The effect of everolimus on neurocognitive aspects of tuberous sclerosis complex

Submitted for the degree of Doctor of Philosophy

Cardiff University

Anurag Saxena

February 2021

Table of Contents

List of Tables

List of Figures

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed …… Date: 24/02/2021……

STATEMENT 1

This thesis is submitted in partial fulfilment of the requirements for the degree of PhD.

Signed …… Date 24/02/2021……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed …… Date 24/02/2021……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… Date 24/02/2021……

Summary

Tuberous sclerosis complex (TSC) is a multi-system genetic disorder. It is characterised by benign tumour development in many organs, epilepsy and TSC-associated neuropsychiatric disorders, termed TAND that are frequent and have a great impact on the lives of patients and their carers.

TSC results from mutations in either *TSC1* (encoding TSC1) or *TSC2* (encoding TSC2). Loss of functional TSC1 or TSC2 leads to activation of mTORC1 (mammalian target of rapamycin complex 1), a regulator of multiple cellular processes, including cell growth and neuronal function. Everolimus is an inhibitor of mTORC1 and is used in clinical practice for managing TSC-associated tumours and epilepsy.

This thesis reports a phase 2, randomised, placebo-controlled trial assessing the safety and efficacy of 6 months of treatment with everolimus in adults with TSC and deficits in memory and/or executive function. It aimed to determine effect sizes in its placebo and everolimus arms to inform future development of larger trials. Responders were defined as those showing an improvement of ≥1SD in at least one of ten measures of memory and executive function. Microstructural changes in cerebral white matter tracts associated with treatment were assessed using diffusion tensor imaging.

Recruitment proved challenging, requiring modification of the original single centre design to a three centre study and revision of eligibility criteria to make them less stringent. Problems with application of one of the eligibility and outcome measures, the 'Test of Everyday Attention', further affected patient numbers for the primary analysis.

In the final primary analysis (after adjustment) of 29 participants, 4 of 9 (44.4%) of the placebo group and 14 of 20 (70%) in the everolimus group were responders, satisfying criteria supporting further study as defined in the protocol.

Twenty-two participants had brain DTI scans at baseline and after 6 months of treatment. In patients treated with everolimus, fractional anisotropy and mean diffusivity demonstrated a statistically significant change in several white matter tracts, but only in a subgroup of these patients.

Adverse events were common and consistent with the known toxicities of everolimus. These findings suggest that mTOR inhibition has potential as a therapeutic strategy in treating TAND aspects of TSC manifestations and warrants further investigation.

Roles and Acknowledgements

Co-Investigators and collaborators

Alison Cooper	Radiographer, CUBRIC, Cardiff	AC
Andrea Liu	Consultant Neuroradiologist, Cardiff	AL
Anurag Saxena	Clinical Research Fellow, Cardiff	AS
Cheney Drew	Trial manager, SEWTU, Cardiff	CD
Derek Jones	Professor, Director of CUBRIC, Cardiff	DJ
Diedre Donnelly	Consultant in clinical genetics, Belfast	DD
Eleri Owen-Jones	Trial manager, SEWTU, Cardiff	EOJ
Elizabeth Randell	Trial manager, SEWTU, Cardiff	ER
Emer McDermott	Assistant Psychologist, Cardiff	EM
John Evans	MRI Research Fellow & MRI Lab Manager, CUBRIC, Cardiff	JE
Julian Sampson	Professor of Medical Genetics, Cardiff	JS
Kerenza Hood	Professor in Statistics & Director SEWTU, Cardiff	KeH
Khalid Hamandi	Consultant Neurologist, Cardiff	KH
Laura Stockwell	Assistant Psychologist, Cardiff	LS
Lianna Angel	Data Manager, SEWTU, Cardiff	LA
Mark Davies	Clinical Lecturer in Medical Genetics , Cardiff	MD
Mark Odgers	Trial Administrator, SEWTU, Cardiff	MO
Matthew Mort	Senior Biostatistician, HGMD, Cardiff	MM
Maxime Chamberland	Research Fellow, CUBRIC, Cardiff	MC
Michelle Smalley	Clinical Psychologist, Cardiff	MS
Nigel Kirby	Data Manager, SEWTU, Cardiff	NK

Peter Hobden	Research Radiographer, CUBRIC, Cardiff	PH
Petrus J de Vries	Professor of Child & Adolescent Psychiatry, Cape Town	PJdV
Rachel McNamara	Senior Trial Manager, SEWTU, Cardiff	RM
Rebecca Cannings-John	Senior research fellow, Statistics, CTR, Cardiff	RCJ
Shelagh Joss	Consultant in clinical genetics, Glasgow	SJ
Sonya Foley	Scientific support officer, CUBRIC, Cardiff	SF

Roles

The TRON clinical study protocol was written by JS, PJdV, MD, with the assistance of the team at SEWTU/CTR. AS contributed to the protocol approval process along with JS and SEWTU/CTR.

AS coordinated the trial and undertook the clinical assessments. Pharmacovigilance was performed by AS in conjunction with JS. AS undertook participant consent, clinical assessments at study visits, reporting and management of adverse events and liaison with patients and their local care teams.

Participants were managed by AS, JS, SJ and DD. AS managed the participants seen in Cardiff in conjunction with JS, while participants seen in Glasgow and Belfast were managed by AS in conjunction with SJ and DD. Neurocognitive assessments were performed by EM and LS supervised by MS. TRON trial was monitored by AS in conjunction with the SEWTU team (South east Wales trials unit) comprising ER, EOJ, CD, LA, NK, MO and KeH.

AS, DJ, and JS wrote the brain imaging study protocol. AS wrote the sections dealing with the study design, safety issues, inclusion and exclusion criteria and assessments. Data analysis section was written by RCJ and AS. The imaging acquisition section was written by AS with help from DJ & AL.

AS and JS obtained funding for the imaging project and permission to conduct the trial with the assistance of the SEWTU team. AS coordinated and monitored the imaging study. AC, PH and SF helped to acquire the imaging data assessments. JE was involved in secure data retention and storage. AS performed the preprocessing of imaging data in conjunction with SF and KH. MC performed the Tracseg analysis. Statistical analysis was done by AS with help from RCJ, DJ & MM.

Acknowledgements

I would like to thank Prof Julian Sampson and Prof Derek Jones for their helpful advice, support and constructive criticism.

I want to thank Mark and Khalid for all the encouragement, especially when things did not go as planned. Special thanks to Emma for being kind to the participants, my consultant colleagues, Frances and Johann, for their kindness to me during the years it took to complete this project. I am particularly indebted to all my family and friends who have been incredibly patient and have let me complete "my hobby".

This project would not have been possible without the study participants' support and commitment and funding from the Tuberous Sclerosis Association (UK), Wales Gene Park and Novartis International AG.

Abbreviations

ACE	Angiotensin-converting enzyme
AD	Axial diffusivity
ADC	Apparent diffusion coefficient
ADNFLE	Autosomal dominant nocturnal frontal lobe epilepsy
ADOS	Autism diagnostic observation schedule
AE	Adverse event
AED	Antiepileptic drug
AFB	Angiofibroma
AKT	protein kinase B
ALT	Alanine Aminotransferase
AML	Angiomyolipomas
AMPK	AMP-activated protein kinase
APA	American psychological association
AR	Adverse reaction
AST	Aspartate aminotransferase
ATP	Adenosine triphosphate
AUC0-т	area under the plasma concentration-time curve
BECTS	Benign epilepsy of childhood with centrotemporal spikes
BIRT	Brain Injury Rehabilitation Trust

BMIPB	BIRT Memory and Information Processing Battery
bvFTD	behaviour -variant frontotemporal dementia
CA1	Cornu Ammonis 1
CADASIL	Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy
CANTAB	Cambridge Neuropsychological Test Automated Battery
CG	Cingulum
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
COVID-19	Coronavirus disease 2019
COWAT	Controlled Oral Word Association Test
CPK	Creatine phosphokinase
CRF	Clinical research facility
CSF	Cerebrospinal fluid
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CUBRIC	Cardiff University Brain Research Imaging Centre
CYP3A4	Cytochrome P450 3A4
DEPDC5	Dishevelled, Egl-10 and pleckstrin domain containing protein 5
DEPTOR	DEP domain-containing mTOR-interacting protein
dMRI	Diffusion MRI

DN	Dysmorphic neurons
DOHSC	Department of Health and Social care
DQ	Developmental quotient
DTI	Diffusion Tensor Imaging
EC	Elevator counting
ECD	Elevator counting with distraction
ECG	Electrocardiogram
ECR	Elevator counting with reversal
EDS	extra-dimensional shift
EEG	Electroencephalogram
EMA	European Medicines Agency
ER	Endoplasmic reticulum
ERK	Extracellular receptor kinases
EU	European Union
EudraCT	European Union Drug Regulating Authorities Clinical Trials Database
FA	Fractional anisotropy
FCD	Focal cortical dysplasia
FDA	U.S. Food and Drug Administration
FFEVF	Familial focal epilepsy with variable foci
FKBP	FK506 binding protein
FLAIR	Fluid-attenuated inversion recovery

FSIQ	Full-Scale IQ
FX	Fornix
G1 phase	Gap1 phase
GC	Giant Cells
GCP	Good clinical practice
GFR	Glomerular filtration rate
GM	Grey matter
GTP	Guanosine-5'-triphosphate
HIF	Hypoxia-inducible factor
HME	Hemimegalencephaly
HMG-CoA	3-hydroxy-3-methyl-glutaryl-coenzyme A
HRCT	High-resolution chest computed tomography
IDED	ID/ED (Intra/extra-dimensional set) Shift
INR	International normalised ratio
IQ	Intelligence quotient
IRDiRC	International Rare Disease Research Consortium
ISRCTN	International Standard Randomised Controlled Trials Number
LAM	Lymphangioleiomyomatosis
LKβ1	Liver kinase β1
LSS	Liverpool Seizure Severity Scale
Lst8	Lethal with SEC13 protein8

LTP	Long-term potentiation
MCAP	Megalencephaly–capillary malformation syndrome
MD	Mean diffusivity
MEK	Mitogen-activated protein kinase
mLST8	mammalian lethal with Sec13 protein 8
MMPH	Multifocal micronodular pneumocyte hyperplasia
MPPH	Megalencephaly–polymicrogyria–polydactyly–hydrocephalus syndrome
MRI	Magnetic resonance imaging
mRNA	messenger ribonucleic acid
MS	Map search
MSEL	Mullen scale of early learning
mSIN1	mammalian stress-activated protein kinase interacting protein 1
mTOR	mammalian target of rapamycin
mTORC1	mTOR complex 1
mTORC2	mTOR complex 2
mTORi	mTOR inhibitor
NART	National adult reading test
NAWM	Normal-appearing white matter
NCI	National Cancer institute
NF1	Neurofibromatosis type I
NICE	National institute for health and care excellence

NIHR	National institute for health research
NMI	No mutation Identified
NORD	National organisation for rare disorders
NS-SEC	National Statistics socio-economic classification
OMIM	Online Mendelian Inheritance in Man
PDK1	3-phosphoinositide-dependent protein kinase-1
PgP	P-glycoprotein
PI3K	phosphatidylinositide-3 kinase
PICs	Patient identification centres
PIKK	Phosphoinositide 3-kinase related kinase
PIP_2	phospholipid phosphatidylinositol (4,5) bisphosphate
PIP_3	phosphatidylinositol (3,4,5) trisphosphate
PIQ	Performance IQ
PK	Pharmacokinetic
PKD	Polycystic kidney disease
PMSE	Polyhydramnios, megalencephaly, and symptomatic epilepsy
PRAS40	Proline-rich Akt substrate
PTEN	Phosphatase and tensin homolog deleted on chromosome ten
QOLIE	Quality of Life in Epilepsy
Raf	Rapidly Accelerated Fibrosarcoma kinase
Raptor	Regulatory-associated protein of mTOR

Ras	Rat sarcoma protein
RCT	Randomised control trial
RD	Radial diffusivity
REC	Research ethics committee
Rheb	Ras homolog expressed in brain
RICTOR	RAPTOR-independent companion of TOR
ROI	Region of Interest
ROS	Reactive oxygen species
RSK	Ribosomal S6 kinase
RVIP	Rapid Visual Information Processing
S phase	Synthesis phase
SAE	Serious adverse event
SAR	Serious adverse reaction
SCL	Symptom Checklist
SCQ	Social communication questionnaire
SD	Standard deviation
SEGA	Subependymal giant cell astrocytoma
SEN	Subependymal nodules
SEWTU	South East Wales Trials Unit (Centre for trial research – CTR)
SLF	Superior longitudinal fasciculus
SMA	Spinal muscular atrophy

SmPC	Summary of product characteristics
SNR	Signal to noise ratio
SOC	Stocking of Cambridge
SoS	Son of Sevenless protein
SPF	Sun protection factor
SRS-A	Social Responsiveness Scale – Adult version
SSP	Spatial memory span
STRADA	STE20-related kinase adaptor α
SUSAR	Sudden unexpected serious adverse reaction
SWM	Spatial working memory
TACERN	Tuberous Sclerosis Complex Autism Center of Excellence Research Network
TAND	Tuberous Sclerosis Associated Neuropsychiatric Disorders
TBC domain	Tre-2/Bub2/Cdc16 domain
TBC1D7	TBC1 domain family, member 7
TBSS	Tract-based Spatial Statistics
TE	Echo time
TEA	Test for Everyday Attention
TESSTAL	Trial of Efficacy and Safety of Sirolimus in Tuberous Sclerosis and LAM
TLE	Temporal lobe epilepsy
TOSCA	TuberOus SClerosis registry to increase disease Awareness
TR	Repetition time

TRON	Treatment of Neurocognitive problems in tuberous sclerosis
TS	Telephone search
TSA	Tuberous Sclerosis Association
TSC	Tuberous Sclerosis Complex
TSwC	Telephone Search While Counting
UF	Uncinate fasciculus
ULN	Upper limit of normal
VABS	Vineland Adaptive Behavior Scale
VBM	Voxel-based Morphometry
VE	Visual elevator
VIQ	Verbal IQ
WAIS	Wechsler Adult Intelligence Scale
WASI	Wechsler Abbreviated Scale of Intelligence
WISC	Wechsler Intelligence Scale for Children
WM	White matter
WT	Wild type

1 Clinical Features of Tuberous Sclerosis Complex

1.1 Phenotype and clinical diagnosis of tuberous sclerosis

Tuberous sclerosis, also known as tuberous sclerosis complex (TSC), is an autosomal dominant disorder affecting up to 1 in 6,000 of the population. (Crino et al. 2006) It is caused by mutations in the *TSC1* or *TSC2* gene (Consortium 1993; van Slegtenhorst et al. 1997) and characterised by the development of hamartomatous lesions or growths in many organs. Clinical manifestations of TSC vary amongst individuals in the number and severity of the organs involved.

Historically, the initial recognition of TSC as an entity was by Bourneville (Bourneville 1899), who felt the brain lesions of affected individuals resembled "root vegetables" or "tubers." Vogt proposed a characteristic triad of epilepsy, facial rash, and learning problems in 1908 (Vogt 1908). For nearly a century, the diagnosis of TSC was based on these criteria until Gomez (Gomez 1988) and Roach (Roach et al. 1992) described revised and more specific clinical criteria. These were reviewed again in 2012 by an international group of TSC specialists that proposed a revised diagnostic pathway to reflect the giant strides made in understanding the mechanistic basis of tuberous sclerosis, including the underlying genetic mechanisms.

Manifestations are more severe in *TSC2* than *TSC1*-associated disease (Jones et al. 1999a; Dabora et al. 2001; Lewis et al. 2004; Sancak et al. 2005) and emerge in different organs with age. Some aspects of brain and heart involvement usually become apparent during the prenatal period, while most signs in the skin, kidney and lung become apparent during childhood or adult life. A diagnosis of tuberous sclerosis can be made according to clinical criteria or genetic testing (Table 1) as per the recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference (Northrup et al. 2013)

1.1.1 Neurological manifestations

The characteristic macroscopic brain pathology in TSC includes cortical and subcortical tubers, subependymal nodules (SEN) in the periventricular zone lining the lateral ventricles and subependymal giant cell astrocytoma (SEGA). Tubers are developmental cerebral cortex abnormalities which are linked to epileptogenesis in TSC. SEGA and SEN are hamartomas arising from the lateral and third ventricles' walls. As SEGAs characteristically grow near the Foramen of Monro, the cerebrospinal fluid circulation could be blocked, leading to progressive ventricular dilatation and increased intracranial pressure, while SEN remain asymptomatic (Ekici et al. 2011).

1.1.2 Seizures and intellectual disability

Epilepsy is the most common symptom of brain involvement and develops in 75-90% of individuals with TSC. (Webb et al. 1991; Thiele 2004) Around 80% of children with tuberous sclerosis have seizures (Joinson et al. 2003). The most typical presentation of the condition is with infantile spasms in infancy or early childhood (Curatolo et al. 2001). Many types of seizure have been reported. The seizures are often refractory to treatment and may require the use of multiple antiepileptic drugs (Saxena and Sampson 2015). Infantile spasms associated with poor cognitive prognosis occur in 20-30% of children with tuberous sclerosis. Later seizure patterns include Lennox-Gastaut Syndrome, focal, complex focal and multifocal seizures and drop attacks. The majority of patients have two or more seizure types. TSC individuals with refractory epilepsy usually have a younger age at diagnosis (epilepsy onset ≤ 1yr), a history of infantile spasms and/or Lennox–Gastaut syndrome, lower educational achievement, a higher prevalence of psychiatric problems, and association with *TSC2* mutation (Vignoli et al. 2013). Mutations in the *TSC1* and *TSC2* genes that cause tuberous sclerosis lead to hyperactivation of the mammalian (or mechanistic) target of rapamycin complex 1 (mTORC1). Inhibitors of mTORC1 have been shown to be effective treatments for many manifestations of tuberous sclerosis and are licensed for use as an adjunctive antiepileptic drug in tuberous sclerosis-related epilepsy (French et al. 2016).

Historically, a bimodal distribution of intelligence quotient (IQ) scores was described in individuals with tuberous sclerosis, suggesting two distinct subgroups (Joinson et al. 2003). Approximately 30% had profound intellectual impairment with very low IQs (too low to be assessed by standardised cognitive function measures); the remaining 70% had IQs with a slight reduction in mean IQ relative to unaffected individuals. This second group was affected by specific cognitive deficits, such as deficiency in long term memory, attention span and executive skills (Harrison et al. 1999b; Prather and de Vries 2004; de Vries et al. 2007). Recent studies have identified profound intellectual disability in a much smaller proportion of children with TSC, an observation that may be attributed to improved control of early seizures (Tye et al. 2018).

1.1.3 TSC – Associated Neuropsychiatric Disorders (TAND)

TAND is an umbrella term coined by the Neuropsychiatry Panel of the 2012 International Consensus Conference to cover all TSC – associated neuropsychiatric disorders (Krueger et al. 2013). The 2012 Neuropsychiatry panel recognised that the multidimensional presentations of TAND potentially leads to clinical and scientific confusion about the different terminologies used. In order to simplify and expedite their recognition, TAND manifestations were conceptualised as ”levels or aspects” in behavioural, psychiatric, intellectual, academic, neuropsychological, and psychosocial manifestations of TSC. It was also advised that TAND symptoms should be treated with pharmacologic and non-pharmacologic interventions, individualised for each patient's specific TAND profile (Krueger et al. 2013).

Patients and their families rate neurodevelopmental and neurobehavioural issues as the most important consequences of TSC (Hunt 1983). Approximately 20-60% of patients with tuberous sclerosis meet autistic spectrum disorder criteria, with the criteria being met in 17% of those with an IQ in the normal range (de Vries et al. 2007). In adults, psychiatric features such as depression and anxiety are also frequently encountered (Lyczkowski et al. 2007). TAND features are underdiagnosed due to a lack of awareness and the time required to make the diagnosis. The 2012 consensus conference recommended screening for TAND symptoms at each clinical visit (at least once per year), reflecting concerns that TAND manifestations may not be evident in the initial assessments and can emerge with time.

While planning the TRON trial, the study group felt that the neuropsychological aspects of learning, thinking, memory, attentional skills and executive function could be evaluated objectively for a clinical trial. In contrast, ascertainment of other TAND levels presentations was more difficult. In addition, as the neuropsychological level may correlate with other TAND levels such as behaviour, academic abilities and psychiatric disorders, any improvement in these aspects may have a wide-ranging effect on TSC individuals (de Vries et al. 2015).

1.1.4 Renal manifestations

Tuberous sclerosis manifestations in the kidneys include angiomyolipomas (AMLs) (Figure 1), oncocytomas, simple cysts, polycystic kidney disease and renal cell carcinoma (O'Callaghan et al. 2004). AMLs develop in approximately 80% of TSC patients and usually are multiple and bilateral. AMLs are frequently asymptomatic but can present with flank pain and impair renal function. They can potentially bleed spontaneously, causing haematuria and occasionally life-threatening

haemorrhage (Bissler and Kingswood 2004). mTOR inhibitor treatment has become the management of choice in all patients with symptomatic AMLs or asymptomatic lesions of >3 cm, as they are at higher risk of rupture/haemorrhage. Surgical options such as resection or selective arterial embolisation are used infrequently, except for managing acute haemorrhagic episodes.

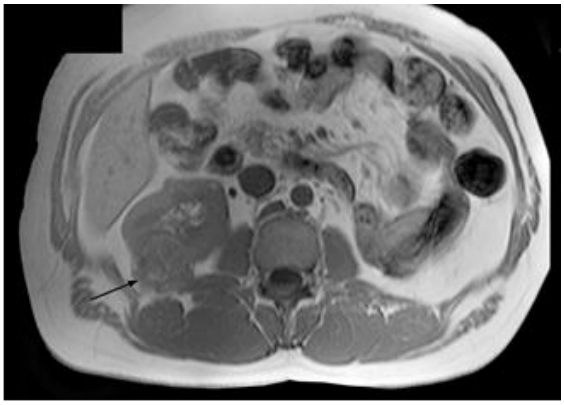


Figure 1 MRI scans showing unilateral AMLs in the right kidney (identified by an arrow) of a patient with tuberous sclerosis (Davies 2011)

Renal cysts are common in tuberous sclerosis and usually asymptomatic. However, around 5% of patients have a contiguous deletion of *TSC2* and the adjacent *PKD1* gene that leads to an early-onset form of polycystic disease (Brook-Carter et al. 1994) which often progresses to renal failure in early adult life (Brook-Carter et al. 1994; Sampson et al. 1997). The incidence of renal cell carcinomas in patients with TSC is thought to be similar to that in the general population. However, renal cell carcinoma associated with tuberous sclerosis tends to occur in younger patients (Tello et al. 1998). A variety of renal cell carcinoma types, including clear cell, papillary, and chromophobe carcinoma, have been reported in patients with tuberous sclerosis.

1.1.5 Pulmonary involvement

Lung complications in TSC include a rare disorder, lymphangioleiomyomatosis (LAM), which is almost exclusively clinically significant in adult females and presents with cough, haemoptysis, shortness of breath or pneumothorax and may progress to respiratory failure (Bissler et al. 2008). Radiographic studies suggest that 40% of female patients with tuberous sclerosis have pulmonary changes, although many of these women are asymptomatic (Costello et al. 2000). The focal proliferation of type II pneumocytes, termed multifocal micronodular pneumocyte hyperplasia (MMPH), may also be a pulmonary manifestation of tuberous sclerosis but is not thought to be clinically significant (Guinee et al. 1995).

1.1.6 Dermatological manifestation

Skin involvement includes hypomelanotic macules and fibrous plaques that often manifest in infancy. Facial angiofibromas are flesh-coloured or red papules, typically occurring over the nose, nasolabial folds, cheeks and chin, which usually manifest in childhood while periungual fibromas develop in older children and adults. (Northrup et al. 2013) Shagreen patches, which are elevated irregular brown or flesh coloured lesions, typically develop in the lumbar-sacral area during childhood. Ungual fibromas are flesh-coloured or pink nodules that occur on the finger or toenail beds. Linear depression in the nail can suggest the presence of a subungual fibroma. Gingival fibromas can also occur.

1.1.7 Cardiac features

Heart involvement is seen in more than half of infants with TSC. The usual manifestation is cardiac rhabdomyoma, often multiple and can be an unexpected finding on routine antenatal scans, appearing between 22 to 28 weeks of gestation (Bader et al. 2003). Cardiac rhabdomyomas usually resolve spontaneously, in contrast to other manifestations in TSC. (Bosi et al. 1996) They typically are asymptomatic but can be associated with obstructive heart failure or life-threatening arrhythmias.

1.1.8 Ophthalmological features

The common ophthalmologic manifestations of tuberous sclerosis are retinal hamartomas (Rowley et al. 2001), which occasionally affect vision if they are in the visual axis.

1.1.9 Neuroendocrine tumours

Rare case reports chronicle neuroendocrine tumours of the pituitary, pancreas and parathyroid, as well as pheochromocytomas in patients with TSC. A molecular analysis demonstrating a second hit mutation in *TSC2* suggested a genuine mechanistic association (Verhoef et al. 1999). The strength of the association between neuroendocrine tumours and tuberous sclerosis remains unclear (Dworakowska and Grossman 2009).

1.1.10Gastrointestinal polyps

Gastrointestinal polyps (generally hamartomatous) may be expected in patients with tuberous sclerosis (Devroede et al. 1988; Gould et al. 1990; Gould 1991; Hizawa et al. 1994). The prevalence of gastrointestinal polyps is underestimated as they usually remain asymptomatic. Gould *et al.* documented that 14 of 18 (78%) adult patients with tuberous sclerosis had intestinal polyps (Gould 1991).

1.1.11Skeletal involvement

Orthopaedic manifestations of tuberous sclerosis include bone cyst-like lesions, hyperostosis of the inner table of the skull, osteoblastic changes, periosteal new bone formation, and scoliosis (Boronat and Barber 2018).

1.2 Clinical assessment and diagnostic criteria

Clinical features remain the principal means of diagnosis; if the individual has two major clinical features or one major feature with ≥2 minor features the diagnosis is considered definite. Identifying a pathogenic mutation in *TSC1* or *TSC2* is sufficient on its own to make a diagnosis. Table 1 provides the diagnostic criteria per 2012 international consensus conference recommendations (Northrup et al. 2013).

