

A5324

Primary Statistical Analysis Plan

Version 3.0

**A Randomized, Double-Blinded, Placebo-Controlled Trial
Comparing Antiretroviral Intensification with Maraviroc and
Dolutegravir with No Intensification or Intensification with
Dolutegravir Alone for the Treatment of Cognitive Impairment in
HIV**

Protocol Version 2.0

ClinicalTrials.gov Identifier: NCT02519777

02/12/2021

This is ACTG A5324 SAP Version 3.0 with names of authors, names of publication writing team members and analysis timeline redacted.

Table of Contents

TABLE OF CONTENTS	1
1 INTRODUCTION	2
1.1 Purpose	2
1.2 Version History	2
2 STUDY OVERVIEW	3
2.1 Study Design	3
2.2 Hypotheses	3
2.2.1 Primary	3
2.2.2 Secondary	3
2.3 Study Objectives	3
2.3.1 Primary	4
2.3.2 Secondary	4
2.4 Overview of Sample Size Considerations	4
2.5 Overview of Formal Interim Monitoring	5
3 OUTCOME MEASURES	5
3.1 Primary Outcome Measures	5
3.2 Secondary Outcome Measures	7
3.3 Other Outcome Measures	8
4 STATISTICAL PRINCIPLES	9
4.1 General Considerations	9
4.2 Analysis Approaches	9
4.2.1 Primary Analysis.....	9
4.2.2 Secondary Analyses	10
4.2.3 Other Analyses.....	10
5 REPORT CONTENTS	11
6 APPENDIX: NEUROPSYCHOLOGICAL NORMS FOR ACTG A5324	12
7 APPENDIX: IADL SCORING	14
8 ASSOCIATED DOCUMENTS	ERROR! BOOKMARK NOT DEFINED.

1 Introduction

1.1 Purpose

This Primary Statistical Analysis Plan (SAP) describes the primary and secondary outcome measures and additional outcome measures of the A5324 study that will be included in the primary manuscript, and which address, at a minimum, the major primary and secondary objectives of the study. The Primary SAP outlines the general statistical approaches that will be used in the analysis of the study. It has been developed to facilitate discussion of the statistical analysis components among the study team, and to provide agreement between the study team and statisticians regarding the statistical analyses to be performed and presented in the Primary Analysis Report. It also describes the results for the primary and secondary outcome measures that will be posted on ClinicalTrials.gov.

Detailed outlines of tables, figures, and coding descriptions that will be included in the Primary Analysis Report are included in the Analysis Implementation Plan (AIP).

Analyses for the Primary Analysis Report will be finalized once the last participant has completed the week 96 study visit, all queries have been resolved, and the study database closure/data lock has been completed.

Outlines of analyses for other objectives and outcome measures not included in the Primary SAP will be provided in a separate SAP.

1.2 Version History

Major changes to the previous version are noted in **bold face** throughout the document.

The SAP has been updated prior to final analysis and based on discussions with the study team.

Version	Changes Made	Date Finalized
1	Original Version	05/19/2017
2	Update for Protocol v2.0, LOA #3	03/28/2019
3	Updated for final analysis and LOA #4: Updated the viremia outcome to exploratory, added a BDI-related exploratory outcome, updated the functional status analysis, added missing data handling, and added appendices describing neurocognitive test scoring procedures and IADL scoring procedures; clarifications throughout	02/12/2021

2 Study Overview

2.1 Study Design

This is a phase IV randomized, double-blinded, placebo-controlled study to assess the efficacy of adding Maraviroc (MVC) and Dolutegravir (DTG) to the current antiretroviral therapy (ART) of HIV-infected individuals with undetectable (<50 copies/mL) plasma HIV-1 RNA, who have mild to moderate neurocognitive impairment, with a primary outcome of improvement in neurocognitive performance.

DURATION 96 weeks

SAMPLE SIZE 186 participants

POPULATION Participants will have HIV-associated neurocognitive disorder (HAND) as defined by the Frascati criteria, plasma HIV-1 RNA <50 copies/mL within 90 days prior to entry, and no more than one plasma HIV-1 RNA ≥ 50 and <200 copies/mL in the past 6 months prior to entry with a subsequent plasma HIV-1 RNA <50 copies/mL, and on stable ART for at least 6 months prior to entry with no plans to change treatment.

