIV-associated neurocognitive disorders in universal health care. HIV Med. 2013 Feb;14(2):99-107.

[13] Pozniak A, Rackstraw S, Deayton J, Barber T, Taylor S, Manji H, Melvin D, Croston M, Nightingale S, Kulasegaram R, Pitkanen M, Winston A. HIV-associated neurocognitive disease: case studies and suggestions for diagnosis and management in different patient subgroups. Antivir Ther. 2014;19(1):1-13.

[14] Atashili J, Gaynes BN, Pence BW, Tayong G, Kats D, O'donnell JK, Ndumbe PM, Njamnshi AK.Prevalence, characteristics and correlates of a positive-dementia screen in patients on antiretroviral therapy in Bamenda, Cameroon: a cross-sectional study. BMC Neurol. 2013 Jul 15;13:86

[15] Kelly CM, van Oosterhout JJ, Ngwalo C, Stewart RC, Benjamin L, Robertson KR, Khoo S, Allain TJ, Solomon T. HIV associated neurocognitive disorders (HAND) in Malawian adults and effect on adherence to combination anti-retroviral therapy: a cross sectional study. PLoS One. 2014 Jun 10;9(6):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 May 3;13(1):126.

[17] Njamnshi AK, Bissek AC, Ongolo-Zogo P, Tabah EN, Lekoubou AZ, Yepnjio FN, Fonsah JY, Kuate CT, Angwafor SA, Dema F, Njamnshi DM, Kouanfack C, Djientcheu Vde P, Muna WF, Kanmogne GD.Risk factors for HIV-associated neurocognitive disorders (HAND) in sub-Saharan Africa: the case of Yaoundé-Cameroon. J Neurol Sci. 2009 Oct 15;285(1-2):149-53

[18] Lawler K, Mosepele M, Ratcliffe S, Seloilwe E, Steele K, Nthobatsang R, Steenhoff A. Neurocognitive impairment among HIV-positive individuals in Botswana: a pilot study. J Int AIDS Soc. 2010 Apr 20;13:15

[19] Paton NI, Kityo C, Hoppe A, Reid A, Kambugu A, Lugemwa A, van Oosterhout JJ, Kiconco M, Siika A, Mwebaze R, Abwola M, Abongomera G, Mweemba A, Alima H, Atwongyeire D, Nyirenda R, Boles J, Thompson J, Tumukunde D, Chidziva E, Mambule I, Arribas JR, Easterbrook PJ, Hakim J, Walker AS, Mugyenyi P; EARNEST Trial Team. Assessment of second-line antiretroviral regimens for HIV therapy in Africa. N Engl J Med. 2014 Jul 17;371(3):234-47.

[20] Carey, C., et al., Initial Validation of a Screening Battery for the detection of HIV-associated Cognitive Impairment. Clinical Neuropsychology, 2004. 18(2): p. 234-48.

[21] Bowie, C.R. and P.D. Harvey, Administration and interpretation of the Trail Making Test. Nat Protoc, 2006. 1(5): p. 2277-81.

[22] (2003) Grooved Pegboard User's Manual. In: Europe LIC, editor. Loughborough, UK.: Lafayette Instrument Co. Europe

[23] D'Elia L, editor (1996) Color Trails test Professional Manual. 2nd ed. Lutz, Florida. USA: PAR Psychological Assessmant Resources, Inc. 82 p

[24] Arenas-Pinto A, Winston A, Stöhr W, Day J, Wiggins R, Quah SP, Ainsworth J, Fleck S, Dunn D, Accoroni A, Paton NI; PIVOT Trial Team. Neurocognitive function in HIV-infected patients: comparison of two methods to define impairment. PLoS One. 2014 Jul 31;9(7):e103498

[25] Robertson KR, Nakasujja N, Wong M, Musisi S, Katabira E, Parsons TD, Ronald A, Sacktor N. Pattern of neuropsychological performance among HIV positive patients in Uganda. BMC Neurol. 2007 Apr 5;7:8

[26] Winston A, Arenas-Pinto A, Stöhr W, Fisher M, Orkin CM, Aderogba K, De Burgh-Thomas A, O'Farrell N, Lacey CJ, Leen C, Dunn D, Paton NI; PIVOT Trial Team. Neurocognitive function in HIV infected patients on antiretroviral therapy. PLoS One. 2013 Apr 30;8(4):e61949

[27] Sacktor N1, 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 Jul 25;67(2):311-4.