TABLE 1 Updated diagnostic criteria for Tuberous Sclerosis Complex 2012

A. Genetic diagnostic criteria 1. The Identification of either a TSC1 or TSC2 pathogenic mutation in DNA from normal tissue is sufficient to make a definite diagnosis of tuberous sclerosis complex (TSC). 2. A normal result does not exclude TSC or affects the use of clinical diagnostic criteria to diagnose TSC. 3. 10% to 25% of TSC patients have no mutation identified by conventional genetic testing
B. Clinical diagnostic criteria (Major features) 1. Hypomelanotic macules (≥3, at least 5-mm diameter) 2. Angiofibromas (≥3) or fibrous cephalic plaque 3. Ungual fibromas (≥2) 4. Shagreen patch 5. Multiple retinal hamartomas 6. Cortical dysplasias* 7. Subependymal nodules 8. Subependymal giant cell astrocytoma 9. Cardiac rhabdomyoma 10. Lymphangioleiomyomatosis (LAM)‡ 11. Angiomyolipomas (≥2)‡
C. Clinical diagnostic criteria (Minor features) 1. "Confetti" skin lesions 2. Dental enamel pits (>3) 3. Intraoral fibromas (≥2) 4. Retinal achromic patch 5. Multiple renal cysts 6. Nonrenal hamartomas
Definite diagnosis: Two major features or one major feature with ≥2 minor features
Possible diagnosis: Either one major feature or ≥2 minor features *Includes tubers and cerebral white matter radial migration lines ‡A combination of LAM and angiomyolipomas without other features does not meet criteria for a definite diagnosis

1.2.1 Clinical management and surveillance

Clinical management requires a tailored approach to the specific problems of the affected individual. Ideally, patients should be seen by specialists in a multidisciplinary team environment in a specialist TSC clinic. There are disease-specific deviations from standard practice for some areas, such as in antiepileptic drug selection.

In individuals newly diagnosed or suspected of having tuberous sclerosis, the initial assessments shown in Table 2 should be considered to establish the diagnosis and to identify complications amenable to treatment as per national and international guidance (Roach and Sparagana 2004; Krueger et al. 2013; Amin et al. 2019).

TABLE 2 Summary of initial assessments for newly diagnosed/suspected TSC

Specialty Area	Recommendation
Genetics	Obtain three-generation family history to assess for additional family members at risk of TSC and offer genetic testing
Brain	Perform magnetic resonance imaging (MRI) of the brain Evaluate for TSC-associated neuropsychiatric disorder (TAND) Obtain baseline routine electroencephalogram (EEG)
Kidney	Obtain MRI of the abdomen Screen for hypertension and evaluate renal function
Lung	Perform baseline pulmonary function and high-resolution chest computed tomography (HRCT), in patients at risk of Lymphangioleiomyomatosis (LAM), typically females ≥18 years
Skin	Perform a detailed clinical dermatologic inspection/exam
Teeth	Perform a detailed clinical dental inspection/exam
Heart	Obtain an echocardiogram in paediatric patients and electrocardiogram (ECG) in all ages to assess for underlying conduction defects
Eye	Perform a complete ophthalmologic evaluation, including dilated fundoscopy, to assess for retinal lesions and visual field deficits

After the initial evaluation, the surveillance summarised in Table 3 has been suggested in patients diagnosed with tuberous sclerosis (Krueger et al. 2013; Amin et al. 2019):

TABLE 3 Surveillance and management recommendations for patients with definite or possible TSC

Speciality Area	Recommendation
Genetics	• Offer genetic testing and family counselling, if not done previously.
Brain	• MRI of the brain every 1-3 yr in asymptomatic TSC patients ≤ 25 yr to monitor for SEGA, with more frequent MRI for asymptomatic patients with large or growing SEGA, or SEGA is causing ventricular enlargement. • Surgical resection should be performed for acutely symptomatic SEGA, while either surgery or medical treatment with mTORi may be considered for growing but asymptomatic SEGA, after discussion in a Neuro-Oncology MDT. • Perform screening for TSC-associated neuropsychiatric disorders (TAND) features at least annually at each clinical visit with comprehensive developmental assessments at key developmental time points. • Management strategies should be based on the TAND profile of each patient. • Obtain routine electroencephalograph (EEG) in individuals with known or suspected seizures. • Vigabatrin is the recommended first-line therapy for infantile spasms, with hormonal therapy if it is unsuccessful. • Anticonvulsant therapy of other seizure types in TSC should generally follow other epilepsies with consideration for epilepsy surgery for refractory epilepsy.
Kidney	• MRI of the abdomen to assess for the progression of AMLs and cystic renal disease every 1-3 yr throughout the patient's lifetime. • Assess renal function (including determining glomerular filtration rate [GFR]) and blood pressure at least annually. • Embolisation followed by corticosteroids is first-line therapy for AMLs presenting with acute haemorrhage. Nephrectomy is to be avoided. • For asymptomatic, growing angiomyolipoma measuring > 3 cm in diameter, treatment with a mTORi is the recommended first-line therapy. Selective embolisation or kidney-sparing resection are acceptable second-line therapy.
Lung	• Perform clinical screening for LAM symptoms at each clinic visit. Counselling regarding smoking risk and oestrogen use should be provided at each clinic visit for individuals at risk of LAM. • Obtain a High-Resolution CT (HRCT) chest every 5-10 yr in asymptomatic individuals if there is no evidence of LAM on their baseline HRCT. • Individuals with lung cysts detected on HRCT should have annual pulmonary function testing and HRCT interval reduced to every 2-3 yr. • mTORi may be used to treat LAM patients with moderate to severe lung disease or rapid progression.

Speciality Area	Recommendation
Skin	• Perform a detailed clinical dermatologic inspection/exam annually. • Patients and families should be counselled to use sunblock (SPF 30) routinely. • Symptomatic TSC-associated skin lesions should be treated as appropriate for the lesion and clinical context by surgical excision, laser(s) & topical mTORi.
Teeth	• Obtain a detailed clinical dental exam at minimum every 6 months and panoramic radiographs by age 7 yr, if not performed previously. • Symptomatic or deforming dental lesions, oral fibromas, and bony jaw lesions should be treated.
Heart	• Obtain an Echocardiogram every 1-3 yr in asymptomatic paediatric patients until regression of cardiac rhabdomyomas is documented. • Obtain ECG every 3-5 yr in asymptomatic patients of all ages to monitor for conduction defects. More frequent or advanced diagnostic assessment, such as ambulatory and event monitoring, may be required for symptomatic patients.
Eye	• Perform annual ophthalmologic evaluation in patients with previously identified ophthalmologic lesions or vision symptoms at the baseline evaluation. More frequent assessment, including those treated with vigabatrin, is of no proven benefit and not recommended in the absence of clinical concerns.

The clinical presentation of TSC is variable, with the progression and severity of organ involvement differing according to the individual's age, genotype and treatment. Although, both International and UK guidelines for surveillance and treatment have been proposed, these recommendations are variably followed across different centres dependent on the treating physician, resource availability and local and national policies and funding agreements (Amin et al. 2019). However, the guidelines do provide a safe framework for surveillance and management of TSC individuals in the absence of TSC-specific clinical expertise. The differences between published guidelines also highlight the subtle differences in clinical opinions across the world and reflect the frequent lack of quality data underpinning these recommendations.

1.3 The genetic basis of Tuberous Sclerosis Complex

1.3.1 TSC1 and TSC2

TSC1 (OMIM #605284) and *TSC2* (OMIM #191092) were identified in 1997 and 1993, respectively (1993; Consortium 1993; van Slegtenhorst et al. 1997) as the genes mutated in tuberous sclerosis. The *TSC1* gene on chromosome 9 consists of 23 exons, the first two of which are untranslated and the second is alternatively spliced. Hamartin or TSC1, the protein product of *TSC1* consists of 1164 amino acids. The *TSC2* gene on chromosome 16 consists of 41 coding exons and a noncoding leader exon, and exons 25, 26 and 31 are alternatively spliced. The protein product, tuberin, or TSC2 has a major isoform of 1807 amino acids.

Studies of the *TSC1* and *TSC2* genes in patients with tuberous sclerosis have revealed a wide spectrum of mutations, but there are no particular regions within the genes in which mutations occur at a higher rate. Tuberous sclerosis-associated *TSC2* mutations include missense, nonsense and frameshift deletions/insertions and splice junction mutations. Also, significant numbers of large (exonic and whole-gene) deletions have been reported (Jones et al. 1999a; Dabora et al. 2001; Sancak et al. 2005; Au et al. 2007; Kozlowski et al. 2007). *TSC1* mutations in tuberous sclerosis, on the other hand are primarily small deletions and insertions and nonsense mutations. Pathogenic *TSC1* missense changes are rare.

Approximately 10-20% of patients who meet the diagnostic criteria for TSC do not have any identifiable mutations. These individuals have 'No Mutation Identified' (NMI) after thorough conventional molecular diagnostic assessment. The NMI cohort has a lower incidence of neurological features and renal findings than those with *TSC2* mutations but a higher incidence of renal angiomyolipomas and pulmonary lymphangioleiomyomatosis than tuberous sclerosis patients with *TSC1* mutations (Camposano et al. 2009). A large proportion of the NMI population has been identified as having mosaicism and intronic mutations in *TSC1* and *TSC2* using next-generation sequencing techniques (Tyburczy et al. 2015).

Mutations in *TSC2* are about five times more common than mutations in *TSC1* in the sporadic tuberous sclerosis population, whereas the ratio is approximately 1:1 in large families with multiple generations affected. Patients with mutations in *TSC2* are more likely to have a higher number and/or

more severe clinical features than those with mutations in *TSC1* (Dabora et al. 2001; Sancak et al. 2005).

1.3.2 Modulation of the mTOR pathway by TSC1/TSC2

The mechanistic target of rapamycin (mTOR, also known as mammalian target of rapamycin) pathway plays a vital role in the ability of cells to sense intracellular and extracellular conditions and to mount appropriate physiological responses (Kim et al. 2013). The TSC1/2 protein complex is a central negative regulator of mTORC1 activity. Several years after discovering TSC1 and TSC2, their intracellular role in regulating the mTOR pathway was identified (Manning et al. 2002). These mechanisms are described in detail in section 1.5.

1.3.3 mTOR hyperactivation and brain pathology in TSC

Cortical tubers, the characteristic brain pathology in TSC, are developmental cerebral cortex abnormalities that histologically appear as loss of the cortex's classical six-layered structure along with Dysmorphic neurons (DN), large astrocytes, and typical cells known as Giant Cells (GC) (Mizuguchi and Takashima 2001). SEGAs are tumours of mixed glial and neuronal lineage. Histologically, SEGA is composed of three types of cells: spindle, gemistocytic and ganglion-like cells. SEGA have areas of calcification and vascular components (Buccoliero et al. 2009). SEN are asymptomatic hamartomas arising from the walls of the lateral and third ventricles.

A diverse group of disorders called the mTORopathies (TSC, FCD, HME and GG) present with epilepsy and have underlying perturbation of mTOR pathway regulation (Crino 2011). The mTORopathies have (common) pathological changes of Balloon Cells (BC) and Dysmorphic Neurons (DN) (Curatolo and Moavero 2013) along with evidence of mTOR hyperactivity (Crino 2015).

The role of mTOR hyperactivity in SEGA growth was demonstrated by gene expression analysis, immunohistochemistry and by the effective modulation of SEGA growth by pharmacological inhibition of mTOR (Krueger et al. 2010a; Tyburczy et al. 2010).

1.4 TSC Associated Neuropsychiatric Disorders

1.4.1 Neurocognitive deficits in Tuberous sclerosis

Individuals with TSC have a high frequency of neurocognitive and neurodevelopmental problems (Prather and de Vries 2004; Crino et al. 2006; de Vries et al. 2007). Historically, these included 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 ten times higher than population expectations). The majority of individuals, even those with average or above-average intellectual ability, have specific neuropsychological deficits of attention, executive or memory skills. For example, in a study of 21 normal IQ adults with TSC, specific neuropsychological deficits (performance in the bottom 2% of the population) were identified in 20/21 (Tierney et al. 2011).

1.4.2 Possible underlying mechanisms of TAND phenotypes

TAND phenotypes include autism, which is associated with brain changes in the temporal lobes and cerebellum in TSC individuals. An increased glucose uptake can be demonstrated on positron emission tomography in these brain areas. These findings indicate both cortical and subcortical dysfunctional circuits (Bolton and Griffiths 1997; Weber et al. 2000; Asano et al. 2001). Autism is more common in TSC individuals with cognitive impairment than individuals with IQ in the normal range (Smalley 1998; de Vries et al. 2007).

Intracellular over-activity of mTORC1 in the brain has been postulated as an etiologic factor in the neurocognitive deficits in patients with TSC (de Vries and Howe 2007). This model proposes that the structural brain lesions and epilepsy in TSC have a lesser role in neurocognitive manifestations of TSC than previously assumed. Mouse models (heterozygous for *TSC1* or *TSC2* mutations) demonstrate spatial working memory and socialisation deficits with a relatively normal brain structure and without epilepsy (Goorden et al. 2007; Ehninger et al. 2008b). If molecular irregularities lead to neurocognitive manifestations, then targeting these processes could potentially be used to improve neurocognition (de Vries 2010).

1.4.3 Pre-clinical studies in TAND

Ehninger *et al*. (2008) presented a study of *Tsc2*(+/-) mice, which demonstrated cognitive deficits including impairment in spatial learning and contextual discrimination, in the absence of neuropathology or seizures, supporting alternative disease mechanisms. There was evidence of hyperactive mTOR signalling in the hippocampal CA1 region correlating to deficits in hippocampal-dependent learning. Sirolimus treatment of adult mice partially reversed these cognitive deficits. In addition, sirolimus treatment rescued synaptic plasticity markers (LTP-long-term potentiation) in hippocampal slices. The results suggest a biological mechanism for some cognitive deficits in TSC and that treatment with an mTOR inhibitor ameliorates cognitive dysfunction in a mouse model (Ehninger et al. 2008a). In another study, sirolimus treated TSC 1 null (-/- homozygous) mice showed improvement in seizures, reversal of brain pathology and increased survival (Meikle et al. 2008).

Sato *et al*. (2012) reported that haploinsufficiency of *Tsc1* or *Tsc2* leads to hyperactive mTOR signalling and impaired social interaction in adult mice (Tsc1+/− and Tsc2+/−). Intracellular markers showed an aberrant mTOR signalling, reduced mRNA and increased downstream protein levels, which were reversed with sirolimus. The study indicated a pathogenic role of abnormal mTOR signalling in social deficits, which responded to mTOR inhibition (Sato et al. 2012).

TSC/mTOR signalling is involved in cell proliferation, synaptogenesis, and growth of dendrites and axons. This suggests that any treatment aimed at correcting problems attributable to these roles should be initiated at an early age. However, in the mature brain, role of TSC/mTOR signalling include the regulation of brain plasticity that is required to preserve an adequately functioning brain (Tang et al. 2002; Kelleher III et al. 2004). Synaptic plasticity in the hippocampus is necessary for learning and memory aspects of brain function (Martin et al. 2000).

Another study looking at the morphological implications of *Tsc1/ Tsc2* mutations on neurons reported that in post-mitotic hippocampal pyramidal neurons of mice and rats, loss of Tsc1 or Tsc2 triggered enlargement of somas and dendritic spines and altered the properties of glutamatergic synapses. Loss of a single copy ofTsc1 was sufficient to results in defects in neuronal morphology. This study provides evidence that cell-autonomous neuronal defects due to haploinsufficiency of *Tsc1* / *Tsc2* contributes to the neuronal structure and function which in addition to cortical tubers contribute to the pathogenesis of the neurological presentation of TSC (Tavazoie et al. 2005)

1.4.4 Clinical studies in TAND

The use of mTOR inhibitors has been proposed to regulate the disinhibited mTOR signalling in ASD associated with TSC and related disorders (such as Lhermitte-Duclos Disease and Cowden Disease) (Ehninger and Silva 2011). In the TESSTAL trial, 7 of 8 neurocognitively tested patients with tuberous sclerosis showed an increase in immediate recall memory scores during treatment of their renal AMLs with sirolimus . By contrast, immediate recognition memory scores fell in 5, and none showed an increase. Executive function scores increased in 5 of the 8 patients. The small patient numbers and nonrandomised design limited the usefulness of the neurocognitive assessments in this trial (Davies et al. 2011a)

1.5 *The mTOR pathway: An intracellular signalling network*

The ability to sense and integrate diverse signals leading to modulation of physiological responses is essential for cellular function. Eukaryotic cells integrate a complex intracellular signalling molecular network assimilating the effects of growth factors, energy, nutrient supply and environmental cues to control cell growth, balance catabolic and anabolic processes, determine aspects of neuronal differentiation and migration, and modulate memory and learning (Ehninger et al. 2008b; Huang and Manning 2008).

The mTOR pathway (Figure 2) is an evolutionarily conserved pathway that plays a central role in integrating environmental cues in the form of growth factors, amino acids, and energy to control multiple cellular functions, allowing a cell to balance catabolic and anabolic processes.

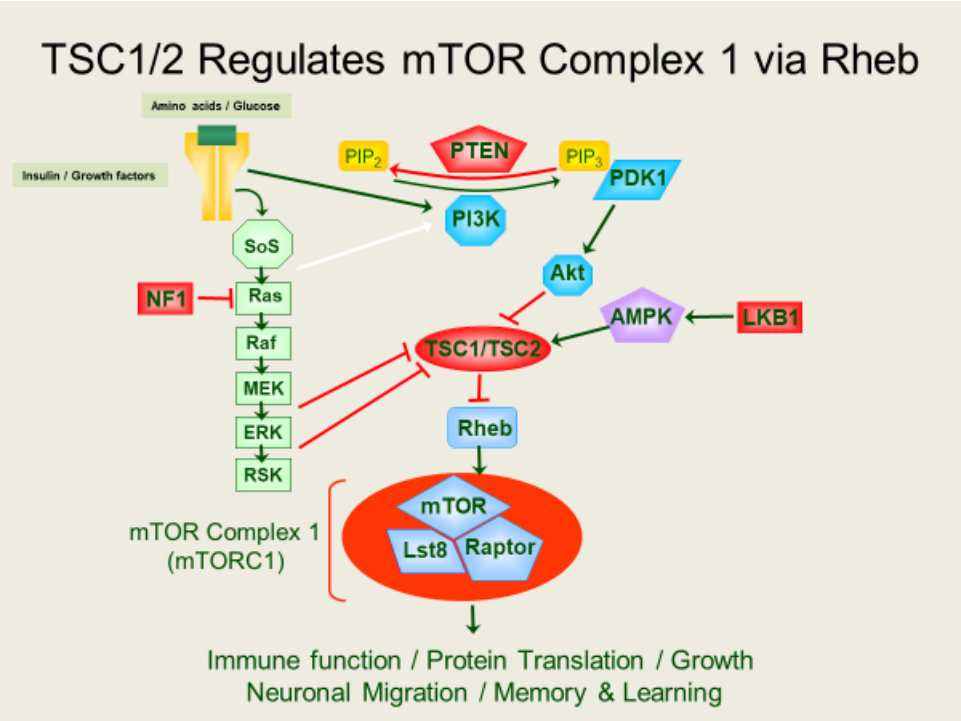


FIGURE 2 Simplified scheme of mTOR pathway signalling via mTORC1. This pathway integrates multiple environmental signals in the control of key cell processes

AMPK – AMP-activated protein kinase; AkT – protein kinase B (a.k.a Akt); ERK - extracellular receptor kinases; LKß1- Liver kinase ß1; mTOR – mammalian target of rapamycin; Lst8 – Lethal with SEC13 protein8; MEK- Mitogen-activated protein kinase; NF1 - Neurofibromatosis type I; PI3K – phosphatidylinositide-3 kinase; PTEN - phosphatase and tensin homolog deleted on chromosome ten; PIP_2- phospholipid phosphatidylinositol (4,5) bisphosphate; PIP_3 - phosphatidylinositol (3,4,5) trisphosphate; PDK1 -3-phosphoinositide-dependent protein kinase-1; Ras- rat sarcoma protein; Raf -Rapidly Accelerated Fibrosarcoma kinase; Rheb – Ras homolog expressed in brain; RSK- ribosomal S6 kinase; Raptor -Regulatory associated protein of mTOR; SoS – Son of Sevenless protein; TSC1/ TSC2 – tuberin/hamartin protein complex

The serine/threonine kinase, mTOR, is a member of the phosphoinositide 3-kinase related kinase (PIKK) family. In response to environmental or physiological stimuli, multiple upstream pathways involving cascades of protein kinases (Figure 2), may either activate (PI3K/Akt, ERK) or inhibit (LKβ1/AMPK) mTOR via modulation of the tuberin-hamartin complex and Rheb GTPase. In anabolic states (growth factor, nutrient stimulation or insulin), PI3K/Akt and ERK activate the mTOR pathway to induce protein synthesis, cell growth and proliferation. Equally, in catabolic states (hypoxia or energy/nutrient deprivation), inhibition of the mTOR pathway slows protein synthesis, metabolism and cell growth (Wong 2010). In the disease manifestations of TSC, mutation of one of the *TSC* genes leads to disinhibition or hyperactivation of the mTOR pathway, causing dysregulated growth and proliferation and predisposing to overgrowth.

1.5.1 mTOR complexes 1 and 2

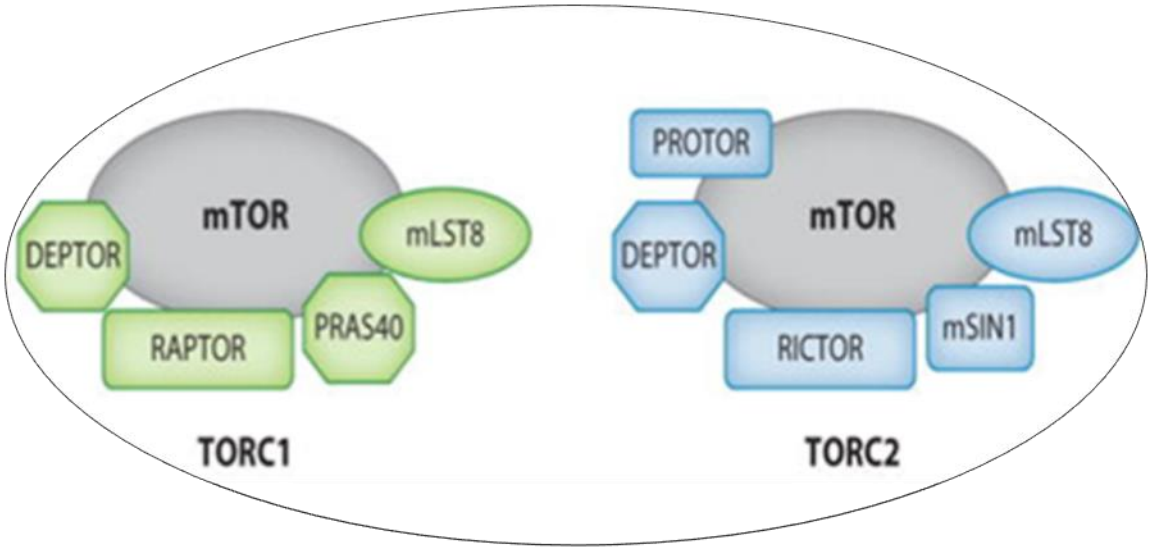


FIGURE 3 Components of mTORC1 & mTORC2

(from Powell, Pollizzi et al. 2012)

In mammalian cells, mTOR forms two structurally and functionally distinct complexes, mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2), each with unique substrate specificity (Figure 3) (Rosner et al. 2008).

The mTORC1 signalling complex (Figure 3) consists of the regulatory-associated protein of mTOR (RAPTOR) and mammalian lethal with Sec13 protein 8 (mLST8). These adapter proteins and mTOR mediate protein-protein interactions (Hara et al. 2002; Kim et al. 2002; Loewith et al. 2002; Kim et al. 2003; Saucedo et al. 2003; Tee et al. 2003; Zhang et al. 2003). The proline-rich Akt substrate 40 kDa (PRAS40) and DEP domain-containing mTOR-interacting protein (DEPTOR) inhibit mTORC1 activity.

The second mTOR containing protein complex, mTORC2 (Figure 3), can also associate with DEPTOR and mLST8. This complex is distinguished by the adapter protein RAPTOR-independent companion of TOR (RICTOR), and protein observed with RICTOR (PROTOR). Another unique component of mTORC2 is mSIN1, which contains a pleckstrin homology domain that is thought to target TORC2 to the membrane, where it can activate myristoylated Akt (Hara et al. 2002; Kim et al. 2002; Loewith et al. 2002; Jacinto et al. 2004; Sarbassov et al. 2004; Frias et al. 2006; Jacinto et al. 2006; Yang et al. 2006; Pearce et al. 2007; Woo et al. 2007).

1.5.2 Functions of the TSC1/TSC2/mTORC1 pathway

The TSC1/TSC2/mTORC1 pathway combines many anabolic and catabolic functions at a cellular level, which are particularly relevant to brain functioning for memory and learning, summarised below. Other processes regulated by mTORC1 include lipogenesis, angiogenesis, glycolysis, autophagy and inflammatory responses (Yecies and Manning 2011; Laplante and Sabatini 2013; Parkhitko et al. 2014).

mTORC1 promotes the transition from G1 to S phase and entry into mitosis, promoting cell growth. In addition, it has a vital role in regulating macromolecule synthesis, such as proteins and lipids required for cell growth (Duvel et al. 2010).

The catabolic processes mediated by the mTOR pathway include regulation of autophagy, cell senescence and stem cell depletion. Autophagy, a process to remove dysfunctional cell components, is up-regulated in mammalian cells under certain cellular stress conditions (nutrient restriction/hypoxia) to sustain anabolic processes and energy production. Activation of mTOR inhibits autophagy, while inhibition of mTOR induces autophagy. mTORC1 hyperactivity leads to stem cell depletion due to stem cell senescence (Yilmaz et al. 2006; Castilho et al. 2009).