STRATIFICATION Participants will be randomized in a stratified manner. Stratification variables will be CD4+ nadir (≤ 100 vs. > 100 cells/mm³), and HAND severity (asymptomatic neurocognitive impairment [ANI] vs. mild neurocognitive disorder [MND] / HIV-associated dementia [HAD]).

REGIMEN At entry participants will be randomized to one of the following:
- Arm A: Add to their existing ART: placebo for MVC and placebo for DTG
- Arm B: Add to their existing ART: DTG and placebo for MVC
- Arm C: Add to their existing ART: MVC and DTG

2.2 Hypotheses

2.2.1 Primary

In HIV-1-infected participants with suppressed plasma viremia and at least mild neurocognitive impairment, improved neurocognition and increased functional capacity over 48 weeks will be seen with addition of Maraviroc (MVC) plus Dolutegravir (DTG) (Arm C) compared to DTG only (Arm B) and compared to placebos (Arm A). Addition of DTG only (Arm B) compared to placebos (Arm A) will result in improved neurocognition over 48 weeks and increased functional capacity.

2.2.2 Secondary

Changes in peripheral blood and cerebrospinal fluid (CSF) biomarkers will parallel improvement in neurocognitive performance following addition of DTG alone or MVC and DTG.

2.3 Study Objectives

Analysis of the study objectives below will be analyzed under a superiority framework, to conduct pairwise comparisons of the three arms.

2.3.1 Primary

1. To determine whether intensification with MVC and DTG (Arm C) will improve neurocognitive functioning at week 48 in participants who have at least mild neurocognitive impairment, are on a stable antiretroviral therapy (ART) regimen, and have plasma HIV-1 RNA <50 copies/mL, over DTG plus placebo for MVC (Arm B) and over placebo of both drugs (Arm A).
2. To determine whether DTG with placebo for MVC (Arm B) will improve neurocognitive functioning at week 48 over placebo of both drugs (Arm A).

2.3.2 Secondary

1. To determine the safety and tolerability of MVC and DTG when added to a stable ART regimen. The safety and tolerability measures will include time to discontinuation of any study medication due to adverse events (AEs), proportion of Grade 3 or Grade 4 clinical AEs, proportion of Grade 3 or Grade 4 laboratory abnormalities, and summaries of all treatment related AEs.
2. To determine whether intensification with MVC and DTG will improve neurocognitive functioning at weeks 24, 72, and 96.
3. To determine whether functional status improves at week 48 after intensification with MVC and DTG (Arm C) to a greater extent than when intensified with DTG and placebo for MVC (Arm B) or placebo of both drugs (Arm A).
4. To assess the effect of adding MVC and DTG on biomarkers in peripheral blood and CSF.
5. To assess the association between **plasma** and CSF biomarkers, neuropsychological performance, and DTG and MVC pharmacokinetics (PK) throughout the course of the study.
6. To assess whether the addition of DTG alone (Arm B) or DTG and MVC (Arm C) is associated with maintenance of plasma HIV-1 RNA <50 copies/mL.
7. To assess changes in blood CD4+/CD8+ counts after intensification with DTG and MVC (Arm C), DTG alone (Arm B) or placebos (Arm A).
8. To determine associations between peripheral blood immunologic markers (s100 β , soluble CD14) and changes in neurocognitive function at week 48.
9. To determine the relationship between changes in CCR5 ligand levels (MIP-1b) and inflammatory marker changes in both peripheral blood and CSF.
10. To determine the effect of DTG intensification on residual viremia, cellular HIV-1 RNA expression, total cellular HIV-1 DNA, and 2 long terminal repeat sequences (LTR) DNA, and to determine whether these virologic changes are associated with week 48 neurocognitive function.