[28] Giesbrecht CJ, Thornton AE, Hall-Patch C, Maan EJ, Côté HC, Money DM, Murray M, Pick N 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 Mar 4;9(3):e89556.

[29] Tozzi V, Balestra P, Serraino D, Bellagamba R, Corpolongo A, Piselli P, Lorenzini P, Visco-Comandini U, Vlassi C, Quartuccio ME, Giulianelli M, Noto P, Galgani S, Ippolito G, Antinori A, Narciso P. Neurocognitive impairment and survival in a cohort of HIV-infected patients treated with HAART. AIDS Res Hum Retroviruses. 2005 Aug;21(8):706-13.

[30] Childs EA, Lyles RH, Selnes OA, Chen B, Miller EN, Cohen BA, Becker JT, Mellors J, McArthur JC. Plasma viral load and CD4 lymphocytes predict HIV-associated dementia and sensory neuropathy. Neurology. 1999 Feb;52(3):607-13.

[31] Eisfeld C1, Reichelt D, Evers S, Husstedt I. CSF penetration by antiretroviral drugs. CNS Drugs. 2013 Jan;27(1):31-55.

[32] Nicholas I Paton, Wolfgang Stöhr, Alejandro Arenas-Pinto, Martin Fisher, Ian Williams, Margaret Johnson, Chloe Orkin, Fabian Chen, Vincent Lee, Alan Winston, Mark Gompels, Julie Fox, Karen Scott, David T Dunn, for the Protease Inhibitor Monotherapy Versus Ongoing Triple Therapy (PIVOT) Trial Team* Protease inhibitor monotherapy for long-term manageme of HIV infection: a randomised, controlled, open-label, non-inferiority trial Lancet HIV 2015; 2: e417–26

[33] Grund B, Wright EJ, Brew BJ, Price RW, Roediger MP, Bain MP, Hoy JF, Shlay JC, Vjecha MJ, Robertson KR; INSIGHT SMART Study Group. Improved neurocognitive test performance in both arms of the SMART study: impact of practice effect. J Neurovirol. 2013 Aug;19(4):383-92.

[34] Evans VC, Iverson GL, Yatham LN, Lam RW. The relationship between neurocognitive and psychosocial functioning in major depressive disorder: a systematic review. J Clin Psychiatry. 2014 Dec;75(12):1359-70.

Table 1: Characteristics and unadjusted associations with NPZ-3 scores at first-line failure

