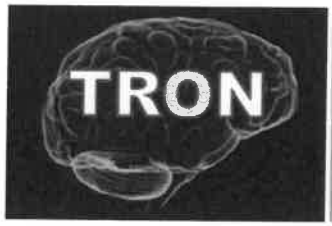
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

sewtu
South East Wales Trials Unit
Uned Ymchwil De-ddwyrain Cymru



Study of Everolimus in theTreatment of Neurocognitive problems in Tuberous Sclerosis

STATISTICAL ANALYSIS PLAN
FOR
TRON (Main study)

TRIAL TITLE:

A randomised, double blind, placebo-controlled study of RAD001 (Everolimus) in the treatment of neurocognitive problems in tuberous sclerosis.

EudraCT number:	2011-004854-25
Sponsor reference:	SPON803-10
ISRCTN number:	ISRCTN09739757

Date: 22 June 2018

A randomised, double blind, placebo-controlled study of RAD001 (Everolimus) in the treatment of neurocognitive problems in tuberous sclerosis (TRON)

Chief Investigator:	Professor Julian Sampson
Co-Investigators:	Dr Petrus de Vries, Dr D Mark Davies, Professor Kerenza Hood, Dr Rachel McNamara, Dr Anurag Saxena, Dr Michelle Smalley
Trial Manager:	Dr Cheney Drew
Data Manager:	Lianna Angel, Tori Drane
Trial Statistician:	Dr Rebecca Cannings-John
	Tim Pickles

R. CANNINGS-JOHN

Author's Name

22/06/2018

Date and Signature

K. Hood

Senior Statistical Advisor

5/7/18

Date and Signature

J SAMPSON

Chief Investigator

5/7/18

Date and Signature

Table of Contents

1. Acronyms and definition of terms

Acronym	Meaning
CACE	Complier Adjusted Causal Effect
IDMC	Independent Data Monitoring Committee
HEAP	Health Economics Analysis Plan
ICH	International Conference on Harmonisation
ITT	Intention to Treat
PI	Principle Investigator
PPA	Per – Protocol Analysis
SA	Safety Analysis
SAP	Statistical Analysis Plan
SEWTU	South East Wales Trials Unit
SOP	Standard Operating Procedure
TSC	Tuberous sclerosis

2. Statistical analysis plan authorship

Rebecca Cannings-John is the Trial Statistician for TRON and the authors of this statistical analysis plan (SAP). Statistical analysis will be carried out by Rebecca Cannings-John under the supervision of Professor Kerry Hood (SEWTU director). This SAP will be finalised for presentation to the Trial Management Group and will be agreed by them and signed off by one author, a senior statistician and the Chief Investigator. A copy will then be sent to the independent statistician from the Trial Steering Committee.

This statistical analysis plan has been developed in compliance with 'Statistical Principles for Clinical Trials' (ICH E9)[i], 'Guidance for Good Clinical Practice' (ICH E6)[ii], 'Structure and Contents for Clinical Study Reporting' (ICH E3)[iii], 'Standard Operating Procedure (SOP) for Statistical Analysis'[iv] and the TRON study protocol[v]. Any amendment to this plan after the commencement of the analysis should be documented in the log provided in the appendix.

3. Introduction to the study

RAD001 (Everolimus) is a selective mTOR (mammalian target of rapamycin) inhibitor, specifically targeting the mTOR-raptor signal transduction complex (mTORC1). mTOR is a key serine-threonine kinase in the PI3K/AKT/mTOR signaling cascade, a pathway known to be dysregulated in the majority of human cancers and in tuberous sclerosis (TSC). Everolimus exerts its activity through high affinity interaction with the intracellular receptor protein FKBP12. The FKBP12/Everolimus complex binds to mTORC1, inhibiting its signaling capacity. mTORC1 signaling is effected through modulation of the phosphorylation of downstream effectors, the best characterised of which are the translational regulators S6 ribosomal protein kinase (S6K1) and eukaryotic elongation factor 4E-binding protein (4E-BP). Inhibition of mTORC1 by RAD001 (Everolimus) has effects on cell growth and proliferation accounting for the anti-tumour and immunosuppressive properties of the drug. mTORC1 signaling also plays important roles in learning and memory in experimental animals but the effects of mTORC1 inhibition on neurocognition in humans are unknown.

Individuals with TSC have a very high frequency of neurocognitive and neuro-developmental problems. These include: profound intellectual disability in 30% and mild to severe intellectual disability in a further 20%; autism spectrum disorder in 40 - 50%; attention deficit hyperactivity disorder in 30-50% (approximately 10 times higher than population expectations). The majority of individuals, even those with normal intellectual ability, have specific neuropsychological deficits of attention, executive or memory skills. Currently there are no approved agents for the treatment of these neurocognitive and neurodevelopmental deficits in TSC. It is therefore clear that there is an enormous unmet need in the TS population for effective treatments for these problems.

Because of the established roles of TSC1 and TSC2 in regulation of mTORC1 it has been hypothesised that over-activity of mTORC1 signalling in the brain is an aetiological factor in the neurocognitive deficits in patients with TSC, but there is no experimental evidence for this in humans. In preclinical studies, constitutional heterozygous knockout mouse models of TSC show deficits in memory and learning that have been corrected by the mTOR inhibitor rapamycin and conditional knockout of both TSC1 or TSC2 alleles in the mouse brain results in a severe neurological phenotype with brain overgrowth, disorganisation, seizures and early death that is improved dramatically by mTOR inhibition.

The current trial, TRON, will investigate RAD001 (Everolimus) in relation to neurocognitive function in adult patients with TSC. It will determine effect sizes to provide evidence for whether

larger trials for this indication are appropriate. It will focus on adult patients who do not have active seizures or who have well controlled seizures.

3.1 Study objectives

This study is designed to determine the effect sizes of treatment with RAD001 (Everolimus) or placebo for 6 months on specific neurocognitive functions - recall memory (verbal and non-verbal) and executive function - in people affected by TSC who have significant deficits in these functions. These data will provide new evidence for whether subsequent larger scale trials are indicated, and will be needed for their design.

Primary objective(s)

- To determine the effect sizes of treatment with RAD001 (Everolimus) or placebo for 6 months on recall memory and executive function in people with tuberous sclerosis.