2.4 Overview of Sample Size Considerations

Based on previous studies, a standard deviation (SD) of 0.7 in the 48-week change in normalized neurocognitive test score is assumed. In the sample size calculation, the null hypothesis is that there is no difference in changes of neurocognitive functioning from baseline to week 48 across the treatment arms, while the alternative hypothesis is that the change of neurocognitive functioning is at least 0.5 higher for the (MVC and DTG) arm. A Bonferroni correction has been applied to account for multiple comparisons between treatments. One interim analysis of efficacy

is planned in the middle of the study and the adjusted sample sizes assume a 15% loss to follow-up. To achieve a power of 90%, we would expect a sample size of 62 participants per group for a total 186 participants on study. See Protocol section 9.4 for further details.

2.5 Overview of Formal Interim Monitoring

Routine SMC reviews by the Neuro SMC assess the study progress via reviewing accrual, study status, screening failures, and toxicity at least annually. In the middle of the study, when 50% of participants are enrolled and followed for 48 weeks, an efficacy interim review will occur. This review will consist of a review of the assumed standard deviation (SD=0.7) and pairwise comparisons between the three blinded treatment arms. The pairwise comparisons will be made using repeated confidence intervals and Bonferroni correction, and will use a threshold p-value of 0.0007 for each comparison to assess significance. See Protocol section 9.5 for details.

3 Outcome Measures

3.1 Primary Outcome Measures

1. The change in normalized composite neurocognitive test score at week 48 from pre-entry. This test score is further defined below.
 - Aligns with Primary Objectives 1 & 2

The normalized neurocognitive test score for domestic participants is defined as the average of z-scores on the following tests:

- Stroop color naming [English: NE7041, Spanish: NE7101]
- Stroop word reading [English: NE7041, Spanish: NE7101]
- Stroop interference trial [English: NE7041, Spanish: NE7101]
- Grooved pegboard dominant [NE7046]
- Grooved pegboard non-dominant [NE7046]
- HVLTR Learning trials [NE7097]
- HVLTR Delayed recall [NE7097]
- HVLTR Delayed recognition [NE7097]
- Semantic verbal fluency [NE7098]
- Letter fluency [NE7098]
- Trail Making A [NE7099]
- Trail Making B [NE7099]
- WAIS-III Symbol search [NE7101]
- Digit Symbol [NLW0016]

The WRAT-4 word reading total [English: NLW0018] (WAT Spanish Word List word reading total [Spanish: NLW0020]) was also administered at domestic sites but will not be included in the average z-score calculation and instead will be analyzed separately.

The normalized neurocognitive test score for international participants is defined as the average of z-scores on the following tests:

- Grooved pegboard dominant [NE7046]
- Grooved pegboard non-dominant [NE7046]
- HVLTR Learning trials [NE7097]

- HVLT-R Delayed recall [NE7097]
- HVLT-R Delayed recognition [NE7097]
- Semantic verbal fluency [NE7098]
- Digit Symbol [NLW0016]
- Timed Gait [NE7090]
- Finger Tapping Dominant [NE7095]
- Finger Tapping Non-dominant [NE7095]
- Color Trail 1 [NE7096]
- Color Trail 2 [NE7096]

The International HIV Dementia Scale (IHDS) [NE7061] was also administered at international sites but will not be included in the average z-score calculation and instead will be analyzed separately.

For time point assessment of impairment (baseline, weeks 24, 48, 72, 96), z-scores will be calculated by subtracting a demographically appropriate (adjusted for age, race, sex, and years of education) norm and dividing by a demographically appropriate standard deviation. Z-score calculations are described in detail in Section 0. The z-scores will be calculated using the raw scores from the neurocognitive tests.

Assessment of change over time in neurocognitive impairment will be calculated by subtracting the average z-scores of the two time points of interest.

The individual scores are separated into six component domains (**plus the two separate tests WRAT and IHDS mentioned above**). The z-score for each domain is the average of standardized individual NP z-scores listed under each domain:

- Fine motor Skills/Complex Perceptual
 - Grooved pegboard dominant – both domestic and international
 - Grooved pegboard non-dominant – both domestic and international
 - Finger Tapping dominant – international
 - Finger Tapping non-dominant – international
 - Timed Gait - international
- Speed of Information Processing
 - Digit Symbol – both domestic and international
 - Stroop color naming – domestic
 - Stroop word reading – domestic
- Executive Function
 - Trail Making B – domestic
 - Stroop interference trial – domestic
 - Letter Fluency – both domestic and international
 - Semantic Verbal Fluency – both domestic and international
 - Color Trails 2 – international
- Verbal Learning
 - HVLT-R Learning Trial – both domestic and international
- Verbal Memory
 - HVLT-R Delayed recall – both domestic and international