During brain development, mTOR activity controls protein expression and cellular processes such as neuronal survival, axon growth and navigation, dendritic arborisation, and synaptogenesis (Miller and Kaplan 2003; Jaworski et al. 2005; Kumar et al. 2005). In the adult brain, mTOR dependent processes are crucial for many forms of synaptic plasticity, such as hippocampal mediated long-term potentiation (LTP) (Hoeffer and Klann 2010). Therefore, the mTOR pathways play an essential role in brain structure and the process of learning and memory via protein synthesis-dependent strengthening of synapses.

1.5.3 mTORC1 independent and rapamycin-insensitive functions of TSC1/2

Several functions of the TSC1/TSC2 complex appear to be mTORC1 independent or rapamycin insensitive. AMPK is reported to be activated in *TSC2*-null cells via Rheb independently of its regulation of mTORC1 signalling (Lacher et al. 2010). Loss of *TSC2* is reported to lead to dysregulation of primary cilia development via a mechanism not responsive to mTORi use (Hartman et al. 2009).

1.5.4 mTOR pathway dysregulation in diverse disease states

Key upstream components in the mTOR signalling pathway include phosphoinositide 3-kinase (PI3K), phosphatase and tensin homolog (PTEN), phosphoinositide dependent protein kinase 1 (PDK1), v-akt murine thymoma viral oncogene (AKT) (also known as PKB – protein kinase B), AMP-dependent protein kinase (AMPK), tuberous sclerosis complex 1 (TSC1) and tuberous sclerosis complex 2 (TSC2) (Figure 4). This signalling pathway is dysregulated in diverse disease states including neurodegeneration, diabetes, epilepsy and cancer (Saxena and Sampson 2014).

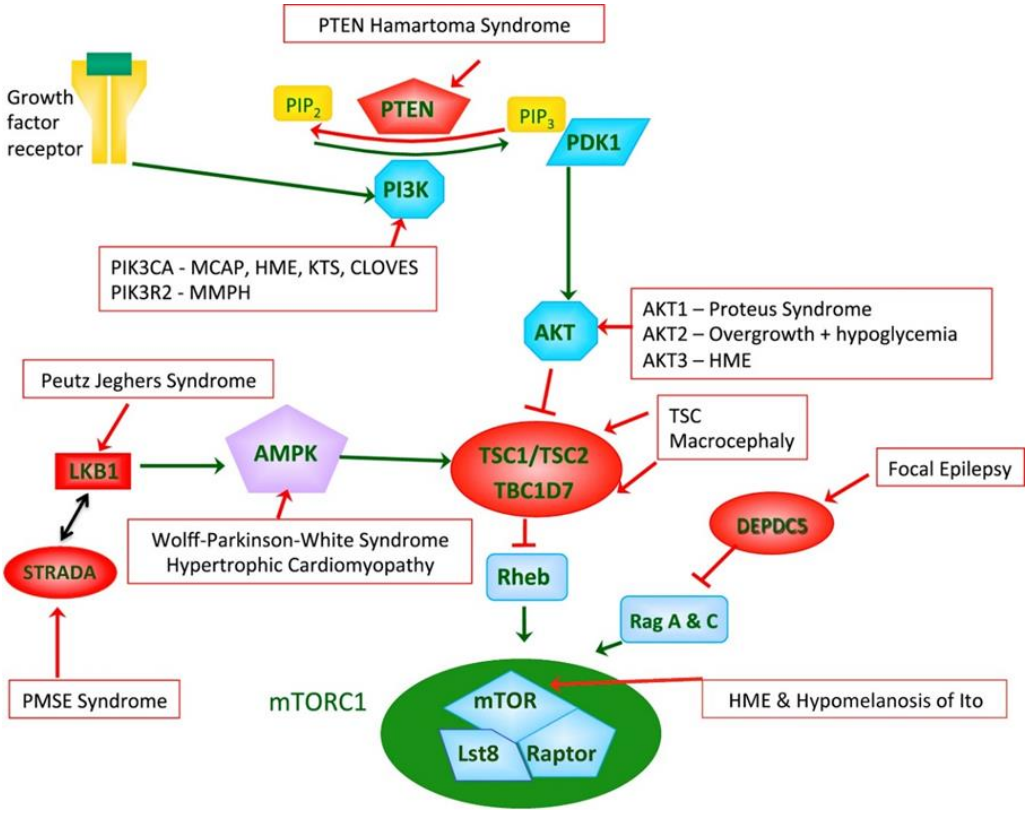


Figure 4 Simplified representation of the signalling pathway upstream of mTORC1

Figure 4 showing disorders associated with mutations affecting the genes encoding pathway components. Clinical disorders/phenotypes are indicated in red boxes. Negative regulators of mTORC1 are shaded red while the positive regulators are shaded blue (Saxena and Sampson 2014)

The TSC1/2 protein complex is a primary negative regulator of mTORC1 activity. Its function is mediated via TSC2's inhibition of Rheb. TBC1D7 (TBC1 domain family, member 7) is a third subunit of the TSC1/2 complex.

Homozygous mutations in TBC1D7 are associated with macrocephaly, intellectual disability and neuropsychological disorders (Capo-Chichi et al. 2013). A reduction in PTEN activity leads to mTORC1 activation. Mutations of PIK3CA have been identified in patients with megalencephaly–capillary malformation syndrome (MCAP) and hemimegalencephaly (HME), whereas activating mutations of PIK3R2 or AKT3 have been identified in most patients with megalencephaly–polymicrogyria–polydactyly–hydrocephalus syndrome (MPPH) and some with hemimegalencephaly (HME) (Mirzaa et al. 2012). Truncating mutations in DEPDC5 (disheveled, Egl-10 and pleckstrin domain containing protein 5) have been described in patients with a broad range of focal epilepsy phenotypes: familial focal epilepsy with variable foci (FFEVF), autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE), familial temporal lobe epilepsy (TLE), benign epilepsy of childhood with centrotemporal spikes (BECTS or Rolandic epilepsy) and cases with focal cortical dysplasia (FCD) (Kurahashi and Hirose 1993; Dibbens et al. 2013). PMSE syndrome (polyhydramnios, megalencephaly, and symptomatic epilepsy), also known as Pretzel syndrome, results from mutations in STRADA (STE20-related kinase adaptor α) results in a reduction in AMPK- (AMP-dependent protein kinase) mediated TSC2 activation and hence mTORC1 activation (Osborne 2010).

Inherited mutations or somatic mutations occurring during organ development cause a spectrum of rare congenital, inherited or developmentally determined conditions as described above (Figure 4). These disorders exhibit highly variable but overlapping clinical features, including localised overgrowth, pigmentary abnormalities, tumour predisposition, cerebral cortical dysplasia, epilepsy and neurodevelopmental disorders. The elucidation of these disorders' mutational bases is leading to new approaches to their treatment, notably using the mTORC1 inhibitor rapamycin (sirolimus) and its derivative everolimus (Crino 2011).

1.5.5 Regulation of the mTORC1 / TSC pathway

mTORC1 acts as a signal integrator for four primary regulatory inputs - growth factors, energy, oxygen and amino acids.

1.5.5.1 Growth factors

Growth factors stimulate mTORC1 through the activation of intracellular signalling pathways. The binding of insulin to its cell surface receptor promotes the tyrosine kinase activity of the receptor, the production of phosphatidylinositol (3,4,5)-triphosphate [PIP3] leading to activation of AKT at the plasma membrane (Figure 2). Activated AKT phosphorylates several downstream substrates and stimulates mTORC1 through TSC1/TSC2 inhibition. AKT directly phosphorylates TSC2, inhibiting the ability of TSC2 to repress mTORC1 activation, possibly by disruption or increased degradation of the TSC1/TSC2 complex or by altered subcellular localisation. Mitogen-activated Ras-ERK signalling (Figure. 2) triggers increased mTORC1 activity via ERK- and RSK-dependent phosphorylation of TSC2 (Roux et al. 2004; Ma et al. 2007).

1.5.5.2 Energy

Depletion of intracellular energy leads to inhibition of mTORC1. The energy status is signalled to mTORC1 by AMP-activated protein kinase (AMPK), a master sensor of intracellular energy status (Fig.2). When intracellular ATP levels decline, and AMP levels increase, which activates AMPK (Shaw et al. 2004). Activated AMPK increases intracellular ATP by up-regulation of catabolic processes, generating ATP, and reduces processes leading to ATP breakdown.

1.5.5.3 Oxygen levels

Hypoxia inhibits mTORC1 by many mechanisms. As oxygen is required for aerobic ATP production via mitochondrial oxidative phosphorylation, hypoxia causes energy stress and AMPK mediated activation of TSC1/TSC2 (Liu et al. 2006). Hypoxia-induced HIF1 activity leads to TSC1/TSC2 complex activation (Brugarolas et al. 2004; Sofer et al. 2005). In response to hypoxia, TSC deficient cells fail to down-regulate mTORC1 activity, exhibit an exaggerated and prolonged increase in HIF activity and continue to proliferate at a rate much higher than their wild type counterparts. These cells also have a tolerance to hypoxia-induced apoptosis (Brugarolas et al. 2004; Sofer et al. 2005; Kaper et al. 2006; Land and Tee 2007; DeYoung et al. 2008).

1.5.5.4 Amino acids

Amino acids, particularly the branched-chain amino acid leucine, represent a strong signal that positively regulates mTORC1. The primary amino acid sensor remains unknown. Intracellular amino acid activation of mTORC1 is known to be independent of TSC1/2 as mTOR remains sensitive to amino acid deprivation in cells that lack TSC1 or TSC2 (Nobukuni et al. 2005). A protein complex called Ragulator, which interacts with Rag GTPases, is necessary for amino acid-dependent mTORC1 activation and mediates a translocation to lysosomal membrane surfaces (Kim et al. 2008a; Sancak et al. 2008).

1.5.5.5 Other regulators of TSC1/2 and of the mTOR pathway

In addition to the critical signals described above, a few other cellular conditions and signalling pathways regulate mTORC1 activity, including genotoxic stress, endoplasmic reticulum stress, and cytokines, to name a few. Over 50 proteins have been shown to interact with TSC1 or TSC2; these interactions' functional significance remains to be fully elucidated (Rosner et al. 2008)

Genotoxic stress reduces mTORC1 activity by several mechanisms. DNA damage activates AMPK, in turn, activating TSC2 (Feng et al. 2005). Oxidative stress in the form of reactive oxygen species (ROS) leads to mTORC1 repression due to activation of TSC2 via LKB1/AMPK (Alexander and Walker).

The endoplasmic reticulum (ER) is an organelle involved in the folding and post-translational modification of proteins. Several physiological and pathological stimuli can cause the accumulation of misfolded proteins in the ER - a condition referred to as "ER stress". When ER stress is not adequately managed by "ER stress response", cell death results (Kim et al. 2008b). ER stress downregulates mTORC1 activity, and in TSC1/2 deficiency, a truncated ER stress response can demonstrate increased sensitivity to ER stress-inducing agents (Ozcan et al. 2008; Di Nardo et al. 2009; Zhou et al. 2009; Kang et al.).

1.5.6 mTORC2

In contrast to mTORC1, relatively little is known about mTORC2. The signalling pathways that regulate mTORC2 are not well characterised. The TSC1/TSC2 complex, whilst inhibiting mTORC1 signalling, promotes mTORC2 activity. Thus, loss of the TSC1-TSC2 complex results in elevated mTORC1 signalling and attenuated mTORC2 signalling. mTORC2 substrates are affected by the loss

of the TSC1-TSC2 complex in cell culture models and kidney tumours from both Tsc2 (+/-) mice and tuberous sclerosis patients (Huang and Manning 2009). Mammalian target of rapamycin complex 2 (mTORC2) is considered a critical downstream mediator of phosphoinositol-3-kinase (PI3K) dependent growth factor signalling. In lymphocytes, mTORC2 has emerged as an essential regulator of cell development, homeostasis and immune responses (Lazorchak and Su 2011).

1.5.7 Mechanism of action of everolimus

Everolimus is a derivative of sirolimus bearing a 2-hydroxyethyl chain. Like sirolimus, everolimus has potent antiproliferative and immunosuppressive effects but with greater stability and solubility as well as favourable pharmacokinetics (Crowe et al. 1999).

The mechanism by which everolimus (sirolimus) exerts effects may depend on the cell type and the mutations that are present. In some cell lines, apoptosis is induced, but in others, the effect is predominantly cytostatic, possibly mediated by inhibition of cell cycle progression at the G1 to S phase. Additionally, in *in vitro* or *in vivo* models, mTOR inhibitors have also been demonstrated to decrease cell size, induce autophagy, promote senescence, inhibit angiogenesis, reduce motility, and selectively target stem cells (Easton and Houghton 2006; Abraham and Eng 2008).

mTORC1 is acutely sensitive to sirolimus, but not all mTORC1 outputs are sirolimus sensitive (Feldman et al. 2009; Garcia-Martinez et al. 2009; Thoreen et al. 2009). Sirolimus weakens the interaction between mTOR and Raptor (Oshiro et al. 2004) and reduces mTORC1 intrinsic kinase activity (Soliman et al.). mTORC2 is not acutely sensitive to sirolimus, but prolonged treatment with the drug can, in some cell types, inhibit mTORC2 activity (Sarbassov et al. 2006).

1.6 Neuroimaging in TSC

In the brain, in addition to TSC lesions of cortical tubers, other brain malformations such as transmantle cortical dysplasia, hemimegalencephaly and schizencephaly have also been described (Galluzzi et al. 2002; Huntsman et al. 2006). These are seen in addition to the classical features of SENs, SEGAs and white matter changes described earlier in section1.2.

Although cranial CT scanning can identify calcification in cortical tubers and SEN better than other imaging modalities, brain MRI is the primary neuroimaging tool in diagnosing and evaluating TSC associated brain lesions in clinical care. Cortical tubers appear as circumscribed areas of signal hypointensities on T1-weighted MRI and signal hyperintensities on T2-weighted (T2-W) sequences (Barkovich 2005). The FLAIR sequence allows the detection of small subcortical and gyral core tubers and white matter lesions (Luat et al. 2007). Subependymal nodules (SENs) are hamartomatous lesions that dot the ependymal surface of the lateral ventricles. The lesions are isointense to grey matter on MR images. SENs calcify in 90% of cases and therefore are easily seen on CT images (Braffman et al. 1992a).

SEGAs are low-grade neoplasms that arise near the foramen of Monro in 10–15% of patients with TSC (Altman et al. 1988). On MRI imaging, these lesions appear similar to SENs. SEGA diagnosis should be considered if the lesion size is ≥ 10 mm in any direction, at the caudothalamic groove or any subependymal lesion at any location that has shown serial growth on consecutive imaging regardless of size. (Clarke et al. 2006; Baskin 2008; Roth et al. 2013).

White matter (WM) lesions, present in more than 80% of TSC patients, feature four distinct patterns seen on MR imaging: 1) Radially extending straight or curvilinear bands stretching from the ventricle through the cerebral WM toward the cortex, 2) wedge-shaped lesions, 3) nonspecific conglomerate foci, and 4) cerebellar radial bands (Braffman et al. 1992a). The radial bands represent heterotopic neuronal and glial elements with impaired cortical migration. They radiologically have similar signal characteristics as cortical tubers (low signal on T1-W images and high signal on T2-W images).

White matter abnormalities seen by MR imaging in patients with tuberous sclerosis include superficial white matter abnormalities associated with cortical tubers, radial white matter bands and cyst-like white matter lesions (Makki et al. 2007b). Superficial white matter abnormalities probably reflect reduced myelin or increased gliotic reaction related to cortical tubers, while radial white matter bands reflect developmental migration defects of neurons and glial cells. White matter cyst-like lesions

probably reflect cystic degeneration of white matter or dilated perivascular spaces. WM changes also involve nonspecific focal conglomerate changes and cerebellar radial bands (Braffman et al. 1992b).

1.6.1 Scientific basis of Diffusion tensor imaging (DTI)

DTI is a variant of conventional Magnetic Resonance Imaging, sensitive to water movements within the tissue's architecture. DTI does not require additional equipment, contrast or chemical tracers (Le Bihan et al. 1991). Diffusion Tensor Imaging (DTI) is a microstructural magnetic resonance imaging (MRI) technique used to examine the white matter tracts between different brain regions. DTI estimates the principal diffusivities parallel and perpendicular to fibre bundles (Basser et al. 1994b).

1.6.2 Concept of molecular diffusion

Molecular diffusion refers to the random translational motion of molecules (Brownian motion) resulting from these molecules' thermal energy (Furth and Cowper 1956). The movement of molecules in an unconstrained environment is presumed to be isotropic (uniform) in all directions. In contrast, an anisotropic pattern would be expected if these movements are limited in one direction more than others due to tissue properties. In a free medium, water molecules at 37 °C, will diffuse 17 μm during 50 ms. However, Diffusion MRI (dMRI) techniques observe that water molecules, on average, move in brain tissues over distances of around 1–15 μm, during comparable diffusion times (Le Bihan 2003). This slowing is explained by the molecules interacting with tissue components, such as cell membranes, fibres and macromolecules, providing *in vivo* clues to the fine structural features and organisation of neural tissues, both in healthy and pathological states.

Diffusion MRI measures diffusion by using a pair of magnetic field gradient pulses(Stejskal and Tanner 1965). In an otherwise homogeneous magnetic field, the first pulse magnetically 'labels' (detectable magnetic field perturbations) hydrogen nuclei (or protons). The second pulse measures the displacement of nuclei during the time interval (or 'diffusion time') between the two pulses. This displacement (movement) of a hydrogen nucleus carried by a diffusing water molecule results in the molecule experiencing spatially varying magnetic field strength, proportional to the displacement. A method to estimate the principal diffusion direction in tissues, using the techniques of NMR spectroscopy, was described by Basser *et al*., in 1994, which forms the basis of approximating the fibre orientation in tissues. The 3D displacement profile is represented by a 3x3 symmetrical matrix – the diffusion tensor. The principal fibre orientation is given by the primary eigenvector, which corresponds to an orientational axis along which there is the least hindrance to diffusion (Basser et al. 1994a).

1.6.3 Diffusion Tensor Imaging

Although water diffusion is a three-dimensional process, water molecules' mobility in white matter tissues is dissimilar (anisotropic) in all directions. Water molecules diffuse more easily along the central axis of a white matter fibre bundle than perpendicular to it (Moseley et al. 1990). This coherent anisotropic movement of water molecules is a central principle for DTI analysis. Diffusion anisotropy in white matter originates mostly from this tissue organisation as bundles of myelinated fibres running in parallel; diffusion in the fibres' direction is approximately three to six times faster than in the perpendicular direction (Le Bihan 2003).

1.6.4 Tractography

Tractography can be used to reconstruct a three-dimensional representation of white matter pathways (Conturo et al. 1999; Jones et al. 1999b; Mori et al. 1999; Basser et al. 2000; Poupon et al. 2000) by sequentially piecing together discrete and shortly spaced estimates of fibre orientation to form continuous trajectories.

The tractography technique has been used to study human brain thalamic connections (Behrens et al. 2003), occipitotemporal connections, and other white matter tracts (Mori et al. 1999; Catani et al. 2002; Lehericy et al. 2004). Although the fibre tracts dissected are virtual and require further anatomical validation, information regarding these brain connections and their functioning cannot be acquired directly. Therefore, we must rely on surrogate techniques such as DTI to infer data about white matter pathways.

1.6.5 DTI measures

DTI measures used to investigate microstructural tissues are fractional anisotropy (FA), mean diffusivity (MD) or apparent diffusion coefficient (ADC), radial (perpendicular) diffusivity, and axial (parallel) diffusivity (Pierpaoli and Basser 1996b). We have analysed FA and MD for TRON study participants, and therefore the subsequent discussion will be limited to these DTI measures.

FA quantifies the anisotropy of white matter tracts and is highly sensitive to change in microstructure; however, it is nonspecific to the cause of change. Mean diffusivity is sensitive to cellular and membrane density where an increase in mean diffusivity indicates neural problems such as oedema or necrosis(Le Bihan et al. 2001; Arfanakis et al. 2002). Diffusion in white matter (WM) is

less restricted along the axon as it tends to be anisotropic (directionally-dependent). In contrast, in grey matter (GM), it is usually less anisotropic and in the cerebrospinal fluid (CSF) is unrestricted in all directions (isotropic, FA value = 0).

Usually, a higher MD and lower FA values indicate damaged or impaired fibre integrity due to increased diffusion and loss of coherence on preferred movement direction(Soares et al. 2013). However, crossing fibres may in some of the tracts lead to increased FA even in disease states such as Alzheimer's disease (Douaud et al. 2011).

1.6.6 DTI techniques

FA and MD data are analysed by whole-brain analysis or tractography based on a region of interest (ROI) analysis, and recently described tract-based spatial statistics (TBSS) (Smith et al. 2006). A region can be defined by anatomical structures (e.g., corpus callosum), pathology (e.g., tumour) or geometry (sphere or cube) (Froeling et al. 2016).

1.6.6.1 Voxel-Based Morphometry (VBM)

A voxel is (a volume element – 3D) the region in a tissue slice that corresponds to a pixel (a picture element – 2D) for a given slice. It is the basic unit of a CT or MRI reconstruction. VBM has been used to evaluate the FA values between two groups of subjects in patients with varied aetiologies such as bipolar disorder (Selvaraj et al. 2012), Alzheimer disease(Busatto et al. 2008), Schizophrenia (Buchsbaum et al. 1998; Honea et al. 2005), chromosome 22q11.2 deletion syndrome (Simon et al. 2005), and epilepsy (Rugg-Gunn et al. 2001).

VBM approach has known issues with alignment as data from a voxel cannot be guaranteed to contains data from anatomically corresponding region in every subject. Smoothing can help ameliorate residual misalignments, though not in a well-controlled way. The need for spatial smoothing, and the problem of arbitrarily choosing the smoothing extent, is a severe limitation of VBM-style approaches. Jones et al. reported that the final results of VBM-style FA analysis of schizophrenia data depend very strongly on the amount of smoothing (Jones et al. 2005b).

1.6.6.2 Tract-based spatial statistics (TBSS)

The TBSS technique attempts to bring together the strengths of VBM while aiming to solve the alignment and smoothing issues. It has the advantage of being fully automated and investigating the

"whole" brain as it does not require pre-specification of tracts of interest. This is achieved by estimating a 'group mean FA skeleton", representing the centres of all fibre bundles common to the subjects involved in a study. It is presumed that the voxel with the highest FA in the local vicinity represents the centre of the tract. Each subject's FA data is then projected onto the mean FA skeleton so that each skeleton voxel takes the FA value from the local centre of the nearest relevant tract, resolving issues of alignment and correspondence (Smith et al. 2006).

TBSS technique could be summarised in the steps below

- Identify a common registration target and align all subjects' FA images to this target using nonlinear registration. Perfect alignment is not expected or required at this stage.
- Create the mean of all aligned FA images and apply "thinning" (suppression perpendicular to the local tract structure) to create a skeletonised mean FA image. Project each subject's (aligned) FA image onto the skeleton by filling the skeleton with FA values from the nearest assumed (highest FA) tract centre.
- Carry out voxelwise statistics across subjects on the skeleton space FA data.

TBSS has been used to analyse patients with various white matter diseases such as schizophrenia, Parkinson's disease and temporal lobe epilepsy where it was found to be more accurate than the VBM based techniques (Afzali et al. 2011; Rae et al. 2012) ((Smith et al. 2006). The study authors also reported that the TBSS technique has less variability than VBM & hand-drawn tractography techniques when assessed across sessions and subjects in healthy subjects.

1.6.6.3 Region of Interest-based (ROI) Tractography

ROI analysis is usually carried out by hand, separately for each subject (Kubicki et al. 2003). FA values are taken from the ROI(s) and then compared across subjects. This is a reliable technique in the centres of the largest tracts. However, it can be hard to place ROIs for smaller, thinner tracts objectively. Besides, this approach limits a study to only being sensitive to change in those few parts of the brain where ROIs are placed.

Another approach uses fibre bundle tracking to identify voxels from which to take FA values for cross-subject comparison(Jones et al. 2005a; Jones et al. 2006). In such approaches, the relevant tracts are usually identified by initialising/constraining tractography using hand-drawn ROIs.

A limitation of tractography using DTI data is that the tensor data cannot adequately represent multiple fibre orientations, leading to erroneous tract reconstructions and inappropriate ROI analysis.

1.6.7 Clinical application of DTI imaging techniques

Many clinical studies on patients with white matter diseases have shown the exquisite sensitivity of DTI indices to characterise these changes in terms of white matter fibres' integrity at an early stage. Examples include leukoencephalopathies (Ay et al. 1998; Eichler et al. 2002), human immunodeficiency virus-1 infection (Filippi et al. 2001), Alzheimer's disease (Hanyu et al. 1997) and CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy) (Chabriat et al. 1999).

DTI indices could also unravel subtle, functional disorganisation that may not necessarily be visualised into anatomical anomalies in T1 and T2 (structural) MRI scans, such as in patients with psychiatric disorders (Lim and Helpern 2002). In TSC patients, DTI studies have uncovered widespread microstructural disorganisation in otherwise normal-appearing white matter (NAWM) (Makki et al. 2007a; Arulrajah et al. 2009). These changes in the frontal and parietal lobes have been reported to correlate with attention deficits in children with TSC (Peng et al. 2004).

1.6.8 Choice of tracts for the TRON imaging study

We chose to analyse the DTI indices (FA and MD) for fibre tracts of Fornix (FX), Uncinate fasciculus (UF), Cingulum (CG) and the three components of Superior longitudinal fasciculus (SLF).

At the trial design stage, no studies had investigated the effect of everolimus on the longitudinal changes in DTI indices and neuropsychological parameters in a TSC population, as was intended in the TRON study (Randell et al. 2016). These tracts were chosen pragmatically, looking at literature from other clinical conditions, where assessments of memory and executive dysfunction were investigated and compared with DTI indices. A few examples are given in the next paragraph.

Alzheimer’s Disease patients had widespread lower FA and higher MD than controls in the corpus callosum, anterior commissure, uncinate fasciculus, cingulum bundle and SLF. Significant differences across groups (AD vs control) in the correlations between diffusion indices (MD) and neuropsychological scores were found in the SLF (Douaud et al. 2011). In another study, recent-onset schizophrenia patients had shown deficits in frontal-parietal connections, with a lower FA value than control subjects across the entire SLF, with particular deficits in the left SLF (Karlsgodt et al. 2008). Higher FA values in SLF (mainly the SLF II) were associated with better performance on the 'n-back' test used to assess young German adults (Burzynska et al. 2011).