Secondary objective(s)

- To assess the effects of treatment with RAD001 (Everolimus) or placebo for 6 months weeks on wider aspects of neurocognitive functioning, seizures and daily life in people with tuberous sclerosis and to assess safety using the NCI CTCAE version 4.0.

Exploratory Objective(s)

- To determine whether an effect of treatment with RAD001 (Everolimus) or placebo is detectible at 1 month and 3 months after starting therapy to establish whether any early markers of change are present.

4. Study design

In order to determine effect sizes and to distinguish drug effects from placebo or practice effects this trial will be a two-arm, individually randomised Phase II, double- blind, placebo-controlled trial of RAD001 (Everolimus) versus Placebo. Randomisation will be 2:1 in favour of RAD001 (Everolimus) to increase data available for effect-size estimation. Everolimus is a licensed medicine within this patient group, but for a different indication. This trial is therefore a proof of principle study for memory and executive cognitive function. The flow chart indicating patient numbers and anticipated drop out is shown in Appendix 1.

4.1 Sample size justification

The sample size has been calculated based on Fleming's single stage procedure[vi] and considers the sample size required to assess the change in the Everolimus group. The control group provides evidence of the learning effect from repeated assessments. The proposed study seeks to determine effect sizes in relation to the primary outcome measures to enable the informed design of subsequent larger studies. This study is therefore not powered for formal statistical comparison of placebo and study drug groups.

We estimate the learning effect (proportion of patients whose memory functioning improved due to familiarity with the assessments) to be slightly less than 0.15. A proportion of less than 0.15 in the intervention group would therefore indicate that the intervention (RAD001/Everolimus in the treatment of neurocognitive problems in tuberous sclerosis) does not warrant further investigation. A proportion of at least 0.35 would provide sufficient evidence for further investigation of this intervention. Values between would be discussed in depth by members of the trial team.

Therefore, to test the null hypothesis that the proportion of patients in the intervention group who improve their memory functioning by one SD is at most 0.15 against the alternative hypothesis that the proportion of patients in the intervention group who improve their memory functioning by one SD is at least 0.35, with 80% power and a one-sided α of 0.05, we require a total sample size of 48 (i.e. 32 interventions and 16 controls). This sample size is inflated to allow for 20% loss to follow-up.

4.2 Randomisation

Minimisation will be carried out by the data manager at the the South East Wales Trials Unit (SEWTU). Allocation ratio will be 2:1 in favour of Everolimus. The algorithm will allocate new participants to the study arm that minimises imbalance with respect to these factors with probability of 0.8. Therefore a random element is retained, further reducing predictability of allocation. Full details of the randomisation are contained in the randomisation protocol[vii].

5. Study outcomes

The primary and secondary outcomes are listed below with full details of the tests in Appendix 2 and 3.

Neuropsychological tests and questionnaires:

Primary outcomes

- BIRT Memory and Information Processing Battery (B-MIPB) - List Learning test
- BIRT Memory and Information Processing Battery - Complex Figure test
- CANTAB - Stockings of Cambridge (SOC)
- CANTAB - Spatial Working Memory (SWM)
- Telephone search dual task (from the Test of Everyday Attention)

Secondary outcomes

- CANTAB - Rapid Visual Information Processing Battery (RVIP)
- CANTAB - Spatial Span (SSP)
- CANTAB - Attentional Set-shifting (IDED)
- Verbal Fluency /Controlled Oral Word Association Test (COWAT)
- Cancellation task
- Symptom Checklist 90R (SCL-90R)
- Quality of Life in Epilepsy (QOLIE)
- Liverpool Seizure Severity Scale (LSSS)
- Vineland Adaptive Behavior Scales-II (VABS-II) (Parent/Caregiver Rating Form)
- Social Responsiveness Scale – Adult version (SRS-A)
- Social communication questionnaire (SCQ)

5.1 Data collection and handling

The data for the study will be collected on CRFs by both the Clinical Study Co-ordinator and Psychology assistant. Each CRF will contain the appropriate tests for both the clinical and psychological assessments as stated in the Trial protocol. The Trial Specific SOP on CRF handling, details who is responsible for the safe keeping of the CRFs. When the CRFs reach SEWTU they will be locked in a safe cabinet until they are ready to be entered. Only persons named on the Trials PRA will have access to the CRFs. They will be entered into the database by firstly being scanned and verified via the data entry system Cardiff Teleform. Once they have been verified the data will be 100% QC'd for accuracy and any changes that are required documented. The data will then be uploaded into the Trials SQL database. Any changes that are required to the

data will be made through SQL as the database has an audit function which will show what changes have been made to the data which is required under ICH GCP.

5.2 Definitions/Calculations

The scales used in the trial are identified in Appendix 2 and 3. Validated questionnaire outcome scales will be calculated according to original coding systems in published papers. For the majority of tests these are calculated electronically but the LSS and SRS-A are completed by hand and calculated by the psychology assistant.

5.3 Analysis population

Primary and secondary comparative analyses will be conducted on an intention-to-treat (ITT) basis without imputation of missing values (complete case). The ITT population uses all randomised participants in the groups they were randomised to regardless of the intervention received and the complete case restricts the population to include all randomised participants with complete follow-up (specific to the outcome and time points). ITT analysis with imputed missing values will be undertaken as a sensitivity analysis.

In a per-protocol or clinically evaluable population, patients are analysed according to the treatment they were randomised to and received at least one dose of study drug.

5.4 Missing data

If a patient attends the 6 month visit (primary end point) it is likely that the individual tests making up the primary outcome will be completed since the tests are done as a battery of tests. In the event that the tests are stopped and partial tests are completed or where a patient does not attend the 6 month session, cases missing follow up will be excluded from the primary analysis and complete case analysis used. A sensitivity analysis will be undertaken to determine the effect of missing of data on the primary outcome. In the case that missing data occurs from individual tests, items will be handled according to the scoring system used for outcome measurement. A comparison of baseline data for complete cases vs. those with missing outcome data will be undertaken to quantify any likely bias.