- HVLIT-R Delayed recognition – both domestic and international
- Attention/Working Memory
 - Trail Making A – domestic
 - WAIS-III symbol search – domestic
 - Color Trails 1 – international
- IHDS – international
- Premorbid Skills
 - WRAT-4 Reading / WAT Spanish Word List – domestic

Based on the z-scores, domain deficit scores (DDS) will be computed for the domains, and the global deficit score (GDS) will be computed using the total battery. The GDS emphasizes impairment, reduces minimal learning/practice and has the advantage of reducing ‘sum to zero’ effects seen with counterbalancing positive and negative scores in a subject’s assessment. However, this approach does minimize the ability to observe subtle changes as scores are grouped into larger categories.

DDS and GDS are derived from the standardized scores by the following algorithm:

- If z-score is greater than -1.0, then DS is 0;
- If z-score is [-1.5, -1.0] then DS is 1;
- If z-score is [-2.0, -1.5] then DS is 2;
- If z-score is [-2.5, -2.0] then DS is 3;
- If z-score is [-3.0, -2.5] then DS is 4;
- If z-score is smaller than -3.0 then DS is 5;
- If z-score is missing then DS is missing;

The DS categories are defined as:

- 0 – Normal
- 1 – Mild
- 2 – Mild to Moderate
- 3 – Moderate
- 4 – Moderate to Severe
- 5 – Severe

3.2 Secondary Outcome Measures

1. Treatment related adverse events
 - Aligns with Secondary Objective 1
 - **Time to discontinuation of any study medication due to adverse events**
 - **Proportion of Grade 3 or Grade 4 clinical AEs**
 - **Proportion of Grade 3 or Grade 4 laboratory abnormalities**
 - **Summaries of all AEs where the site indicated some relationship to study treatment.**
2. Change from baseline of normalized composite neurocognitive test score at weeks 24, 72, and 96. Change from baseline of normalized component neurocognitive test scores at weeks 24, 48, 72, and 96 will also be examined.
 - Aligns with Secondary Objective 2
 - See primary outcome for details on composite neurocognitive test scores
 - Component neurocognitive test scores consist of domain scores

- Change from baseline in the domain deficit scores and GDS to weeks 24, 48, 72, 96
3. Change from baseline in functional status scores based on the IADLs form at weeks 24, 48, 72, and 96.
 - Aligns with Secondary Objective 3
 - **Functional change is defined as a change from baseline in IADL score. The IADL scoring is defined in Section 7.**
 4. Change from baseline of peripheral blood and CSF biomarkers at week 48.
 - Aligns with Secondary Objective 4
 - Potential list of biomarkers includes:
 - i. Blood: sCD14, MIP-1 β , TNFr-II, VCAM-1, NFL
 - ii. CSF: MIP-1 β , NFL, neopterin, IP-10
 5. Plasma HIV-1 RNA \geq 50 copies/mL at weeks 24, 48, and 96.
 - Aligns with Secondary Objective 6
 6. CD4+ T-cell counts and changes from baseline to weeks 24, 48, and 96.
 - Aligns with Secondary Objective 7
 7. CD8+ T-cell counts and changes from baseline to weeks 24, 48, and 96.
 - Aligns with Secondary Objective 7

3.3 Other Outcome Measures

1. Changes from baseline to week 48: residual viremia, cell-associated HIV-1 RNA/DNA/2-LTR circles and SCA, T-cell and monocyte activation.
 - Aligns with Secondary Objective 10
 - Looking specifically at the comparison between DTG only arm and placebo arm to determine the effect of DTG intensification on these factors
 - **NOTE: This outcome measure was listed as secondary in the protocol but the intention was as an exploratory outcome due to lack of funding for testing, specifically for T-cell and monocyte activation**
2. Beck Depression Inventory (BDI) evaluations
 - **New outcome measure**
 - **Change in BDI score (evaluated as the sum of the 21 items) from baseline to weeks 24, 48, 72, and 96**
 - **Change from baseline in BDI score by subscales:**
 - Cognitive (items 1, 2, 3, 5, 6, 7, 8, 9, and 14)
 - Affective (items 4, 10, 12, and 13)
 - Somatic (items 11, 15, 16, 17, 18, 19, 20, and 21)
 - **Percent of participants with a BDI score \leq 13 vs. $>$ 13 at weeks 24, 48, 72, and 96**
3. **Associations of neuropsychological performance with plasma and CSF biomarkers and at weeks 0, 24, 48, 72, and 96 (0 and 48 only for CSF biomarkers); and with DTG and MVC PK at weeks 24 and 48.**
 - **Aligns with Secondary Objectives 5, 8, 9**