A significant correlation between the fornix measurements (FA and MD) was also found with memory z scores to predict memory decline and progression to Alzheimer's disease (Mielke et al. 2012). In Alzheimer's disease, white matter tract disintegrity was evident in the UF, posterior cingulum and fornix compared with controls (Juh et al. 2012). In behaviour-variant frontotemporal dementia (bvFTD) the reduction in the value of left anterior cingulum FA was related to executive function, the right anterior cingulum FA to visual-spatial attention and working memory, the right posterior cingulum to visual-constructional abilities, and the left UF FA to executive skill dysfunction (Tartaglia et al. 2012). Fornix FA values were significantly reduced in patients with multiple sclerosis and identified as a predictor of visual recall problems (Dineen et al. 2012). In the fornix and the Cingulum, FA reductions in the white matter tracts were found in a cohort with Schizophrenia and working memory dysfunction (Sugranyes et al. 2012).

While developing the study protocol for the current work, three studies evaluated DTI indices in TSC populations, with one reporting serial assessments after treatment with everolimus. In the first study, DTI indices (FA, AD, MD, RD) for the splenium of the corpus callosum, internal capsule, superior temporal gyrus, and geniculocalcarine tracts (regions involved in the processing of visual, auditory and social stimuli) were assessed. In general, the FA and AD were reduced in TSC patients. (Krishnan et al. 2010). In a second study, the same group in a later study reported a significant reduction in FA values and higher MD, RD & AD values of the corpus callosum in the TSC population than controls (Peters et al. 2012). The third study reported longitudinal DTI indices changes in 21 children treated with everolimus for 12-18 months for SEGA related to TSC. The authors have reported significant improvement in FA (increased) in the corpus callosum, anterior and posterior limb of the internal capsule, and in the geniculocalcarine tract as compared to baseline. The MD reduced significantly only in the corpus callosum. The control group parameters were unchanged. The study's primary outcome was a reduction in SEGA size; neurocognitive outcomes were not reported (Tillema et al. 2012).

1.7 Pharmacology of Everolimus

1.7.1 Introduction

Everolimus is a derivative of rapamycin which exerts its activity through a high-affinity interaction with the intracellular receptor protein (FK506 binding protein) FKBP12 (Kirchner et al. 2004). Everolimus affects cell growth and proliferation, contributing to its anti-tumour and immunosuppressive properties.

1.7.2 Preclinical data

In safety pharmacology studies, everolimus was devoid of relevant effects on vital organ functions, including the cardiovascular, respiratory and nervous systems. Everolimus had no impact on the QT interval. Furthermore, everolimus showed no antigenic potential. Although everolimus passes the blood-brain barrier, there was no indication of relevant changes in rodents' behaviour, even after single oral doses up to 2000 mg/kg or repeated administration at up to 40 mg/kg/day.

The preclinical safety profile of everolimus was assessed in mice, rats, minipigs, monkeys, and rabbits. There was no indication of kidney toxicity in monkeys or minipigs. Genotoxicity studies showed no evidence of clastogenic or mutagenic activity. For up to 2 years, administration of everolimus did not indicate any oncogenic potential in mice and rats up to the highest doses, corresponding respectively to 3.9 and 0.2 times the estimated clinical exposure (Novartis 2019).

1.7.3 Human safety and tolerability data

Phase I dose-escalating studies, exploratory Phase I/II studies, Phase II/III studies of everolimus in cancer indications as a single agent or in combination with other anti-cancer agents have contributed to an extensive database of human data. Approximately 12,700 patients were treated with everolimus as of 30-Sep-2010.

1.7.4 Teratogenicity data

In reproduction studies, everolimus was toxic to the conceptus in rats and rabbits and was considered potentially teratogenic in rats (Hentges et al. 2001). The potential risk for humans is unknown. Everolimus should be given to pregnant women only if the potential benefit to the mother

justifies the potential risk to the foetus. Therefore, it is recommended that women of childbearing potential use adequate contraceptive measures during treatment with everolimus.

1.7.5 Pharmacokinetics

Everolimus (known initially as RAD001) is rapidly absorbed with a median Tmax of 1-2 hours. The steady-state AUC0-τ is dose-proportional over the dose range between 5 to 70 mg in a weekly regimen and 5 and 10 mg daily. Steady-state was achieved within two weeks with a daily dosing regimen. Cmax is dose-proportional between 5 and 10 mg for both the weekly and daily regimens. At doses of 20 mg/week and higher, the increase in Cmax is less than dose-proportional.

Following oral administration, everolimus is the main circulating component in human blood and contributes most to the overall pharmacologic activity. There was a significant correlation between AUC0-τ and pre-dose trough concentration at steady-state on the daily regimen. The mean elimination half-life of everolimus is approximately 30 hours. No specific excretion studies have been undertaken in cancer patients; however, data available from a transplantation setting found the drug to be mainly eliminated through the faeces (Novartis 2019). In adults, everolimus pharmacokinetic characteristics do not differ according to age, weight or sex (Kirchner et al. 2004).

1.7.6 Adverse reactions

Safety data available from completed, controlled, and uncontrolled studies indicate that everolimus is generally well-tolerated in weekly and daily dose schedules. The AEs are usually reversible and non-cumulative. The data on safety in patients with tuberous sclerosis is very limited but has been consistent with that reported in other patient populations (Krueger et al. 2010b). Non-infectious pneumonitis is reported with mTOR inhibitors but is usually low-grade and reversible.

Adverse Events most frequently observed with everolimus are mouth ulcers, rash, infections, non-infectious pneumonitis, fatigue, headache, anorexia, nausea, vomiting and diarrhoea. The most commonly observed laboratory abnormalities include neutropenia, thrombocytopenia, hypercholesterolemia, and hypertriglyceridemia. The majority of reported AEs have been mild to moderate (NCI CTC grade 1-2).

1.7.7 The recommended treatment of mTOR inhibitor-associated adverse events

The recommended management strategies for common adverse events with everolimus are discussed in the following sections.

1.7.7.1 Management of infections

Everolimus is an immunosuppressant. Patients taking Everolimus are, therefore, at an increased risk of infection. In oncology patients, some infections have been severe and rarely have had a fatal outcome. Physicians should warn patients and their caregivers to be vigilant for signs and symptoms of infection and immediately seek medical attention should such signs or symptoms occur. Should an infection occur, anti-infectives should be prescribed as clinically appropriate. In the case of clinically significant infection, consideration should be given to withholding mTORi treatment until the resolution of the infection (Agricola et al. 2013).

1.7.7.2 Management of mouth ulcers/stomatitis/oral mucositis

Stomatitis/oral mucositis/mouth ulcers due to Everolimus should be treated using appropriate, supportive care. Investigators in earlier trials have described the oral toxicities associated with Everolimus as mouth ulcers rather than mucositis or stomatitis. The paradigm below is recommended for the treatment of stomatitis/oral mucositis/mouth ulcers:

1. For mild toxicity (grade 1), conservative measures such as non-alcoholic mouth wash or saltwater (0.9%) mouth wash several times a day until resolution.
2. For more severe toxicity (grade 2 in which case patients have pain but can maintain adequate oral alimentation, or grade 3 in which case patients cannot maintain adequate oral alimentation), the suggested treatments are topical analgesic mouth treatments (i.e., local anaesthetics such as benzocaine,) with or without topical corticosteroids, such as triamcinolone oral paste 0.1% (Kenalog in Orabase®). Direct application of clobetasol, a high potency steroid, has been associated with rapid symptomatic improvement in mTOR-treated patients with aphthous ulceration (Chuang and Langone 2007).
3. Agents containing hydrogen peroxide, iodine, and thyme derivatives may worsen mouth ulcers. It is preferable to avoid these agents.

4. Antifungal agents must be avoided unless a fungal infection is diagnosed. In particular, systemic imidazole antifungal agents (ketoconazole, fluconazole, itraconazole, etc.) should be avoided in all patients due to their potent inhibition of everolimus metabolism, therefore leading to higher everolimus exposures. Therefore, topical antifungal agents are preferred if an infection is diagnosed. Similarly, antiviral agents such as acyclovir should be avoided unless a viral infection is diagnosed (Davies et al. 2017).

1.7.7.3 Management of hyperlipidemia and hyperglycemia

Management of hyperlipidemia should take into account the pre-treatment status and dietary habits of the patient. Blood tests to monitor hyperlipidemia must be taken in the fasting state. Hyperlipidemia and hypertriglyceridemia should be treated according to best clinical practice. GPs should be informed of any incidences of hyperlipidemia in a participant for management. Grade 3 hypercholesterolemia (> 400 mg/dL or 10.34 mmol/L) or grade 3 hypertriglyceridemia (>5 × ULN) should be treated as clinically indicated with a 3-hydroxy-3-methyl-glutaryl (HMG)-CoA reductase inhibitor (e.g. atorvastatin, pravastatin) or appropriate lipid-lowering medication, in addition to the diet. Patients should be monitored clinically and through serum biochemistry for the development of rhabdomyolysis and other Adverse Events as required in the product label/data sheets for HMG-CoA reductase inhibitors (Davies et al. 2017).

Concomitant therapy with fibrates and an HMG-CoA reductase inhibitor is associated with an increased risk of a severe but rare skeletal muscle toxicity manifested by rhabdomyolysis, markedly elevated creatine phosphokinase (CPK) levels and myoglobinuria, acute renal failure and sometimes death. The risk versus benefit of using this therapy should be determined for individual patients based on their risk of hyperlipidemia's cardiovascular complications.

Grade 3 *hyperglycemia* has been observed in patients receiving everolimus therapy. In almost all cases, the affected patients had abnormal fasting glucose at baseline. Based on this finding, it is suggested that optimal glucose control should be achieved before starting a patient on the study drug and that glucose control should be monitored during the trial.

1.7.7.4 Haematological toxicity

Bone marrow suppression is a common toxicity associated with mTOR inhibitors. Grade 1 effects do not require any interruption of treatment. Thrombocytopenia requires intervention at grade 2 and neutropenia at grade 3. Thrombocytopenia and neutropenia are rarely associated with clinically significant bleeding or infection, and hence do not typically necessitate platelet transfusion or growth factor support. Microcytosis and hypochromia have also been reported in patients with TSC treated with mTOR inhibitors; generally, these effects are self-limiting (Cabrera-López et al. 2012). In cases of grade 3 toxicity, interruption of treatment with everolimus is required, lowering the dose upon resumption. Everolimus should be discontinued in any cases of life-threatening toxicity.

1.7.7.5 Renal Adverse events

In patients with TSC with significant renal involvement, although AEs such as an increased degree of proteinuria may be common, they are generally intermittent and should not trigger treatment cessation. Angiotensin-converting enzyme (ACE) inhibitors and angiotensin receptor blockers can be used to ameliorate microalbuminuria or proteinuria when necessary. Cessation of everolimus should be considered if there is progressively increasing proteinuria to >1 g/day, especially if >3 g/day or if associated with peripheral oedema. Similarly, if GFR progressively declines to <30 mL/min, cessation should be considered, although this may be due to the underlying pathophysiology of TSC rather than everolimus (Davies et al. 2017).

1.7.8 Drug interactions

Everolimus is a substrate of Cytochrome P450 3A4 (CYP3A4) and a substrate and moderate inhibitor of P-glycoprotein (PgP). Therefore, absorption and subsequent elimination of Everolimus may be influenced by by-products that affect CYP3A4 and/or PgP. Drugs that are inhibitors or inducers of CYP3A4 and/or PgP should be avoided if possible or used with caution in patients taking everolimus. Inhibitors of CYP3A4 may decrease the metabolism of everolimus and increase its levels, while inducers of CYP3A4 may increase the metabolism of everolimus and decrease levels. Dose adjustment or interruption during therapy may be required with these agents, and, conversely, dose adjustments or cessation of the following agents may be needed during everolimus therapy.

2 A phase II trial of everolimus as a therapy for Neurocognitive problems in patients with Tuberous Sclerosis (TRON clinical trial)

2.1 The rationale of the clinical trial

The pervasive consequences of the neurocognitive problems in individuals with TSC drive families' worries about this aspect of TSC over and above any other organ system involvement (Hunt 1983). These manifestations lead to poor academic performance and significant challenges with employment, long-term relationships, socialisation and peer interactions. Currently, there are no specific approved agents for managing neurocognitive or neurodevelopmental problems in TSC.

Evidence from pre-clinical studies in adult mouse models of TSC that were treated with mTOR inhibitors, revealed a reversal of cognitive deficits and rescue of physiological markers of synaptic plasticity in heterozygous mice ((Ehninger et al. 2008a). Similarly, in a homozygous TSC mouse model, persistent improvement in epilepsy, brain pathology and survival was noted (Meikle et al. 2008). The TESSTAL phase 2 clinical trial results suggested improvement in certain aspects of recall memory and executive function in adult TSC patients being treated with an mTOR inhibitor (sirolimus) for renal angiomolipoma, but involved only a small group of patients and the trial did not have a control group (Davies et al. 2011a).

The TRON (Trial of everolimus for neurocognitive problems in tuberous sclerosis) trial was designed due to an unmet need for effective treatments for neurocognitive problems in the TSC population. Efficacy studies of everolimus were ongoing for kidney and brain tumours in TSC at the time of trial planning.

2.2 Aims and Objectives of This Research - Neurocognition

The TRON clinical trial was designed to investigate the effects of everolimus on deficits in neurocognitive function in adult patients with TSC. The study aimed to determine effect sizes to inform potential development of future Phase 3 trials for this indication.

2.2.1 Primary Objective

The primary objective was to determine the effect sizes of treatment with everolimus or placebo for 6 months on recall memory and executive function in adults with tuberous sclerosis.

2.2.2 Secondary Objectives

The secondary objectives were to assess the effects of treatment with everolimus or placebo for 6 months on broader aspects of neurocognitive functioning, seizures and daily life in people with tuberous sclerosis and to assess safety using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03(NCI 2010)

2.2.3 Exploratory Objective

An exploratory goal was to determine whether an effect of treatment with everolimus or placebo was detectable at 1 month and 3 months after starting therapy to establish whether any early markers of change were present.

2.3 Timeline & approvals

The Wales REC initially approved the TRON study on 9th November 2011. The initial EU clinical trials Register approval was given on 10th January 2012, with the EudraCT Number of 2011-004854-25. The study was registered with ISRCTN on 28th December 2011.

2.4 Participating centres

The study was conducted across three centres in the UK. Cardiff was the main centre at the initiation of the project in 2012, at the Clinical Research Facility (CRF) of University Hospital for Wales. A further two centres opened in 2016 to make participation easier for patients living in Scotland and Northern Ireland, in order to boost recruitment: Glasgow CRF based at QE University Hospital and the Wellcome Trust-Wolfson NI CRF based at the Belfast City Hospital. For the brain imaging study, all DTI scans were done at the Cardiff University Brain Research Imaging Centre (CUBRIC), Cardiff, Wales. Eligible and willing participants from Glasgow and Belfast travelled to Cardiff for the scans that were undertaken at the baseline and 6 month assessments only.

2.5 Inclusion and exclusion criteria

2.5.1 Screening for eligibility

Entry into the study was a two-stage process. Patients (or their carers) contacted the study team, who did a phone-based screening interview and after that offered a screening visit, if appropriate. At the screening visit, clinical assessment and neurocognitive assessments (Table 6), confirmed or refuted eligibility. For inclusion, participants had to fulfil all inclusion criteria and none of the exclusion criteria as detailed in Table 4 and Table 5.

TABLE 4 Inclusion criteria for TRON study participation

Inclusion criteria TRON study	
1	Definite TSC by current clinical criteria
2	Male or female aged 16 to 60 years
Neuropsychological assessments	
3	IQ over 60 and able to participate in direct neuropsychological tests using • Wechsler Abbreviated Scale of Intelligence (WASI) • Edinburgh Handedness Test • NART Error Scale
4	A score ≤ -1.5 SD in one or more of the primary outcome measures as in Table 6
Clinical assessments	
5	Calculated GFR > 60ml/min/1.73m2 except in case of renal impairment associated with TSC, where an estimated GFR should be ≥ 30 ml/min/1.73m2.
6	INR 1.5 or less (anticoagulants use permitted if target INR on a stable dose for > 2 weeks at the time)
7	Adequate liver function as shown by serum bilirubin ≤ 1.5 x ULN, ALT and AST ≤ 2.5 x ULN
8	Seizure free or stable seizures: defined as no change in the type of AEDs in 6 months before full recruitment and randomisation at baseline.
9	Hepatitis B surface antigen-negative, Hepatitis C antibody negative.
10	All patients able to communicate well with the investigator, understand and comply with the study's requirements, understand and sign the written informed consent.
11	Negative pregnancy test in females at the time of informed consent.
12	Contraception: Female patients of childbearing potential to use two acceptable contraception methods. Male participants to use contraception from the time of screening.

TABLE 5 Exclusion Criteria for the TRON study

Exclusion criteria	
1	Prior treatment with an mTOR inhibitor.
2	Investigational agent < 30 days prior to randomisation.
3	Surgery in the last two months.
4	Previous brain neurosurgery (except for SEGA removal surgery or radiosurgery 5 or more years ago).
5	Urine protein/creatinine >0.02g/mmol (except in case of renal impairment associated with TSC, where Urine protein/creatinine > 0.1g/mmol was exclusion criteria).
6	Serum creatinine > 1.5 x ULN (except in case of renal impairment associated with TSC, where Serum creatinine > 300 µmol/L was exclusion criteria).
7	Uncontrolled hyperlipidaemia (fasting cholesterol > 300mg/dL or >7.75 mmol/L and fasting triglycerides >2.5 x ULN), or diabetes with fasting serum glucose > 1.5 x ULN.
8	History of myocardial infarction, angina or stroke or any other significant cardiovascular disease.
9	Lymphangioleiomyomatosis with FEV1 <70% of predicted or any other restrictive pulmonary disease.
10	Significant haematological abnormality i.e. haemoglobin < 8g/dL, platelets <80,000/mm3, Neutrophil count < 1000/mm3.
11	Bleeding diathesis or on oral anti-vitamin K medication other than low dose warfarin.
12	Pregnancy / Lactation.
13	HIV seropositivity, organ transplant, malignancy other than squamous or basal cell skin cancer.
14	Live vaccine required during the trial.
15	Use of strong inhibitor of CYP3AE.
16	Use of strong inducer of CYP3AE except for antiepileptic drugs.
17	Intercurrent infection at the time of randomisation.
18	Inability to complete study materials (outcome measures) in English.
19	History of significant trauma-related cognitive deficit.
20	Impairment of gastrointestinal function that may alter the absorption of everolimus.
21	Known sensitivity to everolimus or other rapamycin analogues or its excipients.
22	Inability to attend scheduled visits.

2.6 Recruitment strategy

The trial was advertised through the Tuberous Sclerosis Association (TSA) via its website and its magazine that is mailed to its members (approximately 1000 families). Clinical geneticists throughout the UK were made aware of the study through the NIHR Genetics Clinical Research Network. All UK Regional Clinical Genetics Services were approached to act as Patient Identification Centres (PICs). Patients attending TSA supported Tuberous Sclerosis Clinics were also notified of the trial.

2.7 Study design

TRON was a two-arm, individually randomised, Phase II, double-blind, placebo-controlled trial of everolimus versus placebo (allocation ratio of 2:1) in the treatment of neurocognitive problems in adult patients with tuberous sclerosis (TSC). The TRON study was a proof of principle study for memory and executive function outcomes, designed to provide effect size estimates that may inform the design of subsequent trials.

Patients were treated in an outpatient setting for 24 weeks with everolimus or placebo. The primary endpoint for assessment being 24 weeks (6months) from starting treatment. Another evaluation was scheduled at 36 weeks to assess if any gains were sustained after stopping everolimus.

DTI scans for TRON participants eligible for the imaging study were done at CUBRIC at baseline (visit 2) before the start of treatment and at the end of treatment (visit 7). A chronology of TRON study visits and a brief assessment schedule at each visit is depicted in Figure 5.

The treatment duration was selected based upon the expectation that any neurocognitive effects will result from changes either at a molecular level in signalling pathways or at a microstructural level, such as the myelination changes noted in mouse models treated with mTOR inhibitors (Meikle et al. 2008). The observation from clinical trials in patients with TSC was that tumour responses to mTOR inhibitors are most marked in the early months of treatment (Davies et al. 2008; Krueger et al. 2010a).

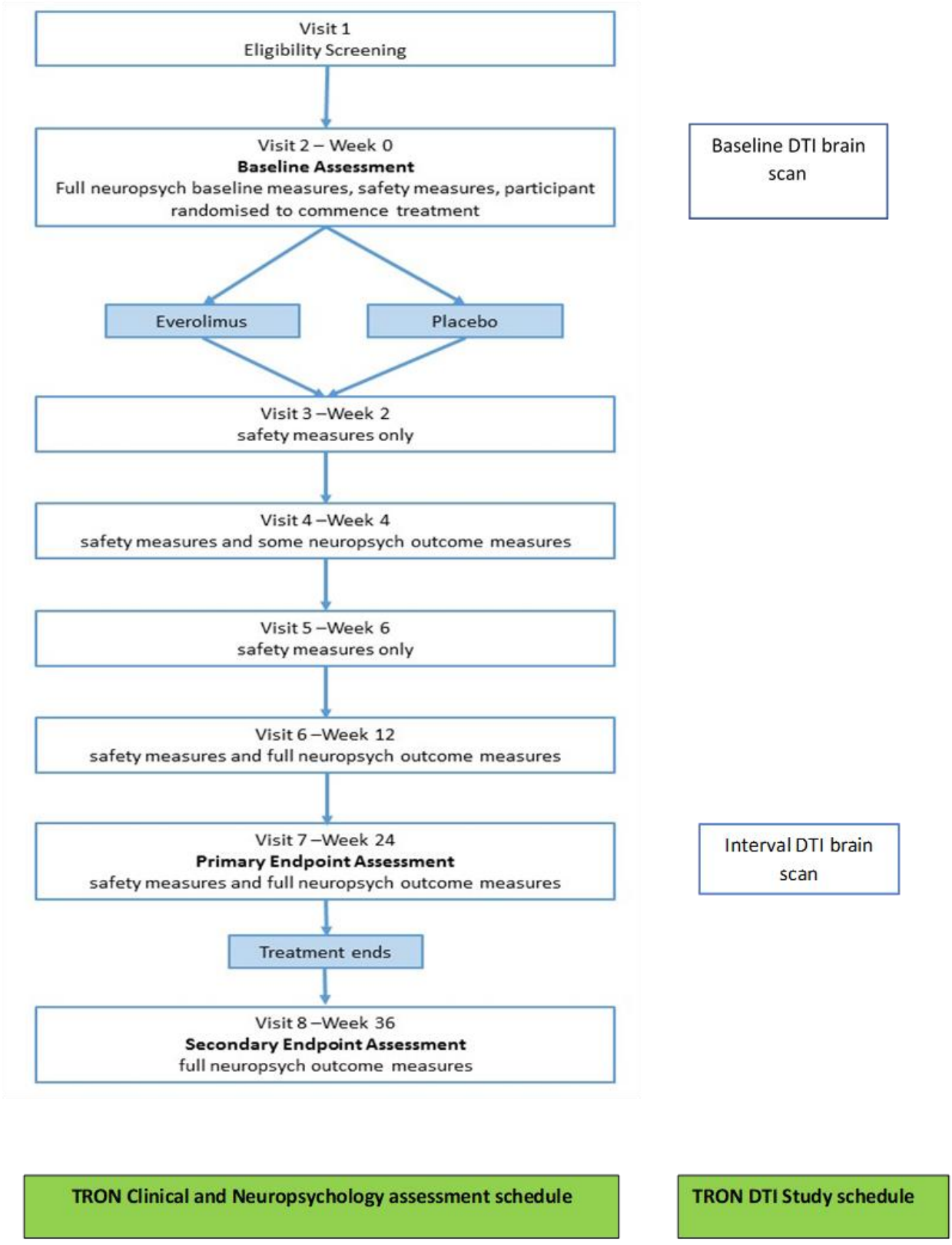


Figure 5 Schematic of study assessments in TRON study with DTI Imaging study visits

2.8 Neuropsychological function testing

Neuropsychological measures listed in Table 6 were used at various study visits. Although the test battery at some visits was extensive, patients with neurocognitive problems completed only the earliest stages of many tests. The difficulty increases until the participant reaches their ability limit. Therefore, participants completed their assessments much more quickly than the "maximum times" for assessments shown in Table 6. A similar test battery was used in TSC patients in a previous non-randomised, non-controlled open-label study at the Institute of Medical Genetics, Cardiff University (Davies et al. 2011b).

A neuropsychology assistant administered the neuropsychological assessments with the supervision of the trial psychologist. Participants were encouraged to travel to the trial centre the day before if travelling over a long distance to ensure they were well-rested for the assessment visits. Environmental changes were kept to a minimum (as much as possible) with the same room used for each visit. The assessment time (morning/ afternoon) was kept the same on every visit, and the same assessor undertook the assessments (leave permitting).

TABLE 6 Neuropsychological assessment measures

Primary outcome measures	Secondary Outcome Measures
1) BIRT Memory and Information Processing Battery a) Complex Figure test i) Immediate recall ii) Delayed recall b) List Learning test i) Immediate recall ii) Delayed recall 2) CANTAB a) Spatial Working Memory (SWM) i) Between Errors ii) Strategy fields b) Stockings of Cambridge (SOC) i) Mean initial thinking time ii) Mean Subsequent thinking time iii) Problem solved in least moves 3) Test of Everyday Attention i) Telephone search whilst counting	1) CANTAB a) Information Processing Battery (RVIP) b) Spatial Span (SSP) c) Attentional Set-shifting (IDED) 2) Controlled Oral Word Association Test (COWAT) a) Cancellation task b) Verbal Fluency 3) Symptom Checklist 90R (SCL-90R) 4) Quality of Life in Epilepsy (QOLIE) 5) Liverpool Seizure Severity Scale (LSS) 6) Vineland Adaptive Behaviour Scales-II (VABS-II) 7) Social Responsiveness Scale – Adult (SRS-A) 8) Social communication questionnaire (SCQ) 9) National Adult Reading Test (NART)

2.8.1 Screening visit Neuropsychological tests

The Wechsler Abbreviated Scale of Intelligence, National Adult reading Test and Edinburgh Handedness Test, described below, were performed only on the first (screening visit) assessment as part of eligibility screening.