Characteristic	Overall (N=1036)	Difference in NPZ-3 Score at first-line failure [a] Difference (95% CI)	P value
Demographics			
Female - n(%)	602 (58%)	-0.19 (-0.41, +0.02)	0.08
Age (years) - mean±sd	38±10	-0.11 (-0.22, +0.00) [b]	0.05
Anthropometric measures			
Weight (kg) mean±sd	58.4±11.4	+0.26 (+0.17, +0.35) [b]	<0.0001
ART history			
Years on Combination ART mean±sd	4.3±2.0	-0.00 (-0.05, +0.05)	0.99
Prior exposure n(%)			
zidovudine	662 (64%)	-0.40 (-0.62, -0.18)	<0.0001
stavudine	664 (64%)	+0.21 (-0.01, +0.43)	0.06
tenofovir	143 (14%)	-0.17 (-0.48, +0.14)	0.27
nevirapine	904 (87%)	-0.09 (-0.41, +0.22)	0.56
efavirenz	315 (30%)	-0.00 (-0.23, +0.23)	0.98
Virology			
Viral load copies/ml			
median (IQR)	65189 (22151-186004)	-0.10 (-0.15, -0.05) [c]	<0.0001
n(%) >=100,000	412 (40%)		
Immunology			
CD4 cells/mm^3			
median (IQR)	73 (29-147)	+0.18 (+0.07, +0.29) [d]	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, +0.34)	0.06
4	234 (37%)	-0.37 (-0.75, +0.01)	
CNS disease n(%)	88 (8%)	-0.32 (-0.70, +0.06)	0.10
CVD n/ total n(%)	69/1035 (7%)	-0.12 (-0.55, +0.30)	0.57
Diabetes n/ total n(%)	19/1033 (2%)	-0.24 (-1.03, +0.55)	0.56
Alcohol and Smoking			
Alcohol (units/week) median(IQR)	0 (0-0)	+0.04 (-0.00, +0.08)	0.06
Ever smoked n/total n(%)	159/1033 (15%)	+0.17 (-0.12, +0.47)	0.25
Laboratory test			
Haemoglobin (g/dl) mean±sd	12.0±2.2	+0.19 (+0.15, +0.24)	<0.0001
Creatinine (mg/dl) mean±sd	0.78±0.26	+0.09 (-0.32, +0.50)	0.67
Socio-economic			
Regular meals available n/total n(%)	678/1033 (66%)	+0.35 (+0.12, +0.57)	0.002
Years of education median (IQR)	11 (7-13)	+0.41 (+0.26, +0.56) [c]	<0.0001
Employment status n(%)			
Available	1033		
Full time	500 (48%)	0	
Part time/occasional work	205 (20%)	-0.40 (-0.68, -0.12)	
Full time student	31 (3%)	+0.03 (-0.59, +0.65)	<0.0001
Unemployed- ill health	134 (13%)	-0.95 (-1.28, -0.62)	
Unemployed- no jobs	163 (16%)	-0.39 (-0.69, -0.08)	
Hours worked per week	27.5±25.5	+0.14 (+0.09, +0.18) [b]	<0.0001
Household monthly income n(%)			
Available	921		
<$50	395 (43%)	0	<0.0001
$50-$200	338 (37%)	+0.72 (+0.47, +0.96)	
≥$200	188 (20%)	+0.85 (+0.56, +1.14)	
Concomitant medication in last 10 weeks			
Dapsone n(%)	26 (3%)	+0.64 (-0.04, +1.32)	0.06
Cotrimoxazole n(%)	952 (92%)	-0.57 (-0.95, -0.18)	0.004
Fluconazole n(%)	77 (7%)	-0.76 (-1.16, -0.36)	<0.0001
Isoniazid n(%)	82 (8%)	-0.46 (-0.85, -0.06)	0.02
Ciprofloxin n(%)	36 (3%)	-0.67 (-1.25, -0.10)	0.02

Characteristic	Overall (N=1036)	Difference in NPZ-3 Score at first-line failure [a] Difference (95% CI)	P value
Ethambutol n(%)	71 (7%)	-0.51 (-0.93, -0.09)	0.02
Pyrazinamid n(%)	61 (6%)	-0.54 (-0.98, -0.09)	0.02
Amoxycillin n(%)	40 (4%)	-0.16 (-0.71, +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.

[a] Difference given is difference in NPZ-3 score between groups or for a 1 unit increase unless specified

[b] Difference in NPZ-3 score given for a 10 unit increase in the characteristic

[c] Difference in NPZ-3 score given for a doubling in the characteristic

[d] Difference in NPZ-3 score given for a 100 unit increase in the characteristic

Table 2: Multivariable associations with NPZ-3 score at first-line failure

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

[a] Also adjusted for centre (p<0.0001). Multivariable linear regression based on multiple imputation, and allowing non-linearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with p<0.2 (exit p=0.1),

Table 3: Changes in NPZ-3 score by second-line regimen

	PI/ NRTI N=390	PI/ RAL N=389	PI mono N=377	Global P value	PI/RAL vs PI/NRTI Difference (95% CI)	PI/RAL vs PI/NRTI p value	PI mono vs PI/NRTI Difference (95% CI)	PI mono vs PI/NRTI p value
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, +0.28)	0.34	+0.05 (-0.15, +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, +0.26)	0.65	-0.04 (-0.26, +0.16)	0.64