5.5 Outliers

Any unusual outcome measurements will be automatically flagged on scanning and checked for scanning error. Any outlier values remaining after data cleaning and checking will be investigated for authenticity. The influence of these outlier values on the primary analysis will be checked. Any significant influence detected will be reported and discussed with the TSC and DMC.

5.6 Analysis Time Frame

Randomisation will occur as soon as the patient receives their second visit and will be followed up for a period of 36 weeks (9 months post-treatment). Follow-up will continue until the last recruited patient is seen at 36 weeks. Syntax for questionnaire outcome scoring, primary and secondary analyses will be written prior to the analysis phase. Descriptive analyses on a whole-group basis will be carried out at the end of each visit of data collection and when each dataset has been closed. Comparative analyses will only commence once the 6 month visit of data collection (24 weeks) has been completed and the database has been closed and exported from the SQL database.

Timings for data management and analysis are outlined in the Gantt (Appendix 5) and are monitored throughout the trial. Any changes to the time devoted to analysis and are to be agreed by the TMG and TSC. The DMC will meet once a year during the study and will receive a report on trial progress prior to the meeting.

6. Statistical analyses

6.1 Descriptive analysis

Participant flow and recruitment
Summary statistics on eligibility, recruitment, withdrawal and dropout will be collated for both trials arms and will form the basis of the CONSORT flow diagram for clinical trial reporting[viii]. Specifically, for each arm, participants randomly assigned, receiving intended intervention, completing the study protocol, and analysed for the primary outcome. Any protocol deviations will be described with reasons as well as dates defining the period of recruitment and follow-up.

Baseline visit (visit 2: week 0)
Appropriate descriptive summaries and graphical illustrations of baseline data and will be presented as means and standard deviations for normally distributed variables, medians and interquartile ranges for skewed, or numbers and percentages. Baseline data will be used to check comparability between study arms and generalisability of the study population. There will be no formal testing of between-arm differences for any variables at baseline.

Study medication
Treatment will be dispensed at visits 2, 4 and 6 and patients will take the medication at the study site at visits 3 to 7 and at home on all other treatment days. Compliance will be monitored by pill counts at each visit.

Blood samples for drug levels (pharmacokinetic assessments) will be taken at visits 3 to 7 inclusive. We will describe and graphically examine the dose of concentration of study drug (ng/ml) given over time and determine trough levels. We will also examine the sham levels in the control group to assess the level of blinding maintained.

6.2 Analysis of primary outcome

Data will be presented descriptively by trial arm at baseline and 6 months, and effect sizes determined. The control group will be recruited for comparison but will not be included in the primary analysis. This will be a 2 stage process.

(1) The proportion of patients (alongside 95% CI) in the control group displaying improved functioning will be reported to highlight the learning effect. An improvement will be defined as at least 1SD (using population norms) response in six month change on the percentiles/standard scores of ANY of the tests used for screening/baseline assessment (table 1). Once the learning effect has been determined, and if different from the pre-hypothesised 0.15, we will discuss with members of the trial team whether the pre-hypothesised improvement of 0.35 in the intervention group is appropriate.

(2) A one sample chi-squared (or goodness-of-fit) test will be used to determine whether the proportion of patients (alongside 95% CI) in the intervention group who improve their recall memory and executive functioning at six months by at least 1 standard deviation (SD) is at least 0.35 (or using the new cut-off if changed).

Using the pre-hypothesised cut-off values:

A proportion < 0.15 in the intervention group (equivalent to that expected in the control group) would indicate that the intervention does not warrant further investigation. A proportion of at least 0.35 would provide sufficient evidence for further investigation of this intervention. Values between 0.15 and 0.35 would be discussed in depth by members of the trial team. We will also describe the proportion of patients who go into deficit (> -1 SD) and are unchanged (<1 SD) at six months.

Table 1. Primary outcomes: Neuropsychological tests and questionnaires

Test		Raw score/count	Age-related percentile band	Standard score	Better than %	Good or better than %
B-MIPB[a]	List learning test - delayed recall - immediate recall	 ✓ ✓	 ✓ ✓			
B-MIPB[a]	Complex figure test - delayed recall - immediate recall	 ✓ ✓	 ✓ ✓			
CANTAB-SOC	Mean initial thinking time Mean subsequent thinking time Number of problem solved in minimum moves	✓ ✓ ✓		✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓
CANTAB-SWM[a]	Between-Errors Strategy fields	✓ ✓		✓ ✓	✓ ✓	✓ ✓
TEA[b]	Telephone search while counting	✓*	✓	✓**		

*** Dual task decrement; ** Scaled score*

[a] outcomes informed by animal research

[b] No TEA score if under 18 years as there are no norms for this test

6.3 Secondary analysis of primary outcome

Secondary analyses will consider the primary outcome measures as dimensions of memory functioning as continuous variables rather than combined into one outcome. For all analyses, we will graphically describe results, both on an individual patient level and also aggregated by trial arm.

Several secondary analyses are proposed:

1. *Six month follow up - continuous percentiles/standard scores*

For each individual test and by trial arm, we will present the mean (SD) scores at baseline and at six months and also the change from baseline to six months using the continuous scores. We will estimate the treatment effect between trial arms using an analysis of covariance (ANCOVA) model using response at six month as outcome with baseline response as a covariate. We will additionally adjust for the balancing factors used in the minimisation. The treatment effect will be estimated using the difference in adjusted means (intervention minus control arm) alongside 95% confidence intervals (CIs) and p-values. The distribution of each continuous outcome will be examined for normality prior to analysis. Where data are clearly non-normally distributed, appropriate transformations will be documented and justified in the model.

Further modelling will use multivariate analysis of variance (MANOVA) to model functioning over time between baseline and six months. MANOVA is used when there are two or more dependent variables as opposed to using several separate ANOVA models which will increase the type 1 error rate (rejecting the null hypothesis of an effect when there is no true effect). These analyses will compare the scores from each individual test at six months between the intervention and control group and will adjust for the balancing factors. The result will be presented as the (adjusted) difference in test outcomes between the intervention and control groups, along with 95% CI and p-value.