2.8.1.1 Wechsler Abbreviated Scale of Intelligence (WASI)

The Wechsler Abbreviated Scale of Intelligence (WASI) (Psychological and PsychCorp 1999) is an individually administered, abbreviated measure of intelligence, designed for use from 6 through 90 years of age. The WASI-II is a parallel short form (with same name subtests) of both the Wechsler Intelligence Scale for Children (WISC) and the Wechsler Adult Intelligence Scale (WAIS) (Ryan et al. 2003; Sandoval 2014). Of the 1,145 adults included in the standardization sample, 248 were evaluated twice, once with the WASI and once with the WAIS-II, after approximately 1 month (Psychological and PsychCorp 1999). There was a correlation of 0.92 between WAIS-II FSIQ and WASI FSIQ-4 scores and of 0.87 with the WASI FSIQ-2 (Axelrod 2002; Wechsler 2011).

Several studies of neuropsychological assessment of individuals with tuberous sclerosis have used the WISC and WAIS (O'Callaghan et al. 2004; Raznahan et al. 2006; Jansen et al. 2008). Although these studies did not use the WASI, the high-reliability coefficients between the WASI and the Wechsler scales indicated that it was realistic to use the WASI in the TSC population.

The WASI yields a Verbal IQ (VIQ), a Performance IQ (PIQ), and a Full-Scale IQ (FSIQ) from the conversion of raw score to age-scaled scores (Ryan et al. 2003). The Verbal Scale (measuring expressive vocabulary, verbal knowledge and abstract verbal reasoning ability) consists of Vocabulary and Similarities subtests. The Performance Scale (measuring spatial visualization, visual-motor coordination and nonverbal fluid reasoning) comprises Matrix reasoning and Block Design subtests. FSIQ estimates are generated using the results from all four subtests (FSIQ-4) (Axelrod 2002). A four-subtest administration requires approximately 30 min.

2.8.1.2 Edinburgh Handedness Test

A measure of right or left-handedness, suggesting cerebral dominance, may be necessary for interpreting other tests (Oldfield 1971). A brief list of questions is presented about which hand would be used to perform everyday tasks, and the patient provides answers. A shorter version, namely, a

10-item inventory, was used in the TRON assessments. This brief quantitative measure has proved to be useful in neuropsychological, general, clinical, and research fields as a screening measure for establishing laterality (Fazio et al. 2012).

2.8.1.3 National Adult Reading Test (NART)

The National Adult Reading Test is an individually administered test of Reading (Nelson and O'Connell 1978). It is intended to assess premorbid levels of intelligence in adults from 20 to 70 years of age who are suspected of suffering from dementia. It consists of 50 words printed in the order of increasing difficulty. All words have irregular pronunciations so that they cannot be phonemically decoded. Thus, the subject can read them only if he or she recognizes them. Standardized intelligence tests can measure current cognitive capacities, while portions of these tests are considered to be less susceptible to dementing processes than others. NART's scores provide estimates of premorbid intellectual functioning and are comparable to scores obtained on the WAIS.

2.8.2 Primary outcome measures

TRON study's primary outcome measures were based on subsets of the BMIPB, CANTAB and TEA as specified in table 6. The details of the various subtests of the test batteries used in the TRON study are provided below.

2.8.2.1 Brain Injury Rehabilitation Trust Memory and Information Processing Battery (BMIPB)

The Brain Injury Rehabilitation Trust Memory and Information Processing Battery (BMIPB) is an individually administered test of memory and information processing skills (individuals aged 16– 89 years). The BMIPB was standardised and developed from the Adult Memory and Information Processing Battery (Coughlan and Hollows 1985).

The BMIPB was designed with serial assessments in mind, using four separate forms of the test to help cope with practice effects and facilitate retesting of recovery or deterioration of memory and information processing speed. A carefully constructed sample of 300 adults aged between 16 and 89 was recruited, matched closely to the UK general population for age, educational level and gender to validate the assessments (A.K Coughlan 2007). The four parallel forms of the BMIPB,

administered in sequential order, consist of seven subtests: Story Recall, List Learning, Word Recognition, Figure Recall, Design Learning, Design Recognition, and Speed of Information Processing.

Two Primary outcomes of the TRON study were based on the BMIPB subtest of List learning and figure recall (Randell et al. 2016). List learning assessments shown to be sensitive to cognitive impairments of memory & executive functioning due to a wide variety of causes such as hippocampal lesions (Henson et al. 2016) are described briefly below. The two subtests are assessed as immediate recall and delayed recall scores. BMIPB, therefore, provided four scores per assessment.

2.8.2.1.1 Complex Figure test (from the BMIPB)

A measure of visuospatial or non-verbal free recall memory (A.K Coughlan 2007).

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.

2.8.2.1.2 List Learning test (from the BMIPB)

A subset of the BMIPB, which measures verbal free recall memory. A list of words are read to the patient, List A and B, where List A is a 15-word list read to the participant and recalled from memory with a discontinuation rule of 5; it is followed by List B, an interference trial. List B is a 15-word List with only one administration followed by a final delayed recall attempt of list A.(A.K Coughlan 2007)

2.8.2.2 Cambridge Neuropsychological Test Automated Battery (CANTAB)

CANTAB is a system used widely in the disciplines of neuropsychology and psychopharmacology (Sahakian and Owen 1992). 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 or terminate test dependent on the serial responses (Fray and Robbins 1996). Table 7 details the CANTAB test battery (Luciana 2003), which also provides a list of subtests used in the TRON study.

TABLE 7 Tests included in the CANTAB battery

Title of subtest	Domain[a]	What it measures
Big Circle/Little Circle	Visual Attention	Basic Visual Discrimination
Delayed Matching to Sample	Visual Memory	Match to Sample Recognition Memory
ID/ED (Intra/extra-dimensional set) Shift[b]	Visual Attention	Discrimination Learning, Set-Shifting
Matching-to-Sample Visual Search	Visual Memory	Visual Recognition Memory
Motor Screening Test	Present in all modules	Reaction Time; Ease of touch-screen use
Paired Associates Learning	Visual Memory	Visual Paired Associates; Matching stimulus to location
Pattern Recognition Memory	Visual Memory	Forced-Choice Recognition of previously-seen patterns
Reaction Time	Visual Attention	Attention, Simple Reaction Time
Rapid Visual Information Processing[b]	Visual Attention	Attention; Continuous Performance Test
Spatial Recognition Memory	Visual Memory	Recognition Memory for Spatial Locations
Spatial Memory Span[b]	Working Memory/Planning	Memory Capacity; Analogue of Corsi Block Task
Spatial Working Memory[b]	Working Memory/Planning	Self-Guided Search; Working Memory; Strategic Search
Stockings of Cambridge[b] (Tower of London)	Working Memory/Planning	Planning; Behavioral Organization

[a] The battery is organized into three modules: Visual Memory, Visual Attention, and Working Memory/Planning

[b] Subtests used in TRON

CANTAB comprises a battery of 13 subtests, which is organized into three modules: Visual memory (Robbins et al. 1994), Visual attention (Downes et al. 1989), Working Memory and Planning(Owen et al. 1990). (table 7)

The visual memory and working memory batteries begin with simple tests, progressing to more complex tests that incorporate the earlier simpler tests' cognitive components. The working memory battery, which includes a test of planning, gives a sensitive measure of executive function (Owen et al. 1990). The attention battery includes tests of selective, divided, and sustained attention (Downes et al. 1989; Sahakian et al. 1993).

All task stimuli are nonverbal, consisting of geometric designs or simple shapes. Language proficiency is needed only to understand the instructions. CANTAB's psychometric properties have been established in diverse populations, in various paediatric studies (Luciana 2003) and studies from Taiwan (Gau and Shang 2010a,b), suggesting language-independent application. The

CANTAB's validity for assessing brain-behaviour relations in adults has been supported by numerous studies of patients with brain lesions and degenerative disorders (Owen et al. 1991; Owen et al. 1997; Rahman et al. 1999).

We used Stocking of Cambridge (SOC) and Spatial working memory (SWM) subtests of CANTAB battery as Primary outcome measures, while the Rapid Visual Information Processing (RVIP), Spatial Memory span (SSP), and Intra–extra-dimensional set (ID/ED) shift were also used as secondary outcome variables (Ni et al. 2013).

CANTAB provided 5 primary outcome scores per assessment (table7). The SWM subtest gives two scores (Between errors and Strategy fields) while the SOC gives three scores (Mean initial thinking time, Mean Subsequent thinking time, Problem solved in the least moves). A summary of the subtests used is provided below.

2.8.2.2.1 CANTAB - Spatial Working Memory (SWM)

The SWM assesses non-verbal working memory. It is based on a self-ordered search test (Petrides and Milner 1982) and it requires the retention and manipulation of spatial information in working memory. The subject must find a blue 'token' hidden in one of the on-screen white boxes by touching the boxes one at a time by trial and error. Once a token was found, there would never be another token inside the same square.

To avoid repeatedly searching in previous locations, the subject had to remember where he/she had searched and found a token. The order of searching was self-determined, and the number of boxes started at two. The subject ultimately completed four trials with two boxes, three boxes, four boxes, six boxes and eight boxes. 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).

Two major indices are included in the results : (1) strategy utilization: the number of search sequences starting with a novel box in the difficult problems (both six- and eight-box problems); (2) errors in the total and three different levels of difficulty (four-, six- and eight-box problems): the total errors for four-, six- and eight-box problems were calculated based on the between-errors, within-errors and double errors of particular box problems (Ni et al. 2013).

2.8.2.2.2 CANTAB - Stockings of Cambridge (SOC)

The SOC is an executive function test reliant on spatial planning based on the Tower of London test (Shallice 1982). Three discs were distributed in both the upper and lower stocking displays. The placement in the upper display was the template for the lower display. Thus, the subject was required to move the disc in the lower display until the three discs were located in the same place, respectively, indicated in the upper display. The starting configuration of the discs was varied with four problems which are 2 or 3 moves deep, four are 4 moves deep, and four are 5 moves deep. The strategy was to reach the required configuration after a minimum of two, three, four or five moves.

Four major indices were presented: (1) problems solved in minimum moves: the number of occasions that were completed in the minimum possible number of moves; (2) mean moves: the number of moves taken over the specified minimum number, but within the maximum allowed; (3) initial thinking time: reaction time taken to select the first disc for the same problem under the two conditions; (4) subsequent thinking time: the difference in time between selecting the first disc and completing the problem under the two conditions (Ni et al. 2013).

2.8.2.2.3 CANTAB -Rapid visual information processing (RVIP)

The RVIP is a 4-min visual continuous performance test modified and simplified from Wesnes and Warburton's task (Wesnes and Warburton 1984) was designed to assess sustained attention capacity (Sahakian et al. 1989).

Digits (ranging from 2 to 9) appear one at a time (100 digits/min) in random order. The subject had to detect three target sequences (3–5–7, 2–4–6, 4–6–8) and respond (within 1800 ms after the onset of the last number) when they saw the last number (7, 6 and 8, respectively). The subject was instructed to detect as many target sequences (27 in total) as possible. The total hit score represented the number of occasions upon which the target sequence was correctly responded to. The total misses score represented the number of occasions the participant failed to respond to a target sequence within the response window. The total false alarms score represented the number of times the participant responded outside the response window of a target sequence. The score of total correct rejections represented the number of stimuli that were correctly rejected.

Five indices were measured such as- probability of hits (h, the participant responding correctly), probability of false alarms (f, the participant responding inappropriately), mean latency along with an

index called – ‘A′, which is a signal detection measure of sensitivity to the target, regardless of response tendency.

2.8.2.2.4 CANTAB - Spatial Memory span (SSP)

The SSP measures spatial short-term memory. It is an analogue of the Corsi blocks task (Milner 1971), which required the ability to remember how visual stimuli were presented. In the beginning, individuals view a display of coloured squares on the computer monitor. One by one, the squares light up in a pre-determined sequence, after which an audible beep is presented. The beep serves as a signal for the examinee to respond by reproducing the sequence.

The primary measure of interest is in the longest sequence successfully recalled in the correct order (the spatial memory capacity). This task would be sensitive to cognitive impairments that impact continuously- developing skills, such as specific learning disabilities or injuries that are acquired during middle childhood (Luciana 2003). Another index reported is the total errors: the number of times an incorrect box was selected.

2.8.2.2.5 CANTAB- Intra–extra-dimensional set shifts (ID / ED)

The IED assessed a subject’s ability to selectively maintain his/her attention on the specific attribute of compound stimuli across different examples, or intra-dimensional shift, and then to shift their attention to a previously irrelevant attribute of stimuli or extra-dimensional shift (EDS) (Downes et al. 1989).

Throughout the ID /ED task, the subject was required to discover rules, initially through trial and error. Once the rule was achieved on six consecutive occasions, the computer established a new rule. Despite these changes, the subject had to try to make as many correct choices as possible.

Four target indices were included, namely: Pre-EDS errors: the number of errors made before the EDS stage; EDS errors: errors made in the EDS stage; Completed stages: the number of stages completed and, Adjusted total trials: the adjustment adds 50 for each stage not attempted due to failure at an earlier stage.

2.8.2.3 Test for Everyday Attention (TEA)

The Test for Everyday Attention battery of tests was published in 1994. It is reported to be the first norm-referenced test that assesses several independent attention systems. These are selective attention, sustained attention, attentional switching, and divided attention.

The authors of the assessment assume that using common, day-to-day activities makes the assessments seem relevant to examinees, ranging from young normal through early Alzheimer's populations. The assessments require about 45 minutes to complete and yield scores from eight subtests. There are three individually administered versions if needed for repeated testing; they are to be administered in a prescribed order because of practice effects.

The subtests include *Map Search*, which assesses selective attention; *Elevator Counting* is a measure of sustained attention, while *Elevator Counting with Distraction* is intended to measure auditory selective attention. The *Visual Elevator* subtest is intended to measure attentional switching, while *Elevator Counting with Reversal* is meant to measure auditory-verbal working memory. The *Lottery subtest* is envisioned as a measure of sustained attention. The *Telephone search* subtest is intended to assess selective attention, while the *Telephone Search While Counting* subtest is intended as a measure of the ability to perform two crossmodal tasks simultaneously, i.e., 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.

The TEA subtests are based on activities that were common for many adults (at the time of publication), such as reading telephone directories, scanning maps, and listening to lottery numbers as if in a broadcast. The tests have different versions that are used to reduce practice effects.

The *Telephone Search While Counting* (TSwC) subtest of the TEA battery of test was used as a primary outcome measure for the TRON study (Randell et al. 2016). We used version A & B forms following each other in the TRON study. Separate norms are used for each version. For test interpretation, the raw scores are converted to scaled scores. A summary of the administration of all the subtest is provided below.

2.8.2.3.1 Telephone search dual-task (from the TEA)

Telephone Search While Counting (TSwC) (Robertson et al. 1996) reflects sustained attention but also yields an estimate of dual-task decrement performance or divided attention. The subtest requires the examinee to again search for pairs of identical symbols on a simulated telephone page, but he or she must perform another simple task at the same time. The examinee simultaneously counts strings of tones presented from the audiotape. The original time per target TS score is subtracted from a TSwC weighted time per target score to estimate dual task decrement performance. This TSwC time per target score is weighted based on the accuracy of the examinee's tone counting. Scores for this subtest can be converted to scaled scores and approximate percentile ranks.

2.8.3 Secondary outcome measures

2.8.3.1 The Controlled Oral Word Association Test (COWAT)

The Controlled Oral Word Association Test (COWAT) (Benton et al. 1983), also known as the ‘‘FAS,’’ is a commonly used neuropsychological tool to assess participants performance on phonemic verbal fluency tasks, which are part of assessing Executive Function. The COWAT has been used to measure verbal fluency in various studies, including participants with frontal lobe lesions, which is a model for executive dysfunction (Janowsky et al. 1989).

The COWAT consists of three-word conditions. Participants were required to orally generate words following Benton’s administration criteria (Spreen and Strauss 1998), beginning with the letters *F, A,* and *S* in 60 secs. Subjects are also instructed to exclude proper nouns, numbers, and the same word with a different suffix (Kemenoff et al. 2002). Each letter (F, A, and S) is allowed one minute. If subjects discontinue before the end of the minute, they are encouraged to think more words. If there is a silence of 15 seconds, instructions and the letter are repeated. For scoring purposes, the actual words in the order in which they are produced are written down. The test administration takes about five minutes. The score is the sum of all admissible words for the three letters. Unacceptable responses occur when a subject repeats a previous response (i.e., a perseveration), or makes an error by including a word that starts with the wrong letter, or other rule violation as stated in the manual (Benton et al. 1983).

2.8.3.2 Cancellation Task

Cancellation tests are used in neuropsychological assessment to measure spatial exploration and awareness and as a simple diagnostic measure for unilateral spatial neglect. It is used as a test of lateralised attention bias. The test is administered by presenting a page of small silhouettes of everyday items. The participant must find the target items (e.g., bells) and cancel them as quickly as possible. Patients presenting with unilateral spatial neglect, most commonly after right-hemisphere lesions, typically perform poorly in these tasks, omitting to cancel targets on the contralesional (usually left) side of the page (Albert 1973)

2.8.3.3 Symptom Checklist 90R (SCL-90R)

The SCL-90R checklist was initially developed as The Hopkins Symptom Checklist (Parloff et al. 1954)to evaluate a broad range of psychological problems and psychopathology symptoms, such as anxiety and depression, and measure change over treatment duration. A set of 90 brief questions are given to the patient to rate. The checklist includes psychometrically nine factors: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, anger-hostility, phobic anxiety, paranoid ideation, and psychoticism (Carrozzino et al. 2019) (Derogatis and Unger 2010).

2.8.3.4 Quality of Life in Epilepsy (QOLIE)

This is an inventory proposed in 1995 to evaluate the overall quality of life, emotional well-being, social isolation, medication effects, perceived physical symptoms and cognitive functions, and health perceptions (Devinsky et al. 1995). A set of either 31 items or 89 are given for the patient to rate. This is an inventory used in individuals with epilepsy in a wide variety of population-based studies (Kovats et al. 2017) as well as clinical trials (Mukuria et al. 2017)

2.8.3.5 Liverpool Seizure Severity Scale (LSSS)

A set of 20 clinical features of seizure symptoms are rated by the patient (Baker et al. 1998). This scale was initially developed to incorporate the patients' perceptions of changes in seizure severity in addition to the alteration of frequency of seizure when evaluating the efficacy of an antiepileptic intervention. The scale has been revised twice since the original proposal in 1991 (Baker et al. 1991; Baker et al. 1998; Scott-Lennox et al. 2001).

The revised LSSS 2.0 avoids the problems of evaluating change over time associated with the major and minor seizure classifications. It detects differences between patients who experienced changes in their seizures associated with disease progression and pharmacotherapy. The revised scoring system is internally consistent, which indicates that its items constitute a homogeneous set that is likely to reduce random measurement error (Scott-Lennox et al. 2001). The Liverpool Seizure Severity Scale 2.0 questionnaire produces a single unit-weighted scale that measures the severity of the patient's most severe seizures during the recall period. As per the scoring procedures, the individuals are scored as per their answers on the questionnaire to produce an ICTAL scale (range of 0-100).

2.8.3.6 Vineland Adaptive Behavior Scales-II (VABS-II) (survey form)

VABS-II is an individually administered assessment of adaptive behaviour (Sparrow et al. 2005). It is considered a low-level measure of functional adaptive behaviour, i.e. ability to cope with personal and social skills in everyday life that may be deficient in autism or developmental delays.

The tests can be applied at any age with the parent/ carer form a questionnaire through an interview with a parent or caregiver. The Vineland scale is designed to assess three adaptive behaviour domains: Communication, Daily Living skills, and Socialization.

Each form uses a 3-point response scale. The paper-and-pencil forms can be scored by hand, and the test manual provides step-by-step directions for converting raw scores to norm-referenced scores such as standard scores, confidence intervals, and percentile ranks.

2.8.3.7 Social Responsiveness Scale – Adult version (SRS-A) (Constantino 2005)

To measure the severity and type of social impairments characteristic of autism spectrum disorders in age 19-89. SRS questionnaire is a useful screen for ASD in a TSC population(Granader et al. 2010). The SRS-A is used in screening and/or as an aid to a clinical diagnosis of ASD and comprises 65 questions, rated on a 4-point Likert scale. In addition to a total score, the SRS consists of five subscales: Social Awareness, Social Cognition, Social Communication, Social motivation and Autistic Mannerisms (South et al. 2017; Torske et al. 2017).

The SRS-2 manual [Constantino & Gruber, 2012] acknowledges concerns about discriminant validity, particularly regarding clients with anxiety, ADHD and other conditions that affect social communication and behavioural flexibility. This aspect of the assessment has been recognized in

clinical studies (South et al. 2017). Social impairments identified on the SRS have been shown to be linked to executive dysfunction in the background of autism (Torske et al. 2017) , which is under-recognized and frequent in individuals with TSC.

2.8.3.8 Social communication questionnaire (SCQ) (Rutter M 2003)

The SCQ is another commonly administered screening instrument for Autism. SCQ has been proven to be a reliable screening tool for autistic spectrum disorder (Chesnut et al. 2017). It has been used in children with TSC to measure social functioning (Granader et al. 2010)

The SCQ is a 40-item (yes/no response format) questionnaire that evaluates communication skills and social functioning, both historically and currently. The TRON study used the Current version (the SCQ AutoScore™ Form: Current). Total scores can range from 0 to 39 (the first question is a language screening item that is not included in the final score), and a total SCQ raw score of ≥15 is highly suggestive of ASD.

2.9 Evaluation of Response

This study's primary outcome was neurocognitive (memory and executive) functioning with improvement defined as at least a one SD response in one or more of the neurocognition assessment tests listed in Table 6

Based upon the information on learning effects in neurocognitive tests in the general population, the TRON investigators together estimated the learning effect (proportion of participants who might be classified as “responders” simply due to familiarity with the assessments) to be approximately 0.15 (15%). Therefore, a proportion of less than 15% improvement in the intervention group would indicate that the intervention (Everolimus in the treatment of neurocognitive problems in tuberous sclerosis) did not warrant further investigation. A proportion of at least 35% was considered to provide sufficient evidence for further investigation of this intervention. Values in between would be regarded as uncertain and discussed in depth by the trial team members for interpretation.

A one-sample chi-square test was to be used to determine whether the proportion of participants in the intervention group who improved their recall memory at 6 months by one SD was at least 20% greater than the proportion that improved in the control group. The effect size was to be presented alongside a 95% confidence interval (CI) and p-value. The proportion of participants in the control group displaying improved functioning by one standard deviation (SD) would highlight the learning effect.

2.10 Statistical Analysis Plan

All randomised participants who received at least one dose of study drug were included in the data analysis. Participants’ assessment scores were analysed according to the treatment they received. Data were reported descriptively and used to determine effect sizes. Results were presented split by trial arm.

2.10.1Descriptive analysis

Summary statistics on eligibility, recruitment, withdrawal and dropout were collated for both trial arms and presented in a CONSORT flow diagram for clinical trial reporting. Specifically, for each arm, numbers of participants randomly assigned, receiving intended everolimus, completing the study protocol, and which were analysed for the primary outcome were documented.

Baseline data were used to check comparability between study arms and generalisability of the study population. There was no formal testing of between-arm differences for any variables at baseline.

2.10.2 Analysis of primary outcome

Data were to be presented descriptively by the trial arm at baseline and 6 months, and effect sizes determined. This analysis was a 2 stage process.

1. The proportion of participants (alongside 95% CI) in the placebo group displaying improved functioning was reported to highlight the learning effect. A responder was defined as showing at least 1SD (using population norms) of improvement in the scores of ANY of the tests used as a primary outcome variable compared to baseline assessment (table 7)

   If the learning effect observed was different from the pre-hypothesised 0.15, it was agreed to discuss within the trial team whether an improvement of 0.35 in the Everolimus group would represent an appropriate threshold to support or not support the case for future larger studies.

2. A one-sample chi-squared (or goodness-of-fit) test was to be used to determine whether the proportion of participants (alongside 95% CI) in the Everolimus group who improved their recall memory and executive functioning at six months by at least 1 standard deviation (SD) was not statistically significantly different from the pre-determined (or revised) threshold.

2.10.2.1 Concept of Effect size

Effect size quantifies the difference between two groups emphasising the size of the difference rather than confounding this with sample size (Coe 2002). Effect size is defined as “a quantitative reflection of the magnitude of some phenomenon that is used for the purpose of addressing a question of interest”. In a research context, “effect” is a quantitative reflection of a phenomenon and “size” as the magnitude of something (Kelley and Preacher 2012). The effect sizes are especially useful in looking at the research in neuropsychological tests. The magnitude of change (how well does the intervention work?) would be of clinical value rather than just the binary outcome (Does the intervention work or not?) if an intervention worked. The American Psychological Association has been encouraging authors to report effect sizes since 1994 (Wilkinson 1999).

2.10.3 Analysis of secondary outcomes

The secondary outcomes were selected to reflect wider aspects of neurocognitive function, seizures, quality of life, and daily life functioning in participants with tuberous sclerosis. The planned primary analysis of the secondary outcomes was similar to that undertaken for the primary outcome. The secondary outcomes were to be presented descriptively at baseline and six months, and effect sizes determined. The proportion of participants that improved over time (defined as at least one SD change in score) for each of the secondary outcomes were examined. A one-sample chi-square test was used to determine whether the proportion of participants in the everolimus group who improved at six months by at least one SD was significantly greater than expected (the null hypothesis of equal proportions is rejected).

The treatment effects on the proportion of participants improving in their wider aspects of neurocognitive function, seizures, quality of life, and daily life functioning at six months were examined using logistic regression models adjusted for the balancing factors, and results presented as adjusted odds ratios (ORs) alongside 95% CIs comparing the odds of improvement in the everolimus compared with the placebo arm.

2.11 Sample Size calculations for the TRON clinical trial

The sample size was calculated based on Fleming's single-stage procedure (Fleming 1982) and considered the sample size required to assess the experimental group's change. This study was not powered for formal statistical comparison of the placebo and study drug groups.

To test the null hypothesis that the proportion of participants 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 participants 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 required a total sample size of 38 (i.e. 25 interventions and 13 controls).