Under 18 excluded from all analyses. P values from ANOVA and t-tests

Figure 1a: Mean absolute NPZ-3 score over time on second-line therapy

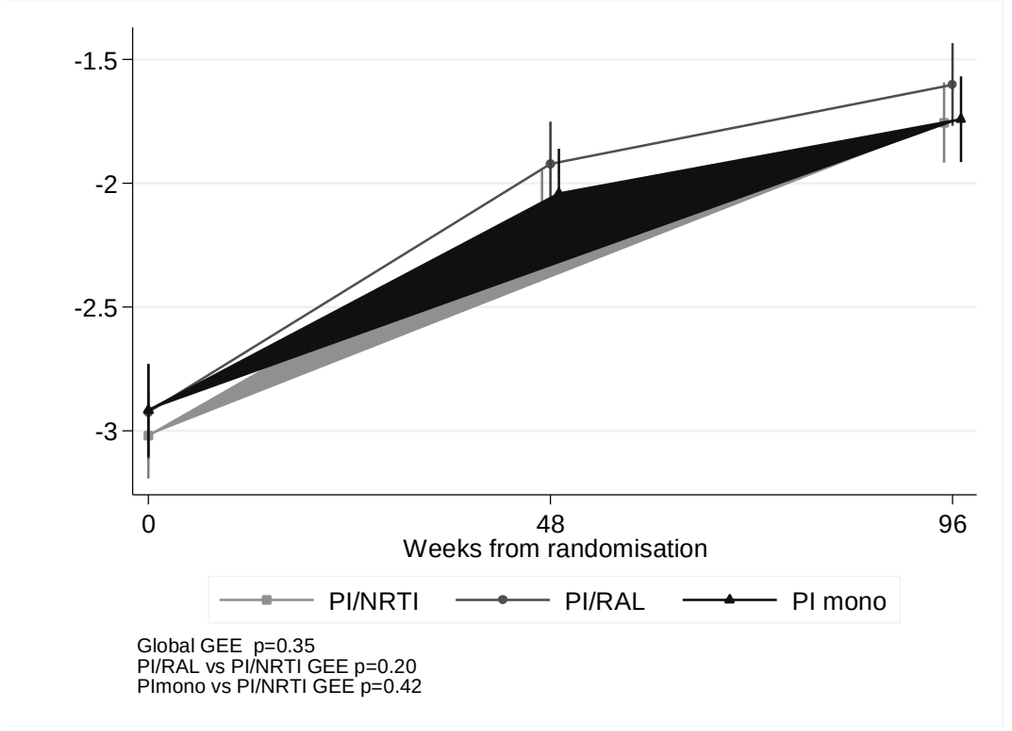
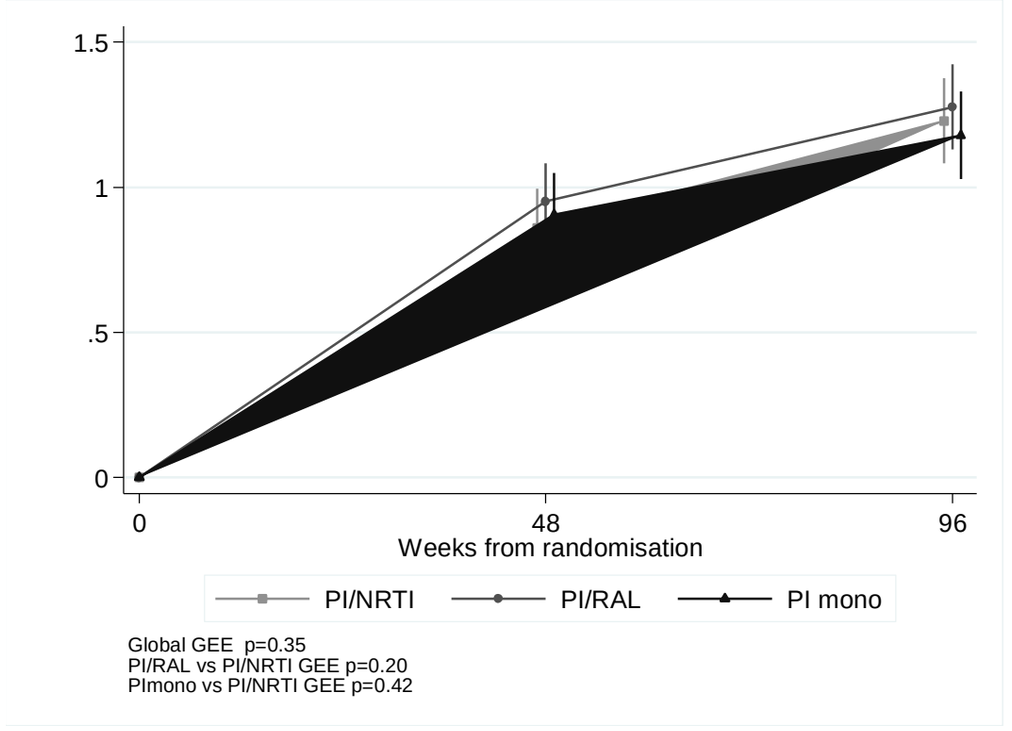