2. *Six month follow up - proportion of improved patients*

For each individual test we will also describe the proportion of patients with a deficit or an improvement (and those unchanged) at six months where deficit / improvement are defined as previously mentioned. The treatment effects of the proportion of patents improving their recall memory and executive functioning at six months will be examined using logistic regression models adjusted for the balancing factors, and result presented as adjusted odds ratios (ORs) alongside 95% CIs comparing the odds of improvement in the intervention compared with the control arm. We will also describe the number of tests a patient improves by at least one SD.

3. *Follow-up at one, three and six months - proportion of improved patients*

We will examine the proportion of patients with an improved recall memory and executive functioning at each time point either on ANY of the five tests and on the individual tests. A repeated measures logistic model will be used with the response outcome of deficit or not at each time point (one, three and six months), adjusted for the balancing factors, and including an interaction term for time and trial arm. This will allow an investigation of any divergent / convergent pattern in outcome over time, an estimation of treatment effect over all time points and at specific time points. The global interaction effect will be tested and where non-significant, the interaction and time term will be both dropped from the model. Both the intercepts and slopes of participants' measures will be allowed to vary randomly where possible. The Akaike Information Criterion (AIC) will be used to determine the best fitting model. We will also map out the individual deficits/improvements over time by examining change at each time point.

4. *Follow-up at one, three and six months - continuous percentiles/standard scores*

For each individual test and by trial arm, we will describe the mean (SD) scores over all time points (at one, three and six months). A repeated measures model (using a generalised linear mixed model) will be used to examine trends over time and estimate treatment effect as previously described. Further modelling will again use repeated measures MANOVA to model patterns of functioning over time at baseline, one, three and six months.

5. *Best response time point*

For each participant we will identify the time point (one, three, six, nine months) that gives the best response. Best response will be defined as the greatest improvement since baseline and examine the proportion with a best response at each time point by trial arm. Logistic or ordinal regression will be performed (depending on the distribution of response across time points) and effects presented as ORs alongside 95% CIs and p-value.

6.4 Analysis of secondary outcomes

Table 2 shows the secondary outcomes represent wider aspects of neurocognitive function, seizures, quality of life, and daily life functioning in patients with tuberous sclerosis.

Table 2. Secondary outcomes: Neuropsychological tests and questionnaires

Test	task	Raw scores	Age-related percentile band	Standard score	Better than %	Good or better than %	T score
Tests collected over baseline, 3, 6 and 9 months							
CANTAB - RVP	Raw correct score A1	✓		✓	✓	✓	
CANTAB - SSP	Span length	✓		✓	✓	✓	
CANTAB - IDED	No of stages completed	✓		✓	✓	✓	
	Total errors	✓		✓	✓	✓	
	Total no of errors adjusted	✓		✓	✓	✓	
COWAT- Cancellation task	Total no of targets cancelled	✓					
	Completion time (mm:ss)	✓					
	No of stars cancelled (left)	✓					
	No of stars cancelled (right)	✓					
COWAT-Verbal fluency	No of words produced (letter 1/2/3)	✓	Total				
QOLIE	For each of the 7 scales and overall score	✓	✓				✓
VABS-II	For each domain and overall	✓		scale score			
SC Questionnaire	Total score and provides cut-offs for autism referral	✓					
Tests collected over baseline, 1, 3, 6 and 9 months							
Symptom Checklist-90R	For each individual symptom dimension and global indices of distress	✓	✓				✓
Liverpool Seizure Scale	Total score	✓					
SRS-Adult	For each of the five treatment subscales and Total Score	✓					✓

Neurocognitive function
Quality of Life
Daily life
Mental health
Seizures

Note: For all scales where applicable, sub-scales will be calculated as well as the overall score.

Primary analysis

The primary analysis of the secondary outcomes will be analysed similarly to the primary outcome. The following secondary outcomes will be presented descriptively at baseline and six months and effect sizes determined. We will examine the proportion of patients improved over time (defined as at least one SD change in score) for each of the secondary outcomes. A one sample chi-squared test will be used to determine whether the proportion of patients in the intervention group who improve at six months by at least one SD is significantly greater than what is expected (the null hypothesis of equal proportions is rejected). The treatment effects of the proportion of patents improving their recall memory and executive functioning at six months will be examined using logistic regression models adjusted for the balancing factors, and result presented as adjusted odds ratios (ORs) alongside 95% CIs comparing the odds of improvement in the intervention compared with the control arm.

Secondary analyses

Secondary analyses of the secondary outcomes will again consider the measures as continuous variables rather than just the proportion of improved patients over time (between baseline and six months). We will estimate treatment effect between trial arms using an analysis of covariance (ANCOVA) model using response at six month as outcome with baseline response as a covariate. We will additionally adjust for the balancing factors used in the minimisation. The treatment effect will be estimated using the difference in adjusted means (intervention minus control arm) alongside 95% confidence intervals (CIs) and p-values. The distribution of each continuous outcome will be examined for normality prior to analysis. Where data are clearly non-normally distributed, appropriate transformations will be documented and justified in the model. Further modelling will examine outcomes over all time points (one, three, and six months) as previously described for the secondary analysis of the primary outcome.

The seizures diary data will be examined and graphically presented by trial arm at 6 months. We will descriptively summarise this data by trial arm (total number of seizures, means and standard deviations for normally distributed variables, medians and interquartile ranges for skewed). We will also report the proportion of patients that had at least a 50% reduction in seizures over the 6 months. If the data quality is sufficient then an Anderson Gill approach will be taken, examining seizures over time.

6.5 Subgroup analyses

Two subgroup analyses will be performed to explore differential treatment effects on the primary outcome measure. Interaction terms will be fitted between trial arm and the following measures thought to be correlated with the primary outcome. These include:

- IQ level as a continuous variable
- Age at recruitment as a continuous variable

6.6 Safety analysis

A summary by trial arm of safety data including abnormal laboratory data (liver, renal, lipid, fasting blood, haematology, blood chemistry and urinalysis), adverse events, and serious adverse events collated throughout the trial will be reported over time in the final report.

6.7 Interim analysis

No interim analyses are planned.

6.8 Methods for handling concomitant medications

A summary by trial arm of concomitant medication and significant non-drug therapies collated throughout the trial will be included in the final report.