The target sample size was initially increased to 48 (i.e. 32 interventions and 16 controls) to allow for 20% loss to follow-up. Subsequently, since the loss to follow up was minimal and recruitment problematic, the target sample size was revised back to 38.

2.12 Randomisation and blinding

Randomisation was carried out by the South East Wales Trials Unit (SEWTU) using a computer-generated allocation sequence according to a ratio of 2:1, intervention to control. The participant's unique identification number and allocation were double-blinded, so neither the participant, clinician-researcher, research psychologist or the trial statistician knew the treatment allocation group of the participant.

The trial pharmacy held the details of participants' treatment allocation and maintained 24-hour cover to unblind in an overdose case. For other emergencies, at the outset, withholding study drug was planned, as unblinding would not affect symptom management, while a member of the clinical team was contacted as soon as possible afterwards. For non-emergency clinical issues, the local physician contacted a trial clinician to discuss the need to unblind or stop treatment.

2.13 Treatment: Dose rationale

The daily investigational drug dose of everolimus was 5mg, administered for 6 months as two oral 2.5 mg tablets once daily, but with adjustment to achieve trough blood levels of 3-10ng/ml.

This dose and target blood level is lower than in adults being treated for renal cancer. Previous research in patients with TSC had suggested that lower doses and blood levels may be as effective as traditional doses and blood levels of mTOR inhibitors in the treatment of renal angiomyolipoma (Davies et al. 2008). However, there was no human data to suggest optimal dosage for treating neurocognitive problems in TSC. Although a lower dose (compared to the recommended dose to treat renal AMLs) was chosen to mitigate side effects, the target trough levels are comparable to those for treatment of AML (3-10 in TRON study vs 3-8 ng/ml recommended for renal AMLs).

The appearance of placebo medication was identical to that of active drug to maintain the blinding. Pharmacy staff were unblinded and were instructed not to reveal treatment allocation to anyone without permission from a clinical team or an investigational team member.

2.14 Safety assessments

Safety assessments included clinical examinations, vital signs and standard clinical laboratory evaluations of haematology, blood chemistry, spirometry, and urinalysis. Adverse event and serious adverse event monitoring were performed at each site visit as per the TRON protocol (Randell et al. 2016).

Samples for drug levels (PK assessments) were taken at visits 3-7. Treatment was dispensed at visits 2, 4 and 6. Participants took the medication at the site at visits 3 to 7 and at home on all other treatment days. Pill counts monitored compliance at each visit. Trough blood levels were measured at each study visit in the treatment phase and reported to the study clinician through the trials unit. Mock levels for patients on placebo were provided to the trial clinician by the trials unit in proportion to those for active drug patients. The study clinician made decisions regarding dosage changes, as per the study guidelines.

2.15 Safety reports and Clinical management

Adverse events (AE) were defined as the appearance or worsening of any undesirable sign, symptom, or medical condition occurring after starting the study drug or placebo, even if the event was not considered to be related to the study drug. Medical conditions/diseases present before starting the study drug were considered Adverse Events if they worsened after starting the study drug. Abnormal laboratory values or test results constituted Adverse Events only if they resulted in clinical signs or symptoms, were considered clinically significant, or required intervention.

AEs were recorded on the Adverse Events form with the signs, symptoms, or diagnosis associated with them. All AEs were reported in accordance with the principles of Good Clinical Practice and the requirements of the Medicines for Human Use (Clinical Trials) Regulations 2004. The severity was graded according to the National cancer institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03 (NCI 2010).

Each adverse event was evaluated to determine:

1. the severity grade (1-4)
2. its relationship to the study drug(s) (suspected/not suspected)
3. its duration (start and end dates and times or if continuing at final examination)

4. action taken (no action is taken; study drug dosage adjusted/temporarily interrupted; study drug permanently discontinued due to this adverse event; concomitant medication taken; non-drug therapy given; hospitalisation / prolonged hospitalisation)
5. whether it constituted a *Serious Adverse Event* (SAE)

An SAE was defined as an event which:

1. was fatal or life-threatening
2. resulted in persistent or significant disability/incapacity
3. constituted a congenital anomaly/birth defect
4. required inpatient hospitalisation or prolongation of existing hospitalisation, unless hospitalisation was for:
    4.1. routine treatment or monitoring of the studied indication, not associated with any deterioration in condition (specify what this includes)
    4.2. elective or pre-planned treatment for a pre-existing condition that is unrelated to the indication under study and has not worsened since the start of study drug
    4.3. treatment on an emergency outpatient basis for an event not fulfilling any of the definitions of an SAE given above and not resulting in hospital admission
    4.4. social reasons and respite care in the absence of any deterioration in the patient's general condition
5. was medically significant, i.e., defined as an event that jeopardies the patient or may require medical or surgical intervention to prevent one of the outcomes listed above.

Unlike routine safety assessments, SAEs were monitored continuously and had special reporting requirements.

An *Adverse Reaction* (AR) was defined as any noxious and unintended response in a clinical trial participant to whom everolimus had been administered, which was related to any dose administered. A "response" to everolimus means that a causal relationship between everolimus and an adverse event is at least a reasonable possibility, i.e. the relationship cannot be ruled out. A *Serious Adverse Reaction* (SAR) was any SAE occurring in a clinical trial participant for which there is a reasonable possibility that it was related to everolimus at any dose administered. A *Suspected Unexpected Serious Adverse Reactions* (SUSAR) was a SAR classified as 'unexpected,' i.e. an adverse reaction, the nature and severity of which is not consistent with the information outlined in the SmPC for everolimus or an expected manifestation of tuberous sclerosis.

2.15.1 Adverse events (AE) management

All Adverse Events and reactions were treated appropriately. Such treatment included changes in study drug treatment, including interruption or discontinuation, starting or stopping concomitant treatments, changes in the frequency or nature of assessments, hospitalization, or any other medically required intervention. Once an adverse event was detected, it was followed until its resolution, unless deemed unnecessary at the end of the study or lost to follow-up and documented in the study file and the AE form. AE assessments were made at each visit (or more frequently, if necessary) for any changes in severity, the suspected relationship to the study drug, the interventions required to treat it, and the outcome.

2.15.2 Dose interruption, modification or discontinuation due to intercurrent events

Information about common side effects that were already known for everolimus was based on the information provided in the Summary of Product Characteristics (SmPC) (Novartis 2019).

The study drug was withheld for at least two weeks before and two weeks following elective surgery, requiring more than three stitches, entry into a body cavity or for optimal healing (e.g. laser dermatologic surgery on the face). The study drug was withheld for at least two weeks following accidents or emergency surgeries that met the criteria above. Table 9 provides the procedure followed for dose modification and re-initiation of treatment in the event of AE, suspected to be related to the study participant treatment with everolimus.

TABLE 8 Study drug dose modification for Adverse Events

Toxicity	Action
Infections	Grade1: No specific dose adjustments recommended. Grade 2 and 3: Interrupt study drug until recovery to ≤ 1 (antibiotics stopped). Restart at the same dose. Grade 4: Interrupt study drug until recovery to ≤ 1. Reintroduce study drug at the next lower dose level.
Stomatitis	Grade 1: No specific dose adjustments recommended. Manage patients based on clinical judgment. Grade 2: Interrupt study drug until recovery to ≤ 1. Restart at the same dose. Grade 3: Interrupt study drug until recovery to ≤ 1. Reintroduce study drug at the next lower dose level. Discontinue study drug if stomatitis doesn't recover to ≤ 1 within 4 weeks. Grade 4: discontinue study drug
Platelet count	≥ 75x 10^9/L No change 50 x 10^9/L to 75 x 10^9/L . Hold study drug until recovery to ≥ 75000/mm3. Reintroduce study drug at the same dose level. < 50 x 10^9/L Hold study drug until recovery to ≥ 75 x 10^9/L reintroduce at the next lower dose level.
Neutrophil count	≥ 1 x 10^9/L No change 0.5 to 1 x 10^9/L. Hold study drug until recovery to ≥ 1 x 10^9/L. Reintroduce study drug at the same dose level. < 0.5 x 10^9/L . Hold until recovery to ≥ 1 x 10^9/ L. Reintroduce study drug at the next lowest dose level.
Febrile neutropenia	Hold study drug until neutrophil count ≥ 1.25 0.5 x 10^9/L and no fever. Then resume at the next lower dose level.
Hyperlipidaemia and/or Hypertriglyceridaemia	Grade1 & Grade 2: Monitor and treat according to local best clinical practice. Consider baseline measures for interpretation. Grade 3: hypercholesterolemia (> 400 mg/dL or 10.34 mmol/L) hypertriglyceridemia (>5 × ULN) - Treat with (HMG)-CoA reductase inhibitor (e.g. atorvastatin, pravastatin) or appropriate lipid-lowering medication, in addition to a diet.
Other toxicities	Grade1: No specific dose adjustments recommended. Manage patients based on clinical judgment. Grade 2 and 3: Interrupt study drug until recovery to ≤ 1. Restart at the same dose. Grade 4: Interrupt study drug until recovery to ≤ 1. Reintroduce study drug at the next lower dose level.
Toxicity requiring interruption for ≥ 6 weeks	Permanently discontinue treatment.

2.15.3 Dose modification and discontinuation of study medication

The dose of study medication was changed by study clinicians if required due to an AE or in response to the blood levels. The dose was adjusted to achieve trough blood levels of 3-10ng/ml. The starting dose was 5 mg once a day; the next dose level change was by 2.5 mg.
Participants with high trough levels or experiencing toxicity could reduce to 2.5mg once a day. Participants on 2.5 mg once daily could be further reduced to 2.5mg every other day. However, if a further dose reduction from 2.5 mg every other day were required, participants were to discontinue study treatment.

2.15.4 Pregnancy

Eligible participants were advised to use adequate contraception during TRON study participation and until 30 days after stopping the study drug. A pregnancy test was performed at every clinical visit, where appropriate.

We planned to record any pregnancy on a Clinical Trial Pregnancy Form and report it to the Sponsor as an adverse event. Pregnancy was a stated reason for the discontinuation of the patient in the clinical trial. Any pregnancy was to be followed to determine the outcome, including spontaneous or voluntary termination, details of the birth, and the presence or absence of any birth defects, congenital abnormalities, or maternal and/or newborn complications. An assessment for a relationship between the study drug to any pregnancy outcome would be performed. Any pregnancy occurring after signing the informed consent and before the patient was enrolled would have been an exclusion criterion. If the father was taking the study drug, we planned to report a pregnancy in the partner occurring up to 3 months after stopping the study drug.

2.16 Results

2.16.1 Study Recruitment

The trial's initial recruitment target was 48 and had been inflated by 20% to allow for withdrawals and loss to follow-up. Owing to recruitment difficulties but better than expected retention, midway through the trial, a decision was made by the trial management group, ratified by the trial steering committee, to drop the 20% inflation to the recruitment target. The revised recruitment target was, therefore, 38. A total of 383 potential participants were approached by an invitation to take part in the study. Sixty-seven attended the screening visit for assessment of eligibility, of whom 38 were considered eligible after medical and neuropsychological assessments at the screening visit.

2.16.2 Addressing Recruitment Issues

The trial was initially planned for completion in three years; however, it had to be extended (Figure. 6). To address slow recruitment, the TRON study team visited TSC clinics across UK, including those in London, Birmingham, Nottingham, Leicester and Leeds, to present the trial rationale and details to clinic staff. The trial team prepared approach packages for participants to be sent by their respective clinicians with patient information leaflets, consent forms for agreement for the study team's approach for more information, and pre-addressed an stamped return envelopes to simplify responses.

The clinical trial team also undertook visits to screen clinical records in genetic centres to identify potential participants who were then approached by their local clinical team for participation in the TRON study. This intervention was in recognition that the regional genetic centre clinical teams, had no spare time for this exercise (it predated the establishment of genetics research nurses employed through the NIHR) and the possibility of pre-screening in some centres due to lack of familiarity with the study protocol. In Feb 2016, a financial incentive of £300, was agreed per patient identification centre to refer potential participants for a screening visit. The study team also presented the trial protocol in national clinical symposia and patient group meetings (with the help of the Tuberous Sclerosis Association) to aid recruitment.

Asides from the efforts above, a dedicated TSC clinic service was launched at Cardiff in Aug 2015. Although this was a clinical service, it provided an opportunity to discuss the TRON study opportunistically, in a face to face setting, with the TSC cohort in south Wales. Six participants were recruited to the TRON study after discussions in this clinic.

Informal feedback received from clinicians in centres based far from Cardiff suggested that some families were potentially open to participating in the study but put off by the challenges of the travel involved. Therefore, in Dec 2015, the trial management group decided to open two more centres in the UK - Belfast and Glasgow to facilitate recruitment in Scotland and Northern Ireland. The clinical trial team travelled to the additional centres for all patient visits, ensuring consistency of clinical and research aspects of the trial, standardised evaluation of all participants and avoidance of inter-observer bias.

2.16.3Timeline of TRON study

Figure 6 depicts the TRON trial recruitment timeline with milestones and changes made to promote recruitment to study. The 1st participant visit to Cardiff was in September 2012, while 1st participant visit in Glasgow was in June 2016, and Belfast opened in October 2016. The final participant visit was in August 2018. The imaging sub-study was open between May 2013 and June 2018. The geographic locations of residence of the TRON study participants and the study centres are depicted in figure 7.

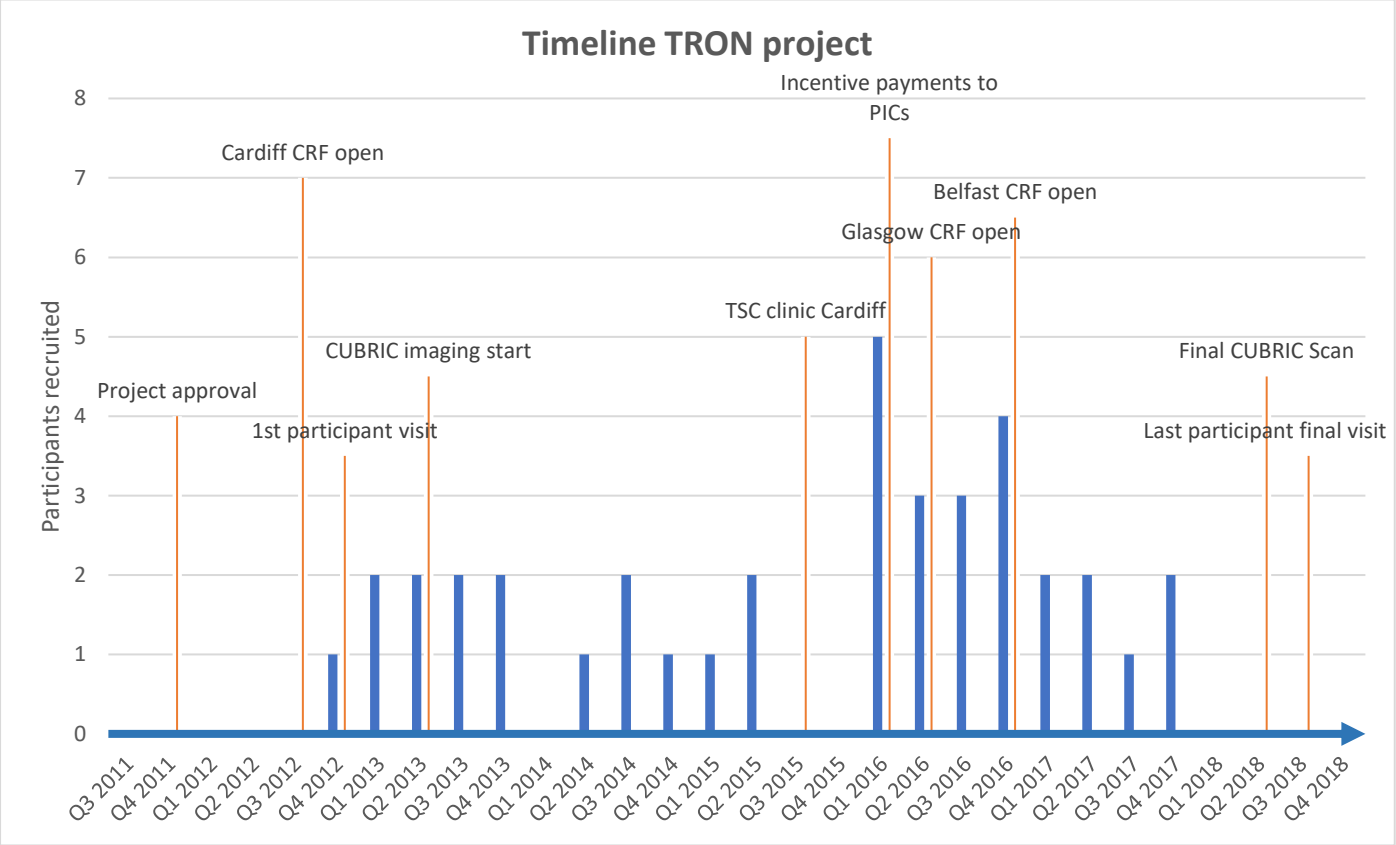


FIGURE 6 Timeline of TRON trial recruitment with milestones of the study

Blue bars represent quarterly recruit numbers

FIGURE 7 MAP OF study centres and study participant residence locations*

*Orange Stars- participant location

*Purple Pins- Study centres

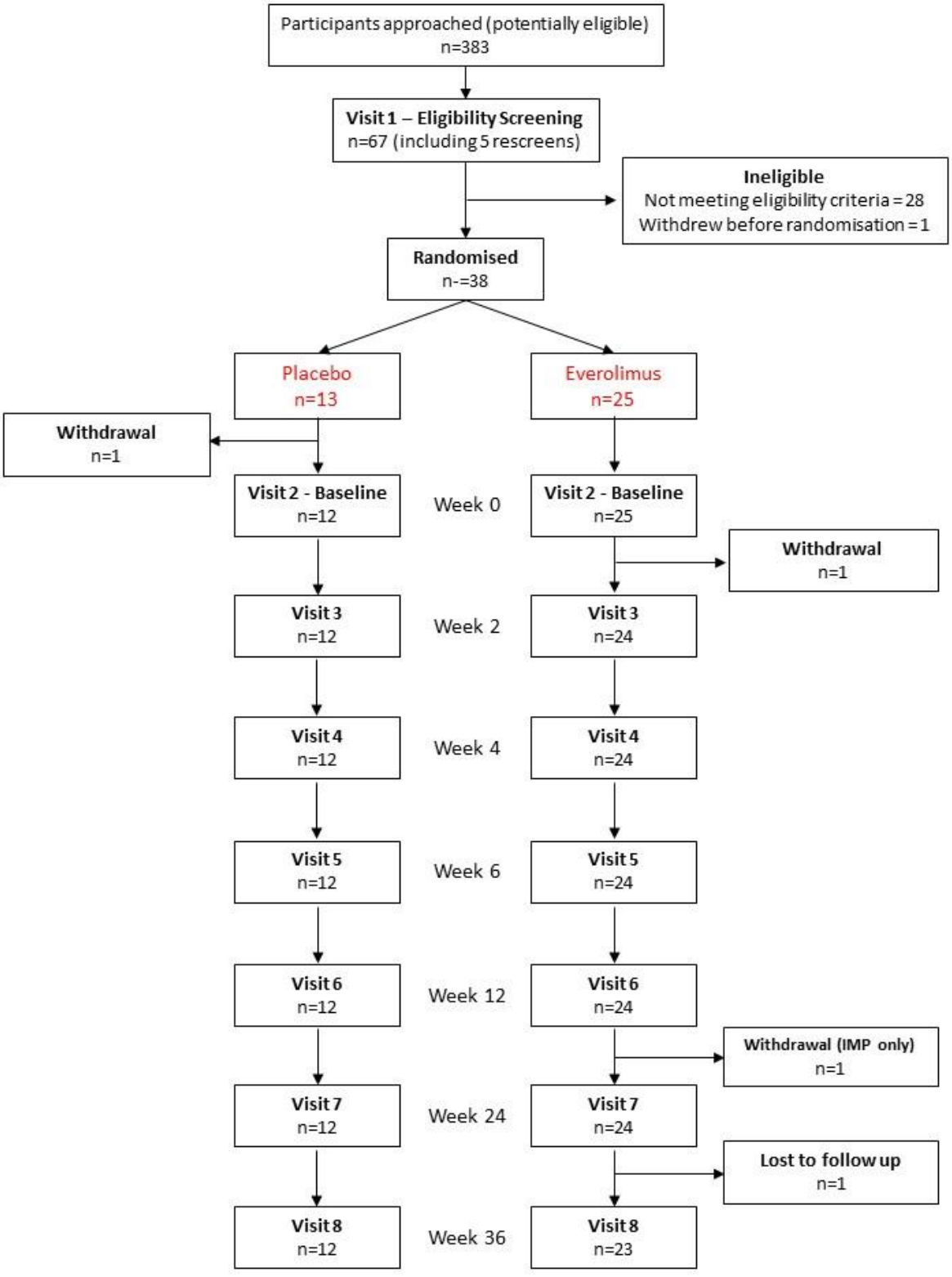


FIGURE 8 CONSORT diagram showing approach, screening and recruitment and retention/loss of participants in the TRON trial

2.16.4 Recruitment and retention over the study period

The recruitment and retention of trial participants are shown in the CONSORT diagram (Figure 8). The specified target of 38 participants were recruited, with 13 participants randomised to the placebo group and 25 participants randomised to the everolimus group. There were 3 withdrawals; 2 in the everolimus arm (1 complete withdrawal and 1 partial withdrawal where the participant stopped taking the study medication, but continued with assessments) and 1 in the placebo arm before baseline assessment. There was one further loss to follow up after the primary end-point visit at 6 months, in the everolimus arm. This resulted in a total of 36 participants for analysis at the primary endpoint.

2.16.5 Baseline characteristics of the study population

Table 9 details the employment categories for TRON participants as per the National Statistics socio-economic classification (NS-SEC)(Rose 2003), and details of participants in education or who had retired. The majority of the participants (19 of 38; 50%) were long-term unemployed or never employed and (11 of 38, 29%) in the lower service classes of L4,L6 and L7. Eight TRON participants were in education or had retired (21%)

TABLE 9 Employment categories as per the (NS-SEC) criteria, in education or post-retirement

	Job categories	Participant number
1	Higher managerial, administrative and professional occupations	0
2	Lower managerial, administrative and professional occupations	0
3	Intermediate occupations	0
4	Small employers and own account workers	3
5	Lower supervisory and technical occupations	0
6	Semi-routine occupations	2
7	Routine occupations	6
8	Never worked and long-term unemployed	19
9	Students	5
10	Retired	3

Table 10 shows the characteristics of the study population in terms of the minimisation variables (age, gender, IQ and antiepileptic drug (AED) use) used in relation to the trial arms (Everolimus, Placebo). The arms were well balanced across all minimisation variables and neurocognitive measures.

The majority of the study participants were <50yr of age, with slightly more females in both arms. The IQ distribution revealed that the many of the participants, 24 / 38 (63% overall; 60% Everolimus vs 69% Placebo group), had an IQ of >80, with 17 / 38 (44.7%) having an IQ>90. 14 / 38 (37%) had an IQ<79 (Borderline or low), while only 4 / 38 (~10%) had IQ below 70. The mean NART scores, and EHI handedness test scores were comparable across the two groups.

TABLE 10 Baseline characteristics of the TRON study population

	Placebo n=13			Everolimus n=25		
Age group	n	%		n	%	
<50 years	10	76.9		21	84.0	
≥50 years	3	23.1		4	16.0	
Gender						
Female	7	53.8		14	56.0	
Male	6	46.2		11	44.0	
IQ Level Group						
<= 69 (Extremely low)	2	15.4		2	8.0	
70 – 79 (Borderline)	2	15.4		8	32.0	
80-89 (Low average)	0	0		7	28.0	
90-109 (Average)	8	61.5		7	28.0	
110-119 (High average)	1	7.7		1	4.0	
120-129 (Superior)	0	0		0	0.0	
130+ (Very superior)	0	0		0	0	
Currently on an anti-epilepsy drug (AED)						
No	4	30.8		9	36.0	
Yes	9	69.2		16	64.0	
Neurocognitive measures at screening						
NART Error Scale	n	Mean	SD	n	Mean	SD
Full scale	13	107.7	10.64	25	106.5	10.79
Verbal	13	106.8	11.39	25	105.6	11.73
Performance	13	107.9	8.72	25	106.8	9.04
Error score	13	23.0	11.74	25	24.1	11.21
EHI Handedness Scale - LQ	13	68.6	24.30	25	84.1	20.17

2.16.6 Neurocognitive deficits and eligibility for the TRON study at screening

Participants with a deficit of ≥ 1.5SD in one or more of the 10 primary outcome neurocognitive measures (table 6) and fulfilling the other eligibility criteria as in tables 4 and 5 were eligible to take part in the study. Nineteen participants were eligible based on a deficit in only one of the measures (table 11)

Table 11 Number of neurocognitive measures for eligibility at screening visit

	Number of Neuropsychological Tests (1.5 SD deficit) at screening visit						
	1	**2**	**3**	**4**	**5**	**6**	**Total**
	n	n	n	n	n	n	n
BMIPB List Learning *(Immediate Recall)*	2	2	2	2	4	1	13
BMIPB List Learning *(Delayed Recall)*	2	4	2	0	3	0	11
BMIPB Complex Figure Test *(Immediate Recall)*	3	4	1	1	4	1	14
BMIPB Complex Figure Test *(Delayed Recall)*	0	1	0	1	0	1	3
SOC (CANTAB) *(Mean initial thinking time)*	0	0	0	0	1	1	2
SOC (CANTAB) *(Mean subsequent thinking time)*	3	1	0	1	1	1	7
SOC (CANTAB) *(Problems solved in least moves)*	1	2	1	3	4	1	12
SWM (CANTAB) *(Between-Errors)*	1	0	0	1	1	0	3
SWM (CANTAB) *(Strategy Fields)*	0	1	1	1	1	0	4
TEA *(Telephone search whilst counting)*	7	1	2	2	1	0	13
Participants (n) eligible with (number of) individual eligibility tests	19	16	9	12	20	6	

2.17 Primary outcome analyses

Initial assessment of placebo arm participants at baseline and 6 months revealed that 9 of 12 (75%) of participants in the placebo arm showed an improvement of at least 1SD (table 12) in at least one of the 10 primary outcome measures (table 7). This was discussed with the trial statisticians. It was agreed we should nevertheless look for a difference between arms of 20% as it would still represent a clinically useful proportion of cases benefitting from everolimus treatment. The observed improvement in the everolimus arm was 87% (Table 12). The improvements observed in both arms were larger than hypothesised, and the 12-percentage point difference between arms did not meet the pre-specified threshold of a 20-percentage point difference.