4 Statistical Principles

4.1 General Considerations

- The primary analysis and secondary analyses will use the ITT principle (i.e. all randomized participants will be included in the primary analyses).
- The primary analyses will use “as randomized” for group assignment.
- **The missing values for neurocognitive test scores will be imputed at the test level in the following manner:**
 - **When missing values can be attributed to neurologic impairment, the worst 1% score will be imputed for the missing values. Note: If the participant was unable to perform the test due to a disability related to HIV-associated neurological disease (e.g., unable to follow or understand directions, severe impairment in motor controls) this was indicated on the appropriate neuropsychological examination form.**
 - **If the reason for missingness is unknown and observations exist both before and after the missing value, these scores will be averaged and that averaged value will be imputed for the missing value.**
 - **If the reason for missingness is unknown and observations do not exist after the missing value, scores will be imputed by the last observation carried forward method.**
- **Missing data prevalence (of visits and of neurocognitive test scores) will be summarized.**
- **For the neurocognitive test scores, Winsorized estimates may be used for extreme outliers after observing the data.**
- All programs that create permanent study-specific derived datasets or that contribute to analyses of the primary or secondary objectives will be validated according to CBAR SOPs. Primary and key secondary objective results will be verified through independent double programming.
- **Visit and analysis window details:**
 - **Visits occurred at screening, entry, and weeks 2(±1), 4(±1), 12(±2), 24(±2), 48(-4,+8), 72(-4,+8), and 96(-4,+8).**
 - **Analysis windows will be formed around each study visit using the midpoints between adjacent visits as cutoffs and potentially including assessments collected outside the recommended visit windows described in the protocol.**

4.2 Analysis Approaches

4.2.1 Primary Analysis

The Kolmogorov-Smirnov test will be employed to check whether the primary outcome is normally distributed. If the normality assumption is not rejected, the two-sample t-test will be used for pairwise comparisons between any two arms. Otherwise, the nonparametric analogue of the two-sample t-test (ie, Wilcoxon rank-sum test) will be used as an alternative. The test will be stratified by CD4+ nadir (≤ 100 vs. > 100 cells/mm³) and HAND severity (ANI vs. MND/HAD). All pairwise comparisons will be conducted among the three study arms.

Treatment comparisons will be summarized by sex and by race. Treatment interactions with sex and race will be tested as required by NIH using multiple variable regression analyses. Subgroup analyses will be performed for the primary outcome in each participating country. In presence of significantly different results by country, the changes of neurocognitive functioning will be compared by adjusting for country effects **by Van Elteren tests**.

Sensitivity analyses will be conducted under different assumptions (e.g., impute missing z-scores of a test at week 48 by the smallest/largest/median of observed values in the same treatment arm) to assess the impact and informativeness of missing data. A per-protocol analysis which will restrict the analysis to participants who adhere perfectly to the instructions in the study protocol will be performed. **In addition, a sensitivity analysis of the effect of the Trails B score imputation (see Section 6 Appendix) will be reviewed by setting imputed scores to incomplete and then re-evaluating the primary outcome.**