Figure 1b: Mean change in NPZ-3 score over time on second-line therapy



Supplementary Table 1: Pattern of missing tests at first-line failure

Pattern (1=test done, 0 =test missing)				Frequency
Pegboard-d	Pegboard-nd	CT 1	CT2	N=1156 [a]
1	1	1	1	1036
1	1	0	0	66
1	1	1	0	23
0	0	0	0	17
1	1	0	1	4
1	0	1	1	3
0	0	1	1	3
1	0	0	0	2
1	0	1	0	1
0	1	1	1	1

[a] Excluding all tests from one site with implementation inconsistencies and patients who were under 18 at enrolment Pegboard-d: dominant hand; Pegboard-nd: non-dominant hand; CT1: colour trails test 1; CT2: colour trails test 2

Supplementary Table 2: Reasons for not completing neurocognitive tests at first-line failure

	Total
Total Randomised	1277
Over 18 & not at 1 excluded site	1156
Colour Trail 1	
Number of valid tests	1067
Number of invalid tests	89
cannot read/write	51
made a mistake	10
physically not capable/ too ill	1
site problem/error	1
unable to complete	7
Vision	13
Other	1
Colour Trail 2	
Number of valid tests	1047
Number of invalid tests	109
cannot read/write	51
made a mistake	19
site problem/error	1
physically not capable/ too ill	6
unable to complete	13
Vision	17
Other	2
Pegboard dominant	
Number of valid tests	1135
Number of invalid tests	21
no reason given	1
physically not capable/ too ill	5
unable to complete	5
Vision	10
Pegboard non-dominant	
Number of valid tests	1130
Number of invalid tests	26
no reason given	1
physically not capable/ too ill	6
unable to complete	8
Vision	11

Supplementary Table 3: Test times and Z-scores at first-line failure

Test	N	Time (sec) mean±sd	Z-score mean±sd
Overall score	**1036**		**-2.95±1.74**
Colour trail 1	1067	102±50	-3.72±2.37
Colour trail 2	1047	189±85	-2.73±2.16
Mean pegboard	1129		-2.63±2.20
Pegboard dominant	1135	93±38	-2.99±2.48
Pegboard non-dominant	1130	100±45	-2.29±2.22

Z-scores were truncated at -8 before calculating ~~z-scores~~means and then taking the mean of colours trails 1 and 2 z-scores and the mean pegboard z-score to form an overall composite z-score

Supplementary Table 4: Sensitivity analysis - complete cases multivariable risk factors at first-line failure

Characteristic	Adjusted Difference in NPZ-3 score (95% CI) (N=983)	P value
Age per 10 year older	-0.23 (-0.35, -0.12)	<0.0001
Weight per 10kg heavier	+0.13 (0.03, +0.23)	0.01
Viral load at failure per doubling	-0.07 (-0.11, -0.02)	0.01
CD4 at failure per 100 cell higher	+0.05 (-0.06, +0.16)	0.38
Haemoglobin per g/dl higher	+0.16 (+0.11, +0.22)	<0.0001
Years of education per doubling	+0.21 (+0.04, +0.38)	0.01
Hours worked per week per 10 hour longer	+0.08 (+0.03, +0.13)	0.001
Household income		
≤$50	0	0.04
$50-$200	+0.34 (+0.08, +0.61)	
>$200	+0.20 (-0.17, +0.56)	
Previous CNS disease	-0.38 (-0.76, +0.01)	0.05
Fluconazole in the last 10 weeks	-0.55 (-0.99, -0.10)	0.01
Dapsone in the last 10 weeks	+0.62 (-0.04, +1.27)	0.07