Table 12 Primary Outcome: Improvement from Baseline to 6 months, whole study population

	Placebo n=12			Everolimus n=24			Total n=36		
Main analysis: Whole TRON population									
Improved in any measure by ≥ 1SD	n	%	95% CI	n	%	95% CI	n	%	95% CI
Yes	9	75.0	46.8 - 91.1	20	87.0	67.9 - 95.5	29	82.9	67.3 - 91.9
No	3	25.0	8.9 - 52.2	3	13.0	4.5 - 32.1	6	17.1	8.1 - 32.7
Missing	0			1*			1		

**1 participant missing Stockings of Cambridge (SOC) response at 6 months so excluded from analysis – this participant also only screened eligible based on TEA task.*

As per the statistical plan, a one-sample χ^2 test for improvement to 95% (revised threshold derived from adding the hypothesised 20-percentage point increase to the observed 75% improvement from baseline to 6 months in the placebo arm) in the everolimus arm provides a p-value of 0.077, implying no significant evidence to suggest that the observed proportion of 87.0% in the Everolimus arm is different to the hypothesised 95.0%. This indicates that there is no statistically significant difference between observing 95.0% or 87.0% improvement in the everolimus arm.

Further inspection of data for the placebo group revealed that the TEA telephone search whilst counting subtest (Robertson et al. 1996) displayed marked intra-participant instability in the scores (Figure 9), which was unexpected. It was noted that the variation might be explained by differences between the two versions of the TEA subtest being used in TRON; version A was used at screening and the 24 weeks (primary endpoint assessment) and version B at baseline. Whilst it is standard practice in neuropsychological testing to vary the available versions between repeated assessments

to minimise learning effects, a raw score of 1.5 on version A produces a scaled score of 8. In contrast, a raw score of 1.5 in version B produces a scaled score of 5. The different output score makes simple comparisons of scores between versions invalid. This is particularly problematic for our analysis as a difference in a scaled score of 3 equates to a 1 standard deviation shift.

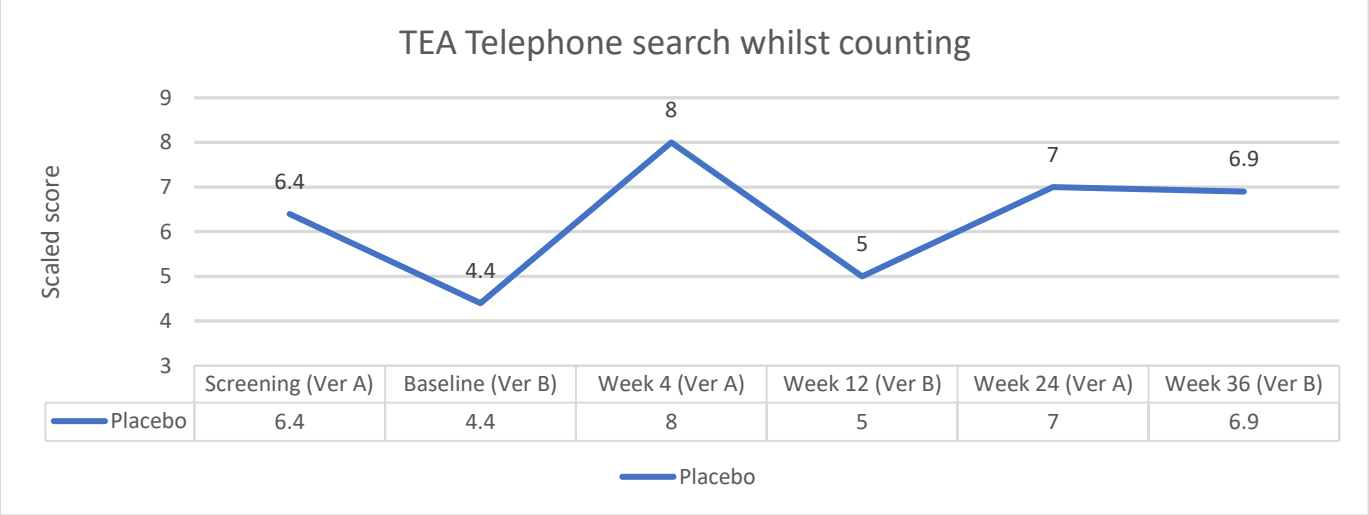


Figure 9 Mean Value (scaled scores) with the Version of TSwC used across the Placebo group visits

Following the trial statisticians' advice, the clinical trial team decided to exclude the TEA from both the screening assessment and primary outcomes analysis. Seven participants had been included in the trial based on eligibility on the TEA alone. Therefore, the numbers used to analyse the primary outcome were reduced to 29 at baseline (n=9 in the placebo group and n=20 in the Everolimus group). An analysis conducted after adjustment (excluding the 7 participants eligible only on TEA) is presented in table 13 and figure 10.

Table 13 Primary Outcome: Improvement from Baseline to 6 months excluding participants eligible on TEA only

	Placebo n=9			Everolimus n=20			Total n=29		
Sensitivity analysis: Population excluding those eligible based on TEA only at screening									
Improved in any measure by ≥ 1SD	n	%	95% CI	n	%	95% CI	n	%	95% CI
Yes	4	44.4	18.9 - 73.3	14	70.0	48.1 - 85.5	18	62.1	44.0 - 77.3
No	5	55.6	26.7 - 81.1	6	30.0	14.6 - 51.9	11	37.9	22.7 - 56.0
Missing	*0*			*0*			*0*		

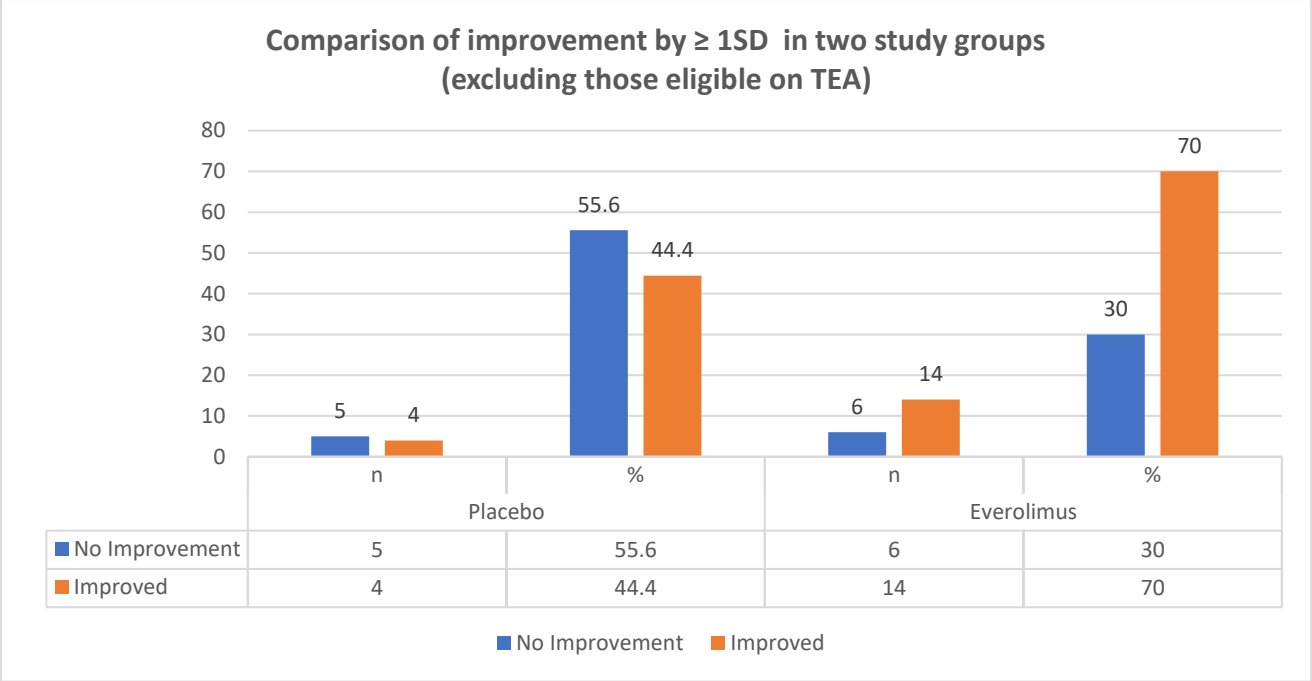


Figure 10 Comparison of frequency of improvement (>1 SD) between Placebo vs Everolimus

2.17.1 Analyses to adjust for participants eligible on TEA

Excluding the 7 participants eligible based only on the TEA test at screening, the proportion of participants in the placebo arm that demonstrated an improvement of a least 1SD in at least one measure between baseline and 6 months was 44.4%, and the proportion of participants in the everolimus arm that demonstrated an improvement was 70.0% (Table 13).

This is a difference of 25.6% points, which meets the pre-specified threshold of a 20-percentage point difference in the intervention group, supporting further investigation of everolimus as a treatment for neurocognitive problems in TSC. Figure 10 depicts the changes across the two groups. While a larger fraction of participants in the everolimus group showed an improvement of at least 1SD in a primary outcome measure, the small numbers are associated with wide and overlapping 95% CIs (Table 13).

A further sensitivity analysis was performed in the whole study population, including participants eligible on the TEA alone, and is summarised in table 14. For this analysis, change was determined between scores on TEA tests performed at screening (rather than baseline) and at 6 months so that the same version (A) was being used and scores could be compared (Table 14). With the high

percentage of improvement in the placebo arm (83.3%), it is not possible to implement a one-sample χ^2 test for 20 percentage points of improvement in the Everolimus arm. The percentage of patients who improved in any 10 measures over time is lower in the Everolimus arm (79.2%).

Table 14 Primary Outcome: Improvement from Screening to 6 months

	Placebo n=12			Everolimus n=24			Total n=36		
Sensitivity analysis: Whole TRON population									
	n	%	95% CI	n	%	95% CI	n	%	95% CI
Yes	10	83.3	55.2 to 95.3	19	79.2	59.5 to 90.8	29	80.6	65.0 to 90.3
No	2	16.7	4.7 to 44.8	5	20.8	9.2 to 40.5	7	19.4	9.8 to 35.0
Missing	*0*			*0*			*0*		

2.17.2 Subgroup analyses

Two subgroup analyses were performed to explore differential treatment effects on the primary outcome of improvement over time by age and IQ. Due to the small numbers of patients that did not improve over the 6-month period, no significant difference was observed between the groups.

2.17.3 Primary outcome measures: performance scores

This section presents the TRON study participants' performance at the study visits (timepoints) where neuropsychological assessments were completed; The analysis is presented illustrated with line diagrams to show the temporal performance of the average score of the participants across the various study time-points. This descriptive presentation is provided to illustrate how the participants' scores varied over the study period. The mean (raw/standard) performance scores of participants over time in both study arms for the subtests of the BMIPB, CANTAB & TEA neuropsychological tests (used for primary outcome analysis) are detailed in table 15, while figures 11 - 20 illustrate the performance scores in the primary outcome measures at screening, baseline, 4-week, 12-week, 24-week and 36-week assessments in the placebo and everolimus arms .

Table 15 Performance (scores) in primary outcome measures at study visits (where completed)

	Arm						Total		
	Placebo			Everolimus					
	n	Mean	SD	n	Mean	SD	n	Mean	SD
BMIPB List Learning Immediate Recall (Raw Score)									
Screening	13	38.5	9.01	25	40.4	10.25	38	39.7	9.76
Baseline	12	41.3	8.98	25	40.4	9.61	37	40.7	9.30
Week 4	12	39.3	14.78	24	41.5	9.71	36	40.8	11.48
Week 12	12	45.4	10.54	24	42.0	11.00	36	43.1	10.82
Week 24	12	43.7	9.17	24	44.6	10.52	36	44.3	9.97
Week 36	12	43.7	12.16	23	41.1	8.53	35	42.0	9.82
BMIPB List Learning Delayed Recall (Raw Score)									
Screening	13	8.5	2.90	25	8.9	3.15	38	8.8	3.04
Baseline	12	8.2	3.07	25	6.8	2.98	37	7.3	3.03
Week 4	12	8.3	3.23	24	7.8	3.27	36	8.0	3.22
Week 12	12	8.8	3.33	24	8.4	3.28	36	8.5	3.26
Week 24	12	8.8	3.11	24	8.8	3.04	36	8.8	3.02
Week 36	12	8.6	3.12	23	8.1	2.41	35	8.3	2.64
BMIPB Complex Figure Test Immediate Recall (Raw Score)									
Screening	13	59.5	11.98	25	52.4	17.72	38	54.8	16.18
Baseline	12	58.3	11.80	25	48.4	18.80	37	51.6	17.34
Week 4	12	58.9	14.71	24	55.5	16.33	36	56.6	15.68
Week 12	12	54.1	18.01	24	60.1	17.96	36	58.1	17.95
Week 24	12	61.5	11.90	24	63.5	20.90	36	62.8	18.23
Week 36	12	59.6	12.62	23	56.3	17.13	35	57.5	15.62
BMIPB Complex Figure Test Delayed Recall (Raw Score)									
Screening	13	56.2	13.52	25	45.9	14.88	38	49.4	15.08

	Arm						Total		
	Placebo			Everolimus					
	n	Mean	SD	n	Mean	SD	n	Mean	SD
Baseline	12	56.8	8.62	25	51.2	15.44	37	53.1	13.74
Week 4	12	52.6	12.31	24	53.6	14.24	36	53.3	13.46
Week 12	12	48.2	19.08	24	60.1	13.44	36	56.1	16.29
Week 24	12	59.4	13.19	24	58.1	16.31	36	58.5	15.16
Week 36	12	59.4	13.19	24	58.1	16.31	36	58.5	15.16
SOC (CANTAB) Mean initial thinking time (Standard Score)									
Screening	13	0.518	1.1983	25	0.711	0.8726	38	0.645	0.9840
Baseline	12	0.956	0.2530	24	0.928	0.4089	36	0.937	0.3608
Week 4	12	1.017	0.2851	24	0.899	0.3490	36	0.938	0.3297
Week 12	12	0.883	0.3085	24	0.890	0.5802	36	0.888	0.5011
Week 24	12	0.927	0.3158	24	0.829	0.3604	36	0.862	0.3448
Week 36	12	1.098	0.1491	23	0.893	0.5137	35	0.964	0.4333
SOC (CANTAB) Mean subsequent thinking time (Standard Score)									
Screening	13	-0.778	3.5370	25	-0.385	1.0618	38	-0.519	2.1965
Baseline	12	0.349	0.3849	24	0.275	0.5465	36	0.299	0.4941
Week 4	12	0.512	0.2085	24	0.276	0.4565	36	0.355	0.4041
Week 12	12	-0.218	2.7340	24	0.270	0.6205	36	0.107	1.6299
Week 24	12	0.363	0.5144	24	0.038	0.8708	36	0.146	0.7782
Week 36	12	0.586	0.1790	23	0.210	0.5756	35	0.339	0.5074
SOC (CANTAB) Problems solved in minimum moves (Standard Score)									
Screening	13	-0.615	1.2578	25	-0.832	0.8319	38	-0.758	0.9864
Baseline	12	-0.143	0.7313	24	-0.298	0.7439	36	-0.247	0.7330
Week 4	12	-0.527	1.3493	24	-0.338	0.9291	36	-0.401	1.0712
Week 12	12	0.217	0.9744	24	-0.255	0.9167	36	-0.098	0.9495

	Arm						Total		
	Placebo			Everolimus					
	n	Mean	SD	n	Mean	SD	n	Mean	SD
Week 24	12	-0.147	1.0466	24	-0.490	0.7773	36	-0.376	0.8765
Week 36	12	-0.209	1.2799	23	-0.402	0.9064	35	-0.336	1.0345
SWM (CANTAB) Between-Errors (Standard Score)									
Screening	13	-0.560	1.3024	25	-0.454	0.9526	38	-0.491	1.0683
Baseline	12	-0.338	0.7811	25	-0.516	1.0067	37	-0.458	0.9323
Week 4	12	-0.358	1.3334	24	-0.399	1.1322	36	-0.385	1.1838
Week 12	12	-0.513	1.0711	24	-0.394	1.0350	36	-0.433	1.0333
Week 24	12	-0.259	1.0964	24	-0.598	0.8303	36	-0.485	0.9258
Week 36	12	-0.227	1.0275	23	-0.520	1.1524	35	-0.419	1.1049
SWM (CANTAB) Strategy Fields (Standard Score)									
Screening	13	-0.400	0.9095	25	-0.356	0.9301	38	-0.371	0.9110
Baseline	12	-0.169	1.3260	25	-0.478	0.7872	37	-0.378	0.9858
Week 4	12	-0.200	1.1393	24	-0.222	0.9423	36	-0.214	0.9957
Week 12	12	-0.155	0.9438	24	-0.233	1.0412	36	-0.207	0.9969
Week 24	12	-0.013	1.2365	24	-0.479	0.8690	36	-0.324	1.0130
Week 36	12	-0.064	1.1040	23	-0.328	0.8146	35	-0.237	0.9164
TEA Telephone search whilst counting (Scaled Score)									
Screening	13	6.4	2.63	24	6.4	3.75	37	6.4	3.36
Baseline	11	4.4	3.53	25	5.8	3.74	36	5.4	3.69
Week 4	12	8.0	2.76	24	8.3	4.23	36	8.2	3.77
Week 12	12	5.0	4.07	24	6.0	4.56	36	5.7	4.37
Week 24	12	7.0	2.98	24	7.7	3.61	36	7.5	3.38
Week 36	12	6.9	4.62	22	6.4	4.56	34	6.6	4.52

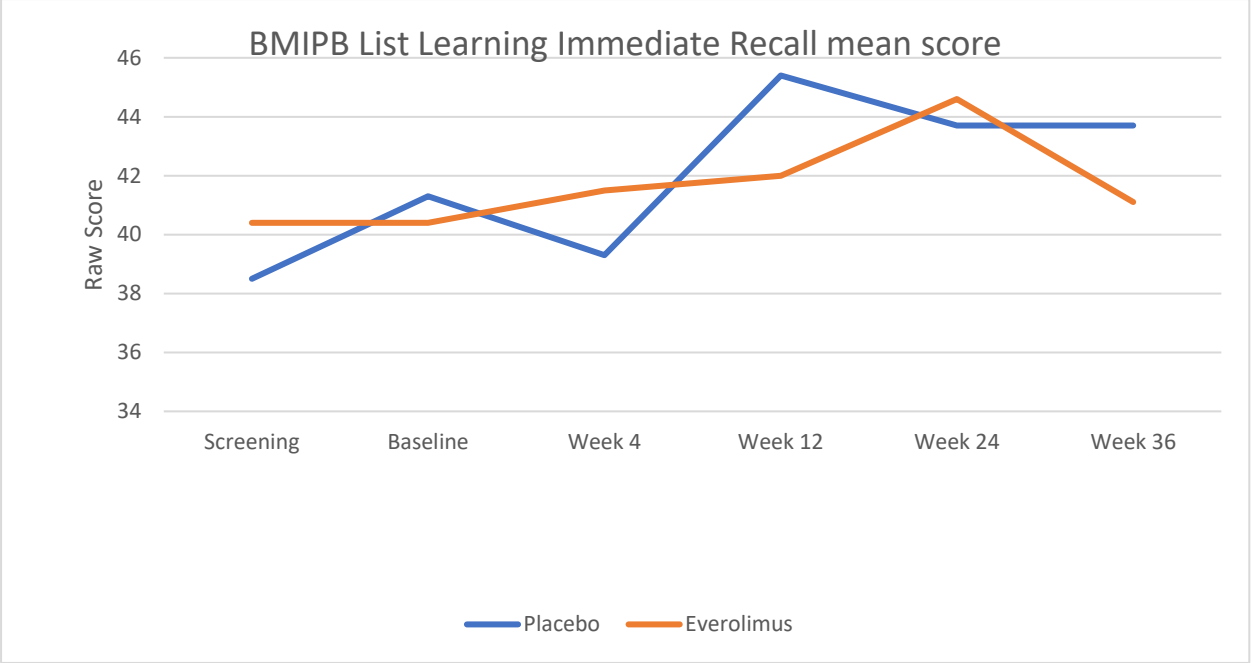


Figure 11 Line chart portraying BMIPB List Learning Immediate Recall mean (Raw) scores

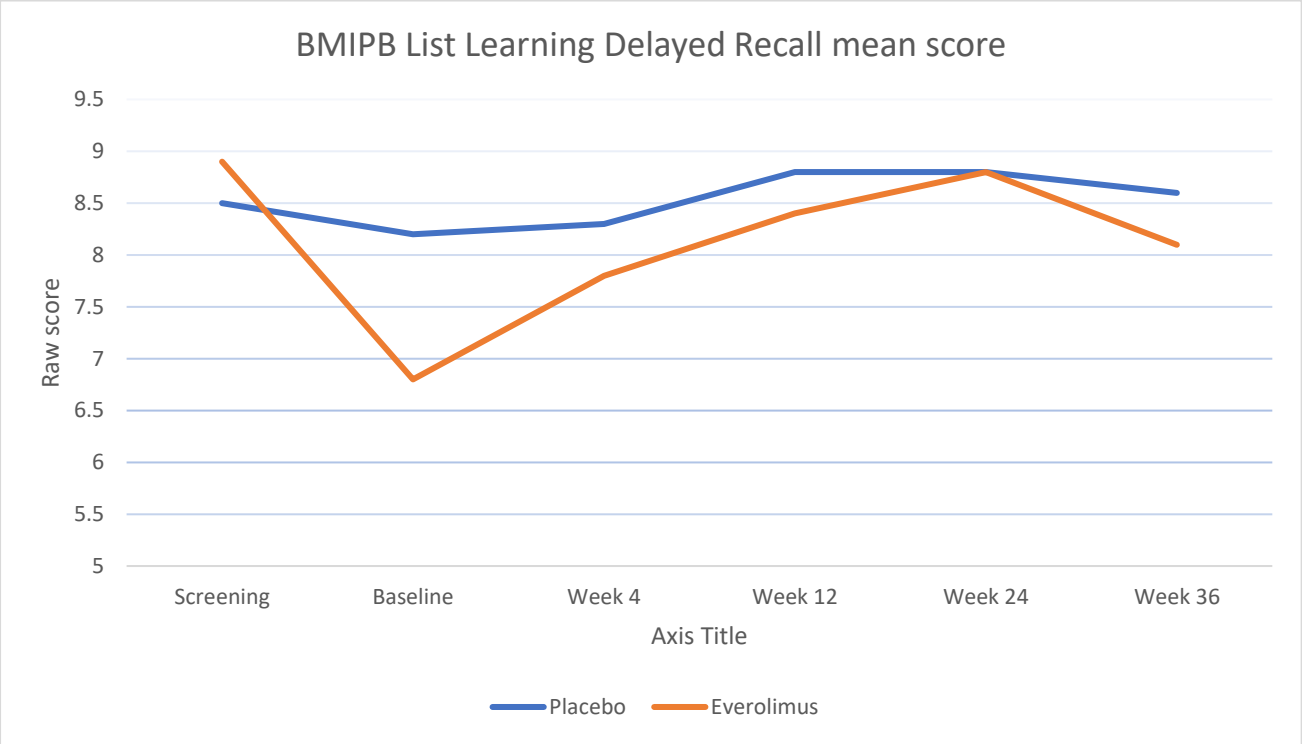


Figure 12 Line chart depicting BMIPB List Learning Delayed Recall mean (Raw) score

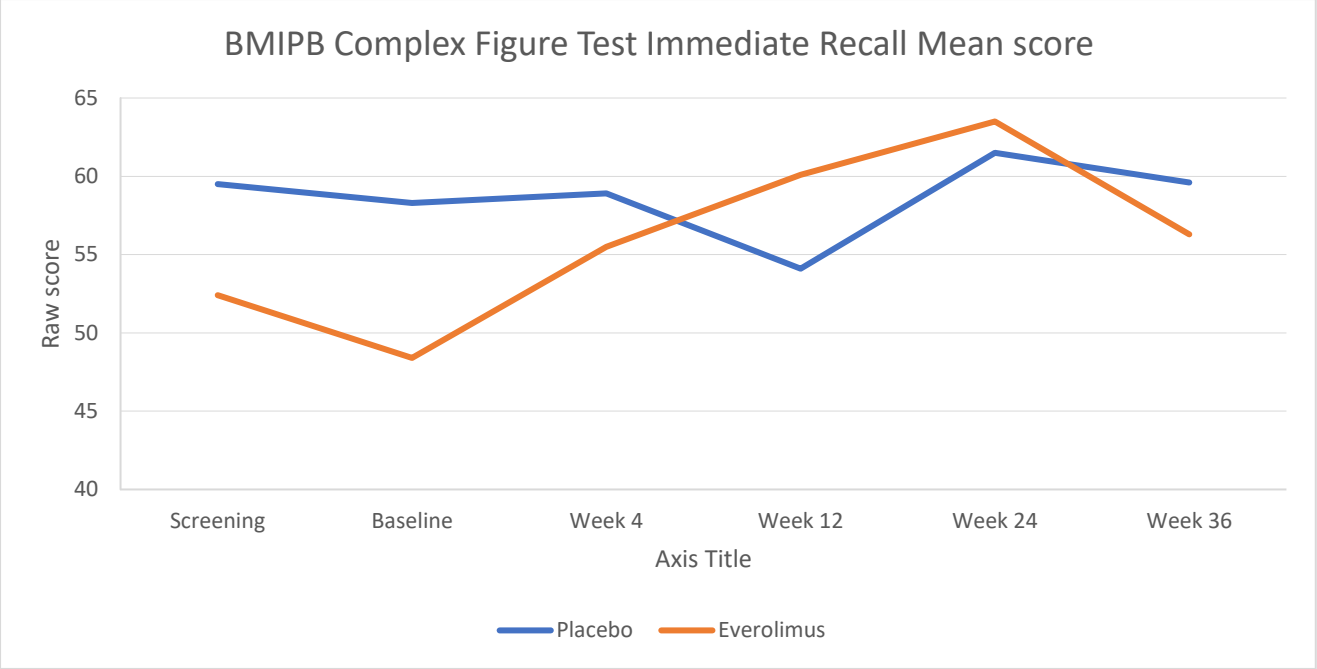


Figure 13 Line chart showing BMIPB Complex Figure Test Immediate Recall Mean (Raw) score