4.2.2 Secondary Analyses

1. Treatment related AEs: Time to discontinuation of medication due to AEs will be compared by treatment arm using the Kaplan-Meier approach and log-rank test; Proportion of participants with Grade 3 or Grade 4 clinical AEs or laboratory abnormalities will be compared by treatment arm, using Fisher's exact tests.
2. Change in neurocognitive scores: See 4.2.1 Primary Analysis above.
3. Change in functional status: Pairwise comparisons will be performed among three treatment arms using the Wilcoxon rank-sum test.
4. Change in biomarkers: The Wilcoxon signed-rank test will be used to assess the changes from baseline to week 48 in specified biomarkers. Comparisons of changes in these biomarkers between any two treatment arms will be conducted through Wilcoxon rank-sum test.
5. Plasma HIV-1 RNA: For each treatment arm, the proportion of participants with ≥ 50 copies/mL and associated 95% CI will be provided in a table and plotted over time; Two-sided 95% CIs will also be provided for the difference between arms.
6. Changes in CD4+ cell counts: Average CD4+ counts and their associated 95% CIs will be provided and plotted over time for each treatment arm; changes in CD4+ counts from baseline to weeks 24, 48, and 96 will be compared pairwise among treatment arms through Wilcoxon rank-sum test.
7. Changes in CD8+ cell counts: Average CD8+ counts and their associated 95% CIs will be provided in a table and plotted over time for each treatment arm; changes in CD8+ counts from baseline to weeks 24, 48, and 96 will be compared pairwise among treatment arms through Wilcoxon rank-sum test.

4.2.3 Other Analyses

1. Changes in virologic measures: Wilcoxon rank-sum test will be used to compare Arm A and Arm B to assess the effect of DTG intensification on virologic changes.
2. **BDI Analyses: Change will be compared pairwise among the three treatment arms using the Wilcoxon rank-sum test. Percent of participants with BDI score >13 vs ≤ 13 will be compared by arm using the Chi-Square test.**

3. Correlation Analyses: Spearman correlations and scatterplots will be used to assess associations between:
 - a. Peripheral blood and CSF biomarkers and neuropsychological performance at weeks 0, 24, 48, 72, 96 (weeks 0 and 48 only for CSF biomarkers).
 - b. Changes in blood immunologic markers and changes in neurocognitive performance from baseline to week 48.
 - c. Changes in blood biomarkers from baseline to week 48.
 - d. **DTG and MVC PK and neuropsychological performance at weeks 24 and 48.**

5 Report Contents

Detailed descriptions of the content for each of the following sections are given in the AIP.

- CONSORT Diagram
- Accrual – summarize enrollment
- Baseline characteristics
 - Age
 - Sex
 - Race
 - IV drug use
 - Country
 - Education
 - Primary Language
 - HAND diagnosis
 - Nadir CD4 count
 - CD4 count
 - Log10 HIV-1 RNA
 - Neurocognitive test score
- Study status
- Treatment status
 - Premature study treatment discontinuation reasons
- Study conduct
 - Missed visits
- Safety Summary
- Analysis of primary outcome measures
 - Subgroup analyses
 - Sensitivity analyses
- Analysis of secondary outcome measures
- Analysis of other outcome measures

6 Appendix: Neuropsychological Norms For ACTG A5324

Domestic Sites

Trail Making A & B, Grooved Pegboard, Semantic and Letter Fluency

Norm source: Heaton RK, Miller SW, Taylor MJ, Grant I. (2004). Revised comprehensive norms for an expanded Halstead-Reitan battery: demographically adjusted neuropsychological norms for African American and Caucasian adults. Lutz, FL: American Psychological Assessment Resource, Inc.

- T scores were generated using the manual look-up tables.
- T scores were then converted to z-scores.
- Norms for African Americans were used for Hispanic and Native American participants.
- Trails B used a cut-off time of 240 seconds instead of the expected time out score of 300 seconds. Scores ≥ 240 are imputed to 300 for calculating the T score (unless a comment from the site provides an exact time).
- Grooved Pegboard did not use a cut-off time during testing. For analysis, all scores >300 are set to 300 for calculating the T score.

Stroop (Commali version)

Norm source: Mitrushina M, Boone KB, Razani J, D'Elia L. (2005). Handbook of normative data for neuropsychological assessment. 2nd ed. New York: Oxford University Press.

- Tables A6.7.-A6.10., pages 666-668 in Mitrushina manual.
- Note- the tables were originally published by Demick and Harkins, 1997.
- Raw scores were converted to z-scores.

WAIS-III Symbol Search and Digit Symbol

Norm source: Weschler, D. (1997). Wechsler Adult Intelligence Scale, Third Edition.