6.9 Sensitivity analyses or model testing

Further modelling of the primary outcome will be carried out as follows:

- Any further imbalances on baseline confounders deemed important to adjust for in the modelling.
- An investigation of the impact of missing data by analysis of multiple imputed complete datasets using the mi set of commands in Stata.

6.10 Exploratory analyses

- Where possible we will examine the 4 week period before baseline (visit 1) and compare outcomes to baseline to check the stability in outcomes. How does this change correlate with outcomes? For example, in patients where a deficit occurs pre-randomisation, do these predict worse outcomes or vice versa?
- As an exploratory part of the Quality of Life data we would examine the time point that gave the patients best response (maximum score) and compare to the best response for the primary outcome.
- The use of composite measures of the tests used in the primary outcome will be explored.
- What baseline data (demographics, clinical data) predicts an improvement at 6 months?

- Case studies – look graphically at patients curves or look at which are clustering together using factor analysis or dendograms.
- 9 months post treatment:
    - Examine the change between 6 and 9 months
    - Does this change the best response time point?
- BMI at 6 months
- Examination of the following measurements over time by arm: Vision, abnormal VF, abnormal Fundus exam, respiratory (FEV1 (% predicted)/FVC (% predicted)),

7.0 Imaging Study: Assessing the effects of treatment with RAD001 (Everolimus) on Brain Diffusion Tensor Imaging (DTI) indices in Tuberous sclerosis

7.1 Imaging study objectives

Primary objective(s)

- To compare the effect size of change in DTI indices in the two groups of participants in TRON trial, on treatment with RAD001 (Everolimus) or placebo for 6 months.

Secondary Objective

- To explore whether the neurocognitive deficits in TSC correlate with DTI abnormalities in white matter tracts.
- To correlate the change in DTI parameters of the white matter tracts with any observed amelioration of neurocognitive deficits.

Exploratory Objective

- To compare the DTI indices with age-matched controls in the normal population using data from the "Genetic Neuroimaging" due to start shortly at Cardiff University, and which will acquire directly comparable data.

7.2 Statistical analysis

We will compare changes in DTI indices of fractional anisotropy and mean diffusivity before and after intervention in both treatment and placebo groups. The assessor would be blinded to group allocation. We would also compare the DTI indices with those from age-matched controls in the normal

8. Software

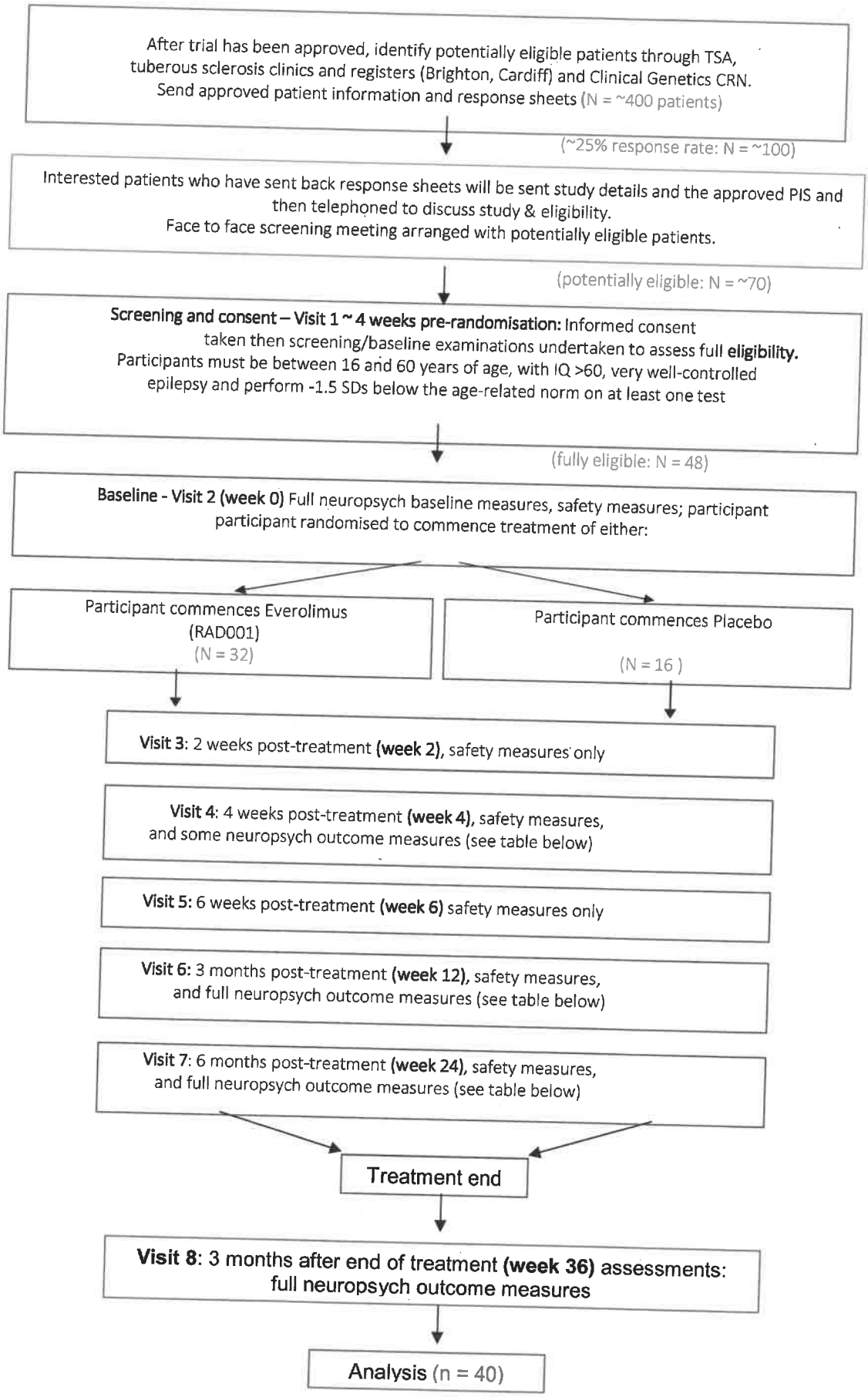
SPSS v20 will be used for all quantitative analyses.