Also adjusted for Centre p<0.0001. Multivariable linear regression based on complete cases, and allowing non-linearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with p<0.2 (exit p=0.1),

Supplementary Table 5a: Sensitivity analysis - African American age adjusted colour trail z-scores

Test	N	Z-score mean±sd
Overall score	**888**	**-3.13 ±1.8**
Colour trail 1	914	-4.38±3.8
Colour trail 2	897	-3.48±3.2

African American Age adjusted norms were limited to those ages between 20 and 50.

Supplementary Table 5b: Sensitivity analysis- African American age adjusted colour trail z-scores multivariable risk factors at first-line failure

Characteristic	Adjusted Difference in NPZ-3 score (95% CI) (N=973)	P value
Age per 10 year older	-0.09 (-0.25, +0.06)	0.25
Weight per 10kg heavier	+0.09 (-0.01, +0.19)	0.09
Viral load at failure per doubling	-0.08 (-0.13, -0.03)	0.002
CD4 at failure per 100 cell higher	+0.09 (-0.02, +0.20)	0.10
Haemoglobin per g/dl higher	+0.16 (+0.11, +0.21)	<0.0001
Years of education per doubling	+0.77 (+0.63, +0.90)	<0.0001
Hours worked per week per 10 hour longer	+0.08 (+0.03, +0.12)	0.001
Household income		
≤$50	0	
$50-$200	+0.43 (+0.15, +0.71)	
>$200	+0.48 (+0.09, +0.86)	0.004
Previous CNS disease	-0.69 (-1.09, -0.29)	0.001
Fluconazole in the last 10 weeks	-0.42 (-0.82, -0.03)	0.04
Dapsone in the last 10 weeks	+0.33 (-0.33, +0.99)	0.33

Adjusted for Centre p<0.0001. Multivariable linear regression based on multiple imputation and allowing non-linearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with p<0.2 (exit p=0.1),

Supplementary Table 6a: Sensitivity analysis - African American Education adjusted colour trail z-scores

Test	N	Z-score mean±sd
Overall score	**1036**	**-3.18±1.8**
Colour trail 1	1067	-4.23±3.4
Colour trail 2	1047	-3.50±2.9

Supplementary Table 6b: Sensitivity analysis- African American Education adjusted colour trail z-scores multivariable risk factors at first-line failure

Characteristic	Adjusted Difference in NPZ-3 score (95% CI) (N=983)	P value
Age per 10 year older	-0.55 (-0.66, -0.45)	<0.0001
Weight per 10kg heavier	+0.13 (+0.03, +0.22)	0.007
Viral load at failure per doubling	-0.07 (-0.11, -0.03)	0.002
CD4 at failure per 100 cell higher	+0.09 (-0.01, +0.19)	0.07
Haemoglobin per g/dl higher	+0.14 (+0.10, +0.19)	<0.0001
Years of education per doubling	+0.69 (+0.56, +0.81)	<0.0001
Hours worked per week per 10 hour longer	+0.09 (+0.05, +0.13)	<0.0001
Household income		
≤$50	0	0.02
$50-$200	+0.34 (+0.09, +0.59)	
>$200	+0.36 (+0.02, +0.71)	
Previous CNS disease	-0.45 (-0.80, -0.09)	0.01
Fluconazole in the last 10 weeks	-0.48 (-0.86, -0.11)	0.01
Dapsone in the last 10 weeks	+0.52 (-0.11, +1.14)	0.11

Adjusted for Centre p<0.0001. Multivariable linear regression based on multiple imputation and allowing non-linearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with p<0.2 (exit p=0.1),