- Scaled scores were defined using the manual look-up table.
- Scaled scores were converted into T scores.
- T scores were then converted to z-scores.

Hopkins Verbal Learning Test-Revised

Norm source: Brandt, J., & Benedict, R. H. B. (2001). Hopkins Verbal Learning Test – Revised. Administration manual. Lutz, FL: Psychological Assessment Resources.

- Raw scores were converted into T scores based on the manual look-up tables.
- T scores were then converted to z-scores.
- Raw scores below the range of T scores in the look-up tables were converted to z-scores using the mean and SD for the reference group as provided at the bottom of each look-up table in the manual.

Wide Range Achievement Test-4 (Green form)

Norm source: Wilkinson, G. S., & Robertson, G. J. (2006). Wide Range Achievement Test: Fourth Edition. Lutz, FL: Psychological Assessment Resources.

- Raw scores were converted into standard scores using the look-up table.
- Standard scores were then converted into z-scores.

Word Accentuation Test

Norm source: Gomar, J. J., Ortiz-Gil, J., McKenna, P. J., Salvador, R., Sans-Sansa, B., Sarró, S., Guerrero, A., & Pomarol-Clotet, E. (2011). Validation of the Word Accentuation Test (TAP) as a means of estimating premorbid IQ in Spanish speakers. *Schizophrenia research*, 128(1-3), 175–176.

- Raw scores were converted into standard scores.
- Standard scores were then converted into z-scores.

International Sites

- Raw scores were converted into z-scores using neurocognitive normative comparison test data from ACTG A5271.

7 Appendix: IADL Scoring

IADL scoring uses the following questions from NRQ0004, and scores them according to a binary scoring system using the 'NOW' response. The IADL score is the sum of the scores as defined below:

1. Housekeeping

- a. I maintain my house/apartment by myself or only need occasional help for larger jobs.....1
- b. I only perform light daily tasks (wash dishes, make bed).....0
- c. I perform some light tasks, but have difficulty keeping my place clean.0
- d. I need help with all housekeeping tasks.....0
- e. I am fully able to do housekeeping, but choose not to do so.....1

2. Managing Monies

- a. I manage all of my monies by myself.....1
- b. I manage routine small purchases, but need help with other checking and balancing accounts.....0
- c. I am not able to handle money accurately.....0
- d. I am able to handle my own monies, but someone else does them for me.....1

4. Cooking

- a. I plan, prepare and serve many of my own meals.....1
- b. I prepare meals if someone else provides me with the right ingredients.....0
- c. I heat and serve meals provided by others.....0
- d. I need to have meals prepared and served to me.....0
- e. I am able to plan, prepare, and serve my own meals but someone else does it for me.....1

7. Transportation

- a. I drive my own car or take public transportation on my own (if you do not own a car).....1
- b. I arrange my own travel using taxis, but do not drive or use public transportation.....0
- c. I can travel on public transportation or use taxis if I am assisted by another.....0
- d. I am entirely dependent on others to take me where I need to go.....0

8. Using the telephone

- a. I handle using the telephone without difficulty (looking up and dialing new numbers, etc.).....1
- b. I only dial a few well-known numbers.....0
- c. I answer the telephone, but do not dial.....0
- d. I do not use the telephone at all.....0
- e. I do not have access to a telephone.....N/A

12. Shopping (e.g., clothes, other non-food goods)

- a. I take care of all of my shopping needs.....1
- b. I only make small purchases.....0
- c. I need someone to go with me on any shopping trip.....0
- d. I am unable to shop.....0

e. I am able to shop, but choose to have someone else do my shopping for me.....1

13. Laundry

a. I do all of my own laundry.....1
b. I need occasional help in doing the laundry.....0
c. I launder only small items (e.g., rinse socks, stockings, etc.).....0
d. All laundry must be done by others.....0
e. I am able to do my own laundry, but choose to have others do it for me.....1

14. Taking/keeping track of medication

a. I can take my medication without help.....1
b. I take medications that are prepared by someone else.....0
c. I am unable to track my own medications.....0
d. I am able to take care of my own medications, but choose to have someone else do it for me.....1