Notes:

Concern that PID013 and their scores on the Neuro tests are a bit all over the place. She seems to think that they are exaggerating their symptom scores and also on some of the tests they were scoring very badly on easy parts but very well on hard parts.

Appendices

Appendix 1. TRON Flow Chart indicating patient numbers and anticipated drop out



Appendix 2. Schedule Showing Breakdown of Neurocognitive Assessments, Questionnaires and Seizure Diary

	Measure	Visit 1 Wk -4 Eligibility visit	Visit 2 Wk 0 Baseline & rando	Visit 4 Wk 4 4wk assess. After 1m of treat	Visit 6 Wk 12 3m assess. After 3m of treat	Visit 7 Wk 24 6m assess. After 6m of treat	Visit Post- 3m aft
Primary outcomes	List learning (BMIPB) -recall after distraction	✓	✓	✓	✓	✓	
	Complex Figure (BMIPB) immediate recall	✓	✓	✓	✓	✓	
	SOC (CANTAB) - minimum moves	✓	✓	✓	✓	✓	
	SWM (CANTAB) ≥4 between errors	✓	✓	✓	✓	✓	
	Dual task (TEA) - dual task decrement	✓	✓	✓	✓	✓	
Secondary outcomes	Rapid Visual Information Processing CANTAB)		✓		✓	✓	
	Spatial span (CANTAB)		✓		✓	✓	
	IDED (CANTAB)		✓		✓	✓	
	Verbal fluency test		✓		✓	✓	
	Cancellation task		✓		✓	✓	
	Symptom Checklist-90R		✓	✓	✓	✓	
	QOLIE		✓		✓	✓	
	Liverpool SS scale		✓	✓	✓	✓	
	Seizure diary (Starts at wks -4 to week 0)	✓	✓	✓	✓	✓	
	VABS -2		✓		✓	✓	
	Social Responsiveness Scale – Adult		✓	✓	✓	✓	
	Social Communication Questionnaire		✓		✓	✓	

Appendix 3. Description of Outcome Measures

Eligibility visit screening measures

1. Wechsler Abbreviated Scale of Intelligence (WASI) (4 subtests)[ix]

A measure of global intelligence, to ascertain that each patient is eligible for the study (IQ more than or equal to 60), and should will be able to perform the required neuropsychological tasks and understand the instructions.
The patient answers questions and has to do simple puzzles. Ages 6 to 89. No parallel forms.
Outcome measures: verbal IQ score, performance IQ score, full scale IQ score, from the conversion of raw score to age-scaled scores.

2. Edinburgh Handedness Test[x]

A measure of right or left handedness, and to what degree, suggesting cerebral dominance. This information may be important in interpreting other tests.
A brief list of questions are presented about which hand would be used for performing everyday tasks, and the patient provides answers.
Outcome measure: scores are converted to percentages indicating handedness.

Primary outcome measures

3. List Learning test (from the BIRT Memory and Information Processing Battery)[xi]

A measure of verbal free recall memory.
A list of words are read to the patient, who says as many as he/she can recall, and the process is repeated five times, followed by a distracter list and a delayed recall trial.
Outcome measures: raw scores are converted to age-related percentile bands.

4. Complex Figure test (from the BIRT Memory and Information Processing Battery)

A measure of visuospatial or non-verbal free recall memory.
The patient has to copy a two-dimensional line drawing and then reproduce it from memory, both immediately after copying, and after a delay of 30 minutes.
Outcome measures: the reproductions are scored according to strict criteria and converted to age-related percentile bands.

(CANTAB)[xii]

This is a touch-screen based, computerised battery of non-verbal tests developed from well-documented paradigms in animal studies. Some tests are 'self-adjusting' to the person's ability, generating more demanding items if serial correct responses are recorded, or terminating the test if serial wrong responses are given. Any age. All subtests have several parallel forms.

5. CANTAB - Stockings of Cambridge (SOC)[xii]

An executive function spatial planning test based upon the 'Tower of London' test.
The patient is presented with three vertical patterns on screen, and must plan how to shift colored discs between one and another, in as few moves as possible, whilst following given rules.
Outcome measures: no. of problems solved in the minimum moves; average moves for each set, each compared to the mean of the normative group, matched for age and gender.

6. CANTAB - Spatial Working Memory (SWM)[xii] A self-ordered task requiring retention and manipulation of spatial information in working memory. The patient must find a blue 'token' hidden in one of the onscreen white boxes by trial and error. On the next trial, he/she must then find the next token, whilst avoiding visiting any boxes visited before in this trial ('within'

errors), and avoiding visiting any boxes that contained the token on previous trials ('between' errors).
Outcome measure: the total number of 'raw' between-errors on all sets of >4 trials length, compared to the norm mean, matched for age and gender.

7. Telephone search dual task (from the Test of Everyday Attention)[xiii]
A measure of ability to perform two crossmodal tasks simultaneously, ie, divided attention. The patient must look for key symbols while searching a telephone directory and simultaneously count strings of auditory tones presented. The telephone directory search task alone is presented for comparison with the dual task.
Outcome measure: dual task scores are compared to single task scores to produce a dual-task decrement score, converted to a age-related percentile.

Secondary Outcome Measures
8. CANTAB - Rapid Visual Information Processing Battery (RVIP)[xii]
A measure of sustained attention and general performance.
A series of digits are presented onscreen. Patients must watch for target sequences and press a key pad as quickly as possible when they see the required sequence.
Outcome measure: raw correct score, latency and sensitivity of response, converted to age-related percentiles.

9. CANTAB - Spatial Span (SSP)[xii]
A test of spatial memory span.
An array of white squares are shown, some of which light up serially, and the patient must remember which ones, and in which order, and duplicate the correct sequence after a brief delay. The number administered increases up to 9.
Outcome measure: highest number of squares correct in one sequence (=span) compared to the norm mean, matched for age and gender.

10. CANTAB - Attentional Set-shifting (IDED) [xii]
This is an executive function test of rule acquisition and reversal, cognitive flexibility and set-shifting, essentially of a non-verbal nature.
Patterns composed of one of two artificial 'dimensions' (color-filled shapes and white lines) are presented, one of which is 'correct' according to a simple, secret rule. The patient must detect and apply the rule, which then changes over a number of stages, and the patient needs to discover what logic the computer is following.
Outcome measure: the no. of stages achieved; the total number of errors, considering how many stages were attained, compared to the mean of the normative group, matched for age and gender.