Supplementary Table 7a: Sensitivity analysis – Ugandan norm z-scores

Test	N	Z-score mean±sd
Overall score	**1036**	**-0.93±1.48**
Colour trail 1	1067	-1.25±2.25
Colour trail 2	1047	-1.75±2.31
Pegboard- dominant hand	1133	-0.31±1.79
Pegboard- non-dominant hand	1130	+0.07±1.79

Supplementary Table 7b: Sensitivity analysis- Ugandan norm z-scores multivariable risk factors at first-line failure

Characteristic	Adjusted Difference in NPZ-3 score (95% CI) (N=1137)	P value
Age per 10 year older	-0.04 (-0.05, -0.04)	<0.0001
Weight per 10kg heavier	+0.11 (+0.04, +0.19)	0.004
Viral load at failure per doubling	-0.07 (-0.11, -0.03)	<0.0001
CD4 at failure per 100 cell higher	+0.08 (-0.01, +0.16)	0.08
Haemoglobin per g/dl higher	+0.11 (+0.07, +0.15)	<0.0001
Years of education per doubling	+0.66 (+0.55, +0.77)	<0.0001
Hours worked per week per 10 hour longer	+0.09 (+0.05, +0.12)	<0.0001
Household income		
≤$50	0	0.04
$50-$200	+0.25 (+0.4, +0.46)	
>$200	+0.27 (-0.01, +0.56)	
Previous CNS disease	-0.45 (-0.75, -0.15)	0.003
Fluconazole in the last 10 weeks	-0.32 (-0.64, 0.00)	0.05
Dapsone in the last 10 weeks	+0.50 (-0.02, +1.02)	0.06

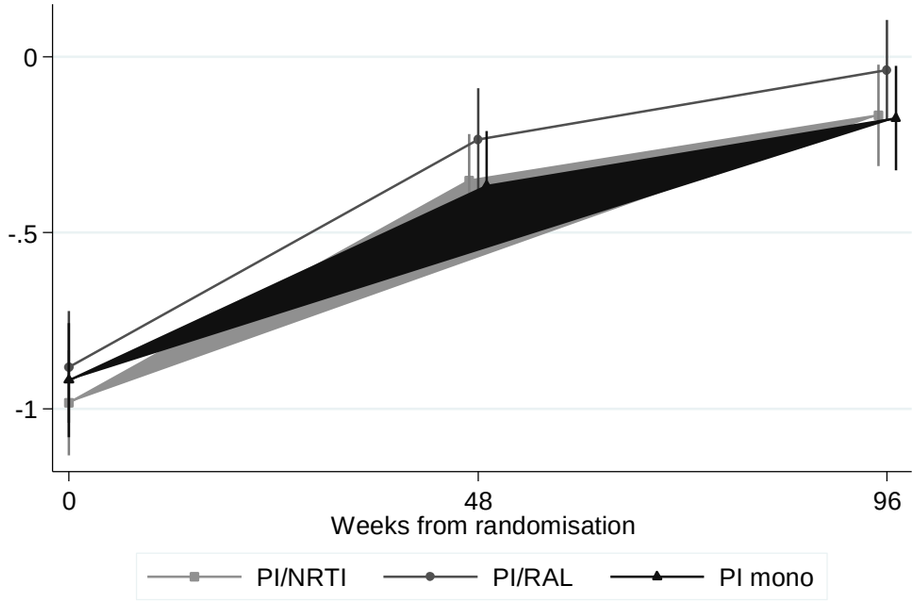
Adjusted for Centre p<0.0001. Multivariable linear regression based on multiple imputation and allowing non-linearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with p<0.2 (exit p=0.1),

Supplementary Table 7c: Sensitivity analysis- Ugandan norm z-scores changes in score by second–line regimen

	PI/ NRTI N=390	PI/ RAL N=389	PI mono N=377	Global P value	PI/RAL vs PI/NRTI Difference (95% CI)	PI/RAL vs PI/NRTI p value	PI mono vs PI/NRTI Difference (95% CI)	PI mono vs PI/NRTI p value
Week 0								
Available	359	345	332					
mean score ±sd	-0.98±1.45	-0.88±1.51	-0.92±1.50					
Week 48								
Available	324	315	304					
mean change ±se	+0.60±0.05	+0.62±0.05	+0.60±0.06	0.96	0.02 (-0.13, +0.16)	0.84	-0.00 (-0.16, +0.15)	0.96
Week 96								
Available	311	306	298					
mean change ±se	+0.81±0.06	+0.80±0.06	+0.75±0.06	0.76	-0.00 (-0.17, +0.16)	0.95	-0.06 (-0.23, +0.11)	0.51