11. Verbal Fluency /Controlled Oral Word Association Test (COWAT)[xiv]
This is an executive function test of speed and flexibility of verbal thought processes, or verbal fluency.
The patient is given one minute to say as many words as possible beginning with a given letter. This is repeated for other letters.
Outcome measures: the no. of words produced for each letter, compared to the normative mean for age and converted to a percentile band.

12. Cancellation task[xiv]
A test of lateralised attention bias (as above).

A page is presented of small silhouettes of everyday items. The patient must find all the target items (e.g. bells) and cancel them as quickly as possible.
Outcome measure: the number of targets cancelled on each side of the page.

13. Symptom Checklist 90R (SCL-90R)[xv]
To evaluate a broad range of psychological problems and symptoms of psychopathology, such as anxiety and depression, and measure change over treatment duration.
A set of 90 brief questions for the patient to rate.
Outcome measures: scores can be collated on indices to give a measure of intensity of symptoms, or converted to T scores for comparison with adult norms.

14. Quality of Life in Epilepsy (QOLIE)[xvi]
To evaluate overall quality of life, emotional well-being, social isolation, medication effects, perceived physical symptoms and cognitive functions, health perceptions, etc.
A set of either 31 items or 89 are given for the patient to rate.
Outcome measures: scores are converted into indices scores for comparison with adult norms.

15. Liverpool Seizure Severity Scale (LSSS)[xvii]
To evaluate severity and nature of seizures.
A set of 20 clinical features of seizure symptoms are rated by the patient.
Outcome measures: scores are summed for each patient.

16. Vineland Adaptive Behavior Scales-II (VABS-II) (survey form)[xviii]
A low-level measure of functional adaptive behavior, i.e. ability to cope with personal and social skills in everyday life that may be deficient in autism or developmental delays.
Tester completes a questionnaire through interview with parent or caregiver. Any age. No parallel forms.
Outcome measures: raw scores converted to scaled scores, then a summary value for a number of domains.

17. Social Responsiveness Scale – Adult version (SRS-A)[xix]
To measure the severity and type of social impairments characteristic of autism spectrum disorders in children and adolescents (up to 18).
65 items are completed by a parent or carer.
Outcome measures: total score

18. Social communication questionnaire (SCQ)[xx]
To evaluate communication skills and social functioning in those who may have autism or autism spectrum disorders.
A set of 40 questions completed by the parent or carer to address long-term ("Lifetime") or present ("Current") observations.
Outcome measures: total score

19. National Adult Reading Test (NART)[xxi]
To determine the premorbid intellectual function in English-speaking adults.
Pronunciation of 50 phonetically irregular words which are presented to the patient is assessed.
Outcome measures: NART generated IQ score

Appendix 4. Dummy tables

Table 1. Randomisation data by minimisation variables

	Age		Gender		IQ level		Anti-epilepsy drug		Total
	<50	≥50	M	F	60-79	≥80	Yes	No	
Placebo									
Everolimus									
Total									

Table 2a. Baseline descriptives of randomised participants

	Everolimus N=	**Placebo N=**	**Total N=**
Demographics			
Age at recruitment - mean (sd)			
<50 years			
≥50 years			
Gender - N (%) Male			
Currently on AED - N (%) Yes			
Physical Examinations/vital signs			
Weight (kg)			
General physical exam			
Lymphadenopathy - N (%) Yes			
TS skin lesions - N (%) Yes			
ENT – infection			
Visual Feed by confrontation			
N(%) of abnormalities per patient (Upper left/right and lower left/right)			
Fundus Exam – direct ophthalmoscopy			
N(%) of abnormalities per patient (Upper left/right and lower left/right)			
Respiratory/Spirometry			
FEV1 (%Predicted)			
FVC (%Predicted)			
Cardiovascular exam			
Abdominal exam			
CNS Exam			
Any abnormality / Neurological Deficit /CN palsy			
Abnormal Haematology - N(%)			
Abnormal renal function- N(%)			
Abnormal liver function - N(%)			
Abnormal lipid profile- N(%)			
Abnormal fasting glucose- N(%)			

	Everolimus N=	Placebo N=	Total N=
Epilepsy – Seizure			
Stable[1] – N (%) Yes			
Seizure Type – N (%)			
GTCS			
CPS			
Healthcare consultations			
Starting dose (tablets/per day)			

Table 2b. Baseline neuropsychological outcome of randomised participants

	Everolimus N=	Placebo N=	Total N=
Measured at visit 1-pre baseline			
WASI (IQ)			
Verbal			
Performance			
Full-4 scale			
IQ - N(%) 60+			
EHI Handedness test			
LQ			
Decile			
NART Error Scale			
Full scale			
Verbal			
Performance			
Error score			
List learning (BMIPB)			
Immediate recall (percentile)			
Delayed recall (percentile)			
Complex Figure (BMIPB)			
Immediate recall (percentile)			
Delayed recall (percentile)			
Dual task (TEA)			
Telephone (percentile)			
Telephone search while counting (percentile)			
SOC (CANTAB)			
Mean initial thinking time			
Mean subsequent thinking time			
Number of problems solved in min moves			
SWM (CANTAB)			
Between-errors			
Strategy fields			
Neuropsychological test[2] that showed a deficit > -1.5 SD			
1 test			

[1] Seizure free/AED unchanged for 6 months

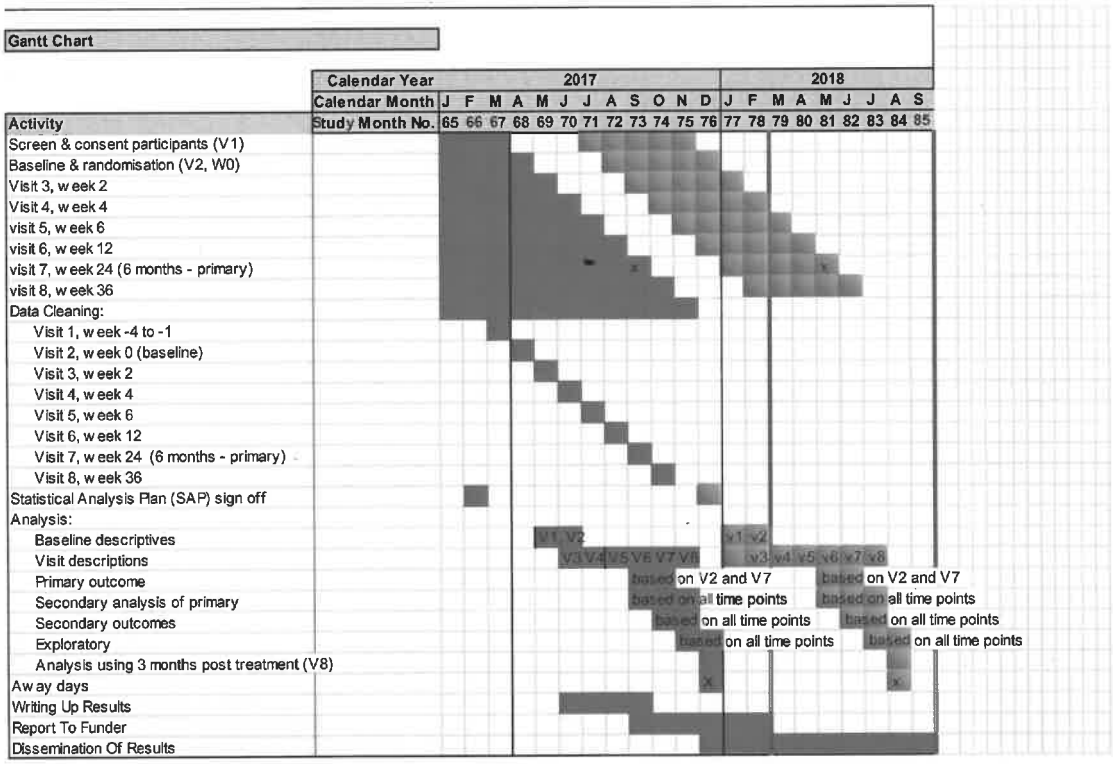
[2] List learning test, complex figure test, TEA, SOC (CANTAB), SWM (CANTAB)

	Everolimus N=	Placebo N=	Total N=
2 tests			
3 tests			
4 tests			
All 5 tests			
Rapid visual Information (CANTAB)			
Spatial Span (CANTAB)			
IDED (CANTAB)			
Verbal fluency test			
Cancellation task			
Symptom checklist 90R			
QOLIE			
Liverpool SS scale			
Seizure diary			
VABS-2			
Social Responsiveness - Adult			
Social Communication Questionnaire			

Table 2. Primary outcome: Mean (standard deviation) of individual tests at baseline, 6 months and the change by treatment group

Test	Task	Treatment group	N	Baseline	6 months	Change (6m - Baseline)	N (%) deficit (≤1SD in change score)	% improvement (≥1SD in change score)
B-MIPB	*List learning test – recall after distraction*	Everolimus						
		Placebo						
B-MIPB	*Complex figure test - delayed recall*	Everolimus						
		Placebo						
		Placebo						
CANTAB-SOC	*Minimum moves*	Everolimus						
		Placebo						
CANTAB-SWM[a]	*Between-Errors*	Everolimus						
		Placebo						
TEA	*Dual task decrement*	Everolimus						
		Placebo						
OVERALL		**Everolimus**						
		Placebo						

Appendix 5. Gantt chart for data management and analysis



Log of changes to SAP

References

[i] ICH E9: Statistical principle for clinical trials (Notes for Guidance on Statistical Principles for Clinical Trials) September 1998, CPMP/ICH/363/96.

[ii] ICH E3: ICH Harmonised Tripartite Guideline: Structure and Content of Clinical Study Report E3

[iii] ICH E6 (R1): Guideline for Good Clinical Practice. *European Medicine Agency*; CPMP/ICH/135/95

[iv] Playle, R. Standard Operating Procedures for Statistical Analysis Version 0.2, 9/11/07

[v] TRON Protocol version 4.0, 01/029/2012.

[vi] Fleming TR. One-sample multiple testing procedure for Phase II clinical trials. Biometrics 1982 38. p. 143-151.

[vii] TRON randomisation protocol v0.1 03/04/2012.

[viii] Campbell MK, Elbourne DR and Altman DG (2004). CONSORT statement: extension to cluster randomised trials. BMJ. 328:702-708.

[ix] Wechsler D. Manual for the Wechsler Abbreviated Scale of intelligence (1999) The Psychological Corporation, San Antonio, TX

[x] Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971 9(1):97-113.

[xi] Coughlan AK, Oddy MJ, Crawford JR (2007), BIRT Memory and Information Processing Battery (BMIPB), Brain Injury Rehabilitation Trust.

[xii] Cambridge Neuropsychological Test Automated Battery (CANTAB) (2008) Cambridge Cognition, UK

[xiii] Robertson IH, Ward T, Ridgeway V, Nimmo-Smith I (1994) Test of Everyday Attention (TEA), Psychological Corporation, USA.

[xiv] Benton AL, Hamsher de SK, Sivan AB (1983) Multilingual aplasia examination. Second Edition, Iowa City, AJA Associates. [COWAT]

[xv] Derogatis LR, Symptom Checklist-90-Revised, Pearson Assessments..

[xvi] Quality of Life in Epilepsy Inventory (QOLIE), RAND Health, Santa Monica, USA.

[xvii] Baker GA, Smith, Jacoby A, Hayes J, Chadwick D. Liverpool Seizure Severity Scale revisited Seizure 1993;3: 201-205.

[xviii] Sparrow SS, Cicchetti DV, Balls DA (2005) Vineland Adaptive Behavior Scales, Second Edition ('Vineland II'). AGS Publishing, Circle Pines, USA.

[xix] Constantino JN (2005) Social Responsiveness Scale (SRS), Western Psychological Services, USA.

[xx] Rutter M, Bailey A, Lord C (2003) Social Communication Questionnaire (SCQ) Western Psychological Services, USA

[xxi] Nelson, HE (1982). The National Adult Reading Test (NART): test manual. :NFER-Nelson.