Under 18 excluded from all analyses. P values from ANOVA and t-tests

Supplementary Figure 1: Sensitivity analysis- Ugandan norm z-scores over time on second–line therapy



Global GEE p=0.64
PI/RAL vs PI/NRTI GEE p=0.51
PImono vs PI/NRTI GEE p=0.67

Supplementary Table 8: Sensitivity analysis - Including site with implementation inconsistencies multivariable risk factors at first-line failure

Characteristic	Adjusted Difference in NPZ-3 score (95% CI) (N=983)	P value
Age per 10 year older	-0.25 (-0.36, -0.15)	<0.0001
Weight per 10kg heavier	+0.12 (+0.02, +0.21)	0.02
Viral load at failure per doubling	-0.07 (-0.11, -0.02)	0.004
CD4 at failure per 100 cell higher	+0.10 (-0.01, +0.20)	0.08
Haemoglobin per g/dl higher	+0.16 (+0.11, +0.21)	<0.0001
Years of education per doubling	+0.40 (+0.27, +0.52)	<0.0001
Hours worked per week per 10 hour longer	+0.09 (+0.05, +0.13)	<0.0001
Household income		
≤$50	0	0.10
$50-$200	+0.27 (+0.02, +0.55)	
>$200	+0.20 (-0.15, +0.55)	
Previous CNS disease	-0.46 (-0.83, -0.09)	0.01
Fluconazole in the last 10 weeks	-0.58 (-0.96, -0.19)	0.003
Dapsone in the last 10 weeks	+0.46 (-0.15, +1.08)	0.14

Adjusted for Centre p<0.0001. Multivariable linear regression based on multiple imputation and allowing non-linearity using fractional polynomials. Factors selected based on backwards elimination from all Table 1 factors with p<0.2 (exit p=0.1),

Supplementary Table 9: Multivariable Risk factors associated with NPZ-3 score at week 48 and 96

Characteristic	Adjusted difference in NPZ-3 score (95% CI) (N=830)	P value
Baseline NPZ-3 score	+0.57 (+0.52, +0.61)	<0.001
Weeks on second-line ART		0.02
48 weeks	0	
96 weeks	+0.38 (+0.07, +0.69)	
Age at failure per 10 year older	-0.12 (-0.20, -0.04)	0.003
Weight at current time per 10 kg heavier	+0.06 (-0.00, +0.11)	0.05
Viral load per doubling		
At failure	+0.01 (-0.02, +0.04)	0.38
At current time	+0.00 (-0.02, +0.02)	0.69
CD4 per 100 cell higher		
At failure	-0.01 (-0.07, +0.05)	0.84
Change in CD4 from failure to current time	+0.08 (+0.04, +0.13)	<0.0001
Haemoglobin at current time per g/dl higher	+0.04 (+0.00, +0.07)	0.03
Years of education at failure per doubling	+0.07 (-0.03,+0.18)	0.19
Hours worked per week at current time per 10 hour longer	+0.02 (-0.01, +0.04)	0.20
household income at current time		
≤$50	0	0.04
$50-$200	+0.17 (+0.01, +0.33)	
>$200	+0.25 (+0.05, +0.46)	
CNS disease prior to current time	-0.05 (-0.30, +0.20)	0.70
Fluconazole use prior to current time	-0.05 (-0.29, +0.19)	0.70
Dapsone use in the 10 weeks before failure	+0.11 (-0.26, +0.49)	0.55

Adjusted for Centre (p<0.0001). Model of change in NPZ-3 scores from baseline to week 48 and 96 using generalised estimating equations with an independent working correlation structure and robust variance estimators. Based on complete cases and allowing non-linearity using fractional polynomials.