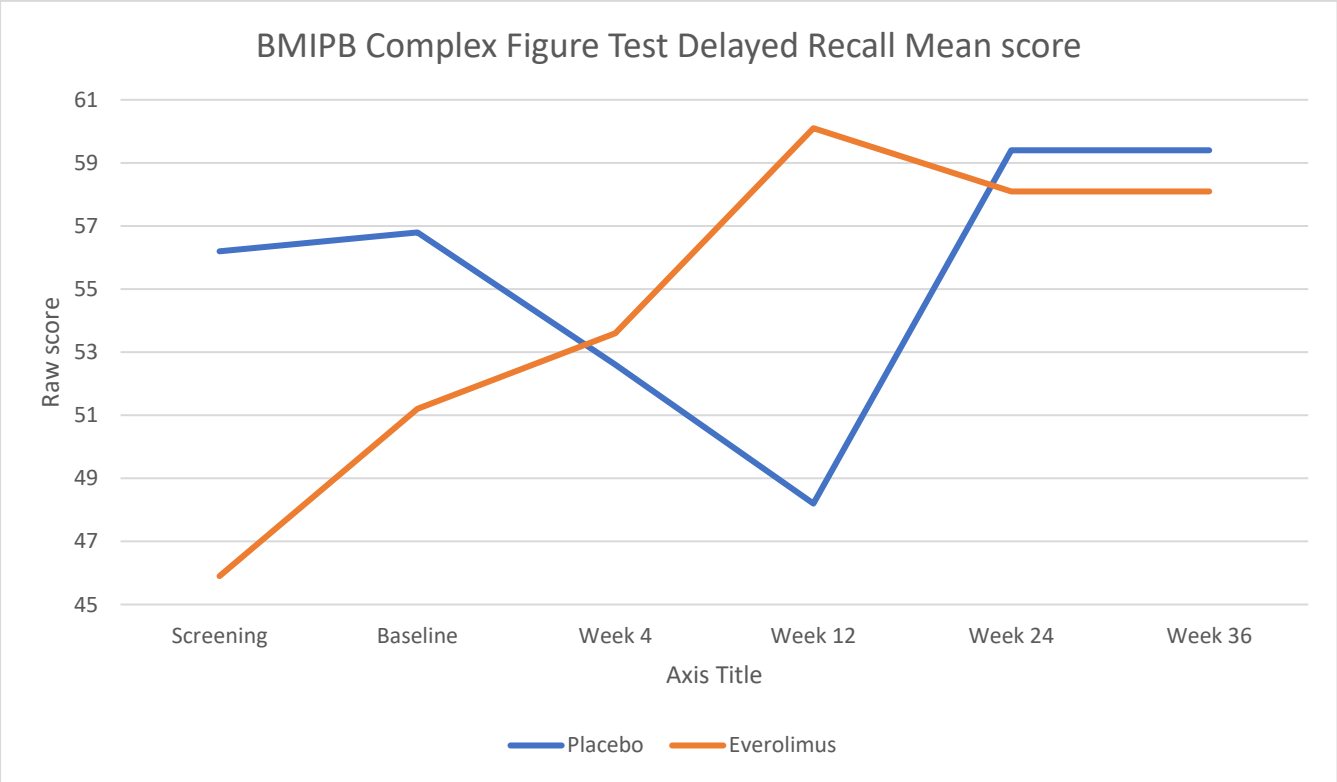


Figure 14 Line chart describing BMIPB Complex Figure Test Delayed Recall Mean (Raw) score

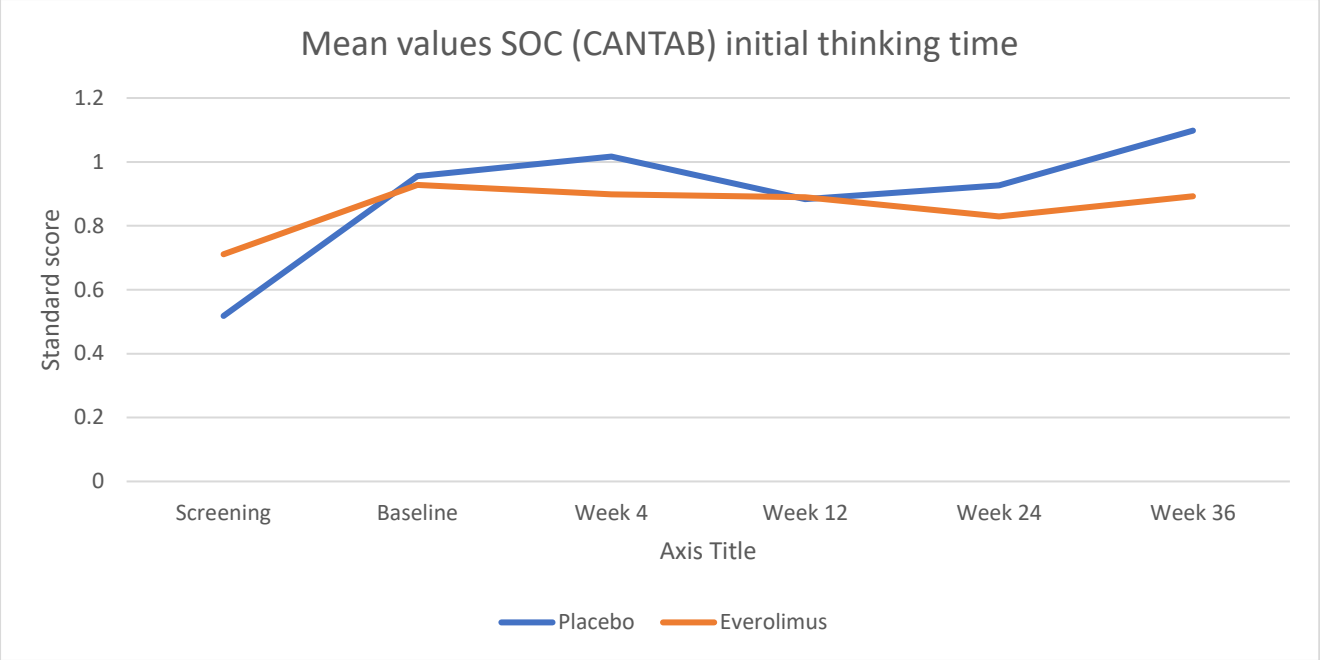


Figure 15 Line chart displaying mean values SOC (CANTAB) - Initial thinking time

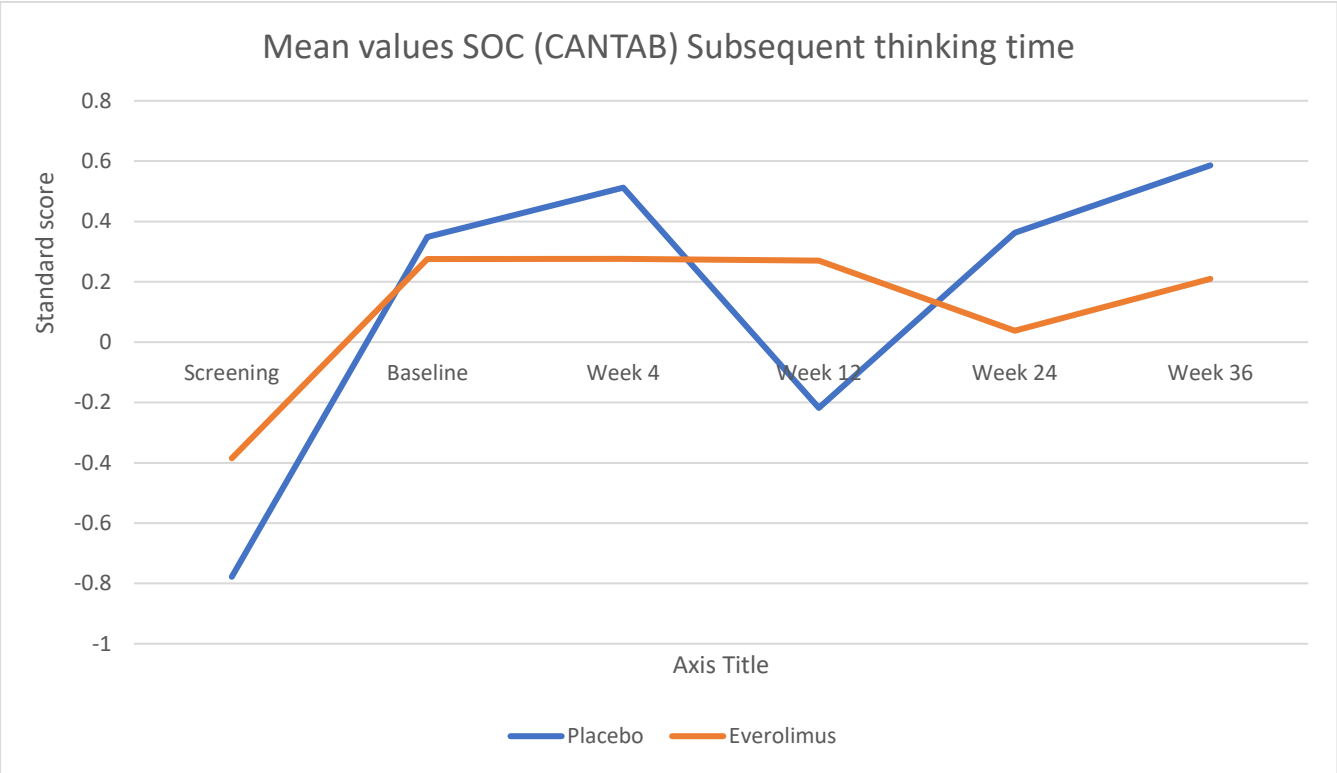


Figure 16 Line chart representing mean values SOC (CANTAB) – Subsequent thinking time

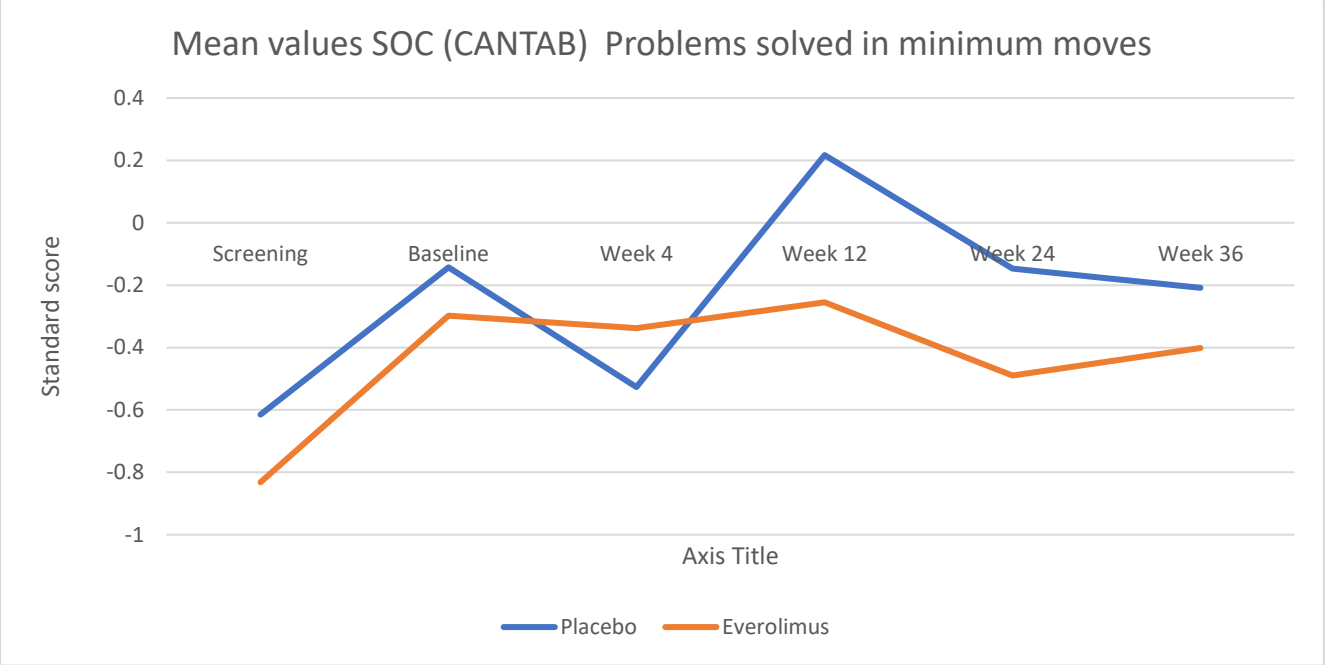


Figure 17 Line chart indicating mean values SOC (CANTAB) - Minimum moves

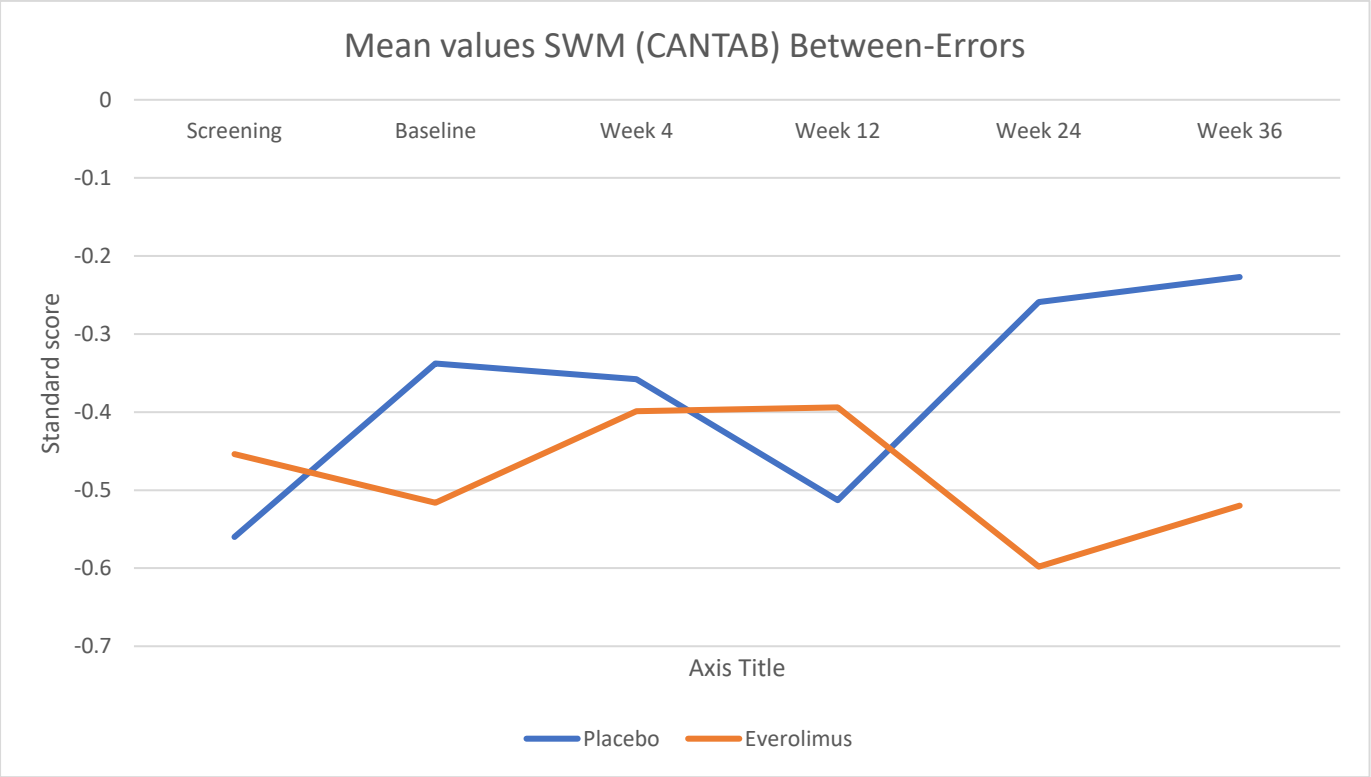


Figure 18 Line chart demonstrating mean values SWM (CANTAB) Between-Errors

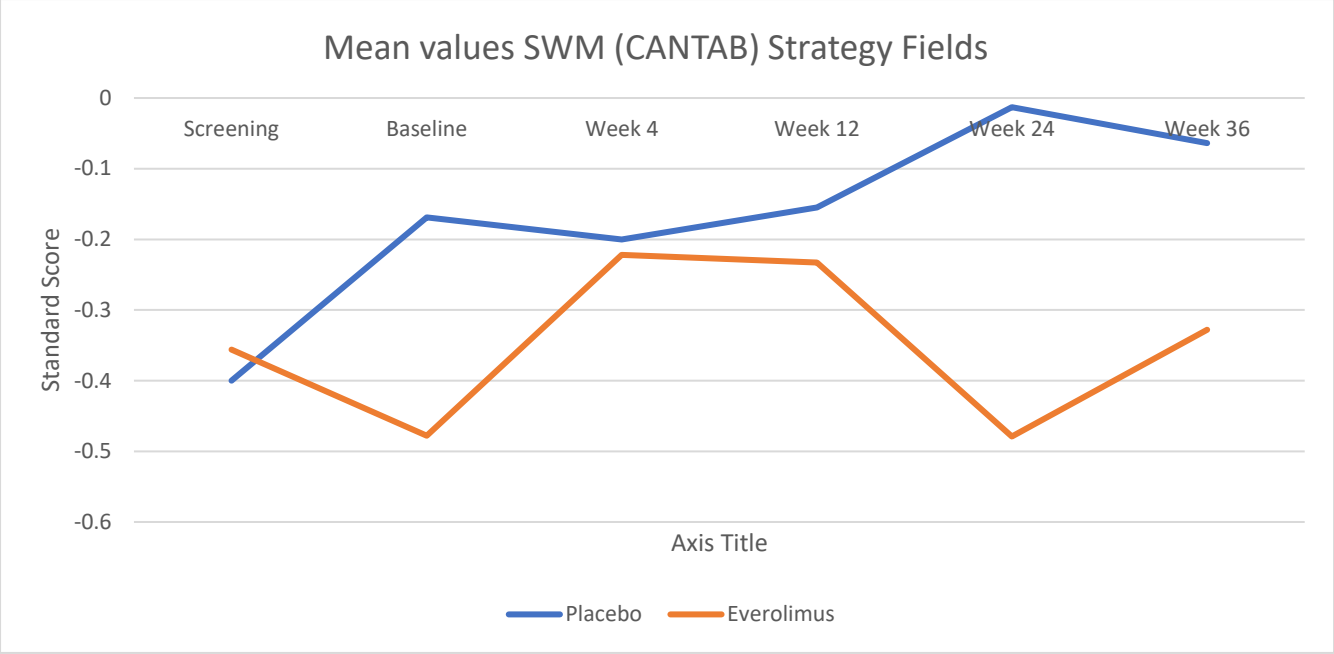


Figure 19 Line chart showing mean values SWM (CANTAB) Strategy Fields

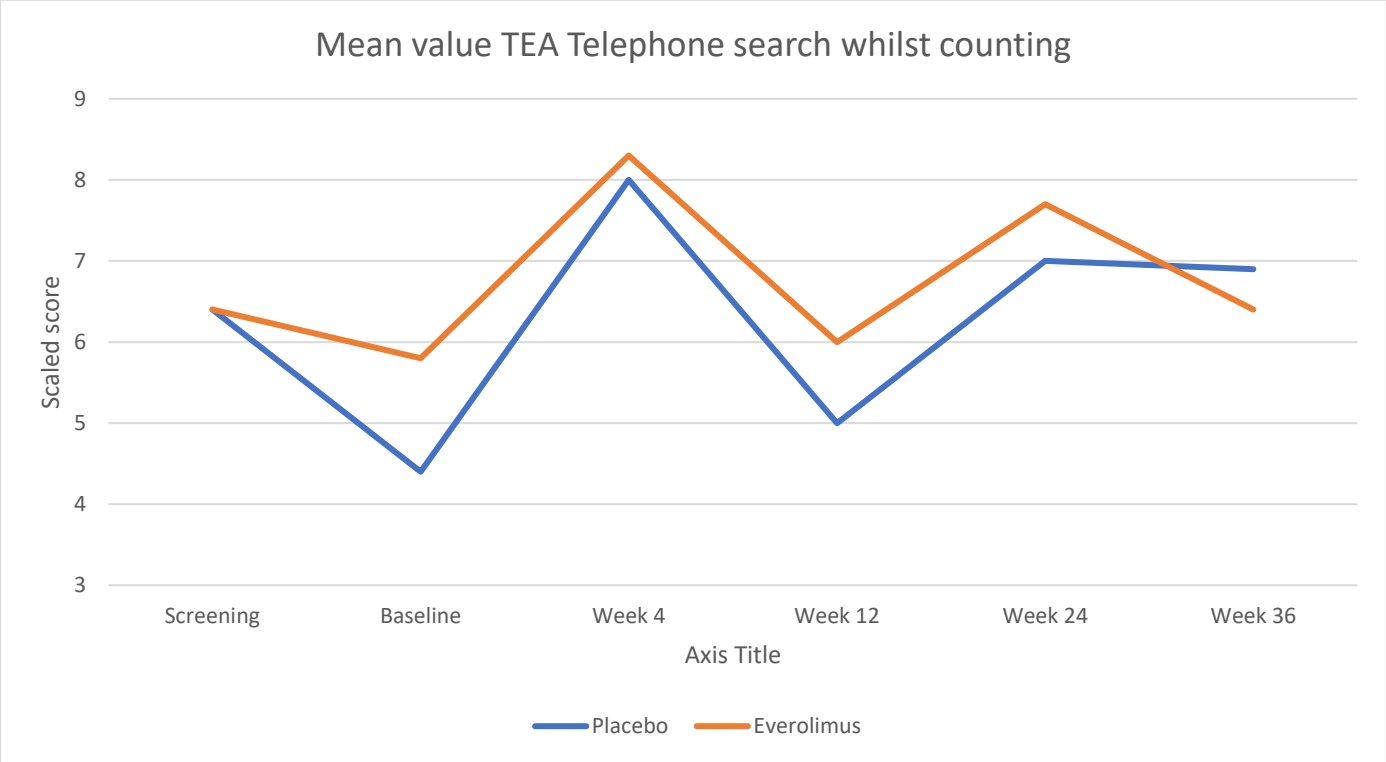


Figure 20 Line chart depicting mean (scaled) score TEA Telephone search whilst counting

2.17.4 Primary Outcome assessments: Improvement by ≥ 1SD

Tables 16 and figures 21-30 show the numbers of participants improving (Yes) or not improving (No) in their performance scores by ≥1SD at follow up assessments (timepoints) in both study arms for the subtests of the BMIPB, CANTAB & TEA neuropsychological tests (used for Primary Outcome analysis). Treatment ceased at 24 weeks and the planned 36-week assessment was undertaken to see if any changes occurring during treatment were sustained after its cessation.

Table 16 Improvement from baseline (≥1SD) in Neuropsychological scores at study visit timepoints (whole-population)

		Arm				Total	
		Placebo		Everolimus			
		n	%	n	%	n	%
BMIPB List Learning Immediate Recall (Raw Score)							
Week 4	No	11	91.7	23	95.8	34	94.4
	Yes	1	8.3	1	4.2	2	5.6
Week 12	No	10	83.3	23	95.8	33	91.7
	Yes	2	16.7	1	4.2	3	8.3
Week 24	No	10	83.3	20	83.3	30	83.3
	Yes	2	16.7	4	16.7	6	16.7
Week 36	No	10	83.3	22	95.7	32	91.4
	Yes	2	16.7	1	4.3	3	8.6
BMIPB List Learning Delayed Recall (Raw Score)							
Week 4	No	11	91.7	19	79.2	30	83.3
	Yes	1	8.3	5	20.8	6	16.7
Week 12	No	8	66.7	15	62.5	23	63.9
	Yes	4	33.3	9	37.5	13	36.1
Week 24	No	10	83.3	16	66.7	26	72.2
	Yes	2	16.7	8	33.3	10	27.8
Week 36	No	9	75.0	15	65.2	24	68.6
	Yes	3	25.0	8	34.8	11	31.4
BMIPB Complex Figure Test Immediate Recall (Raw Score)							
Week 4	No	10	83.3	16	66.7	26	72.2
	Yes	2	16.7	8	33.3	10	27.8

Week 12	No	9	75.0	15	62.5	24	66.7
	Yes	3	25.0	9	37.5	12	33.3
Week 24	No	9	75.0	11	45.8	20	55.6
	Yes	3	25.0	13	54.2	16	44.4
Week 36	No	10	83.3	17	73.9	27	77.1
	Yes	2	16.7	6	26.1	8	22.9
BMIPB Complex Figure Test Delayed Recall (Raw Score)							
Week 4	No	12	100.0	22	91.7	34	94.4
	Yes	0	0.0	2	8.3	2	5.6
Week 12	No	12	100.0	18	75.0	30	83.3
	Yes	0	0.0	6	25.0	6	16.7
Week 24	No	12	100.0	22	91.7	34	94.4
	Yes	0	0.0	2	8.3	2	5.6
Week 36	No	12	100.0	23	100.0	35	100.0
	Yes	0	0.0	0	0.0	0	0.0
SOC (CANTAB) Mean initial thinking time (Standard Score)							
Week 4	No	12	100.0	22	95.7	34	97.1
	Yes	0	0.0	1	4.3	1	2.9
Week 12	No	12	100.0	22	95.7	34	97.1
	Yes	0	0.0	1	4.3	1	2.9
Week 24	No	12	100.0	22	95.7	34	97.1
	Yes	0	0.0	1	4.3	1	2.9
Week 36	No	12	100.0	21	95.5	33	97.1
	Yes	0	0.0	1	4.5	1	2.9
SOC (CANTAB) Mean subsequent thinking time (Standard Score)							
Week 4	No	12	100.0	21	91.3	33	94.3
	Yes	0	0.0	2	8.7	2	5.7
Week 12	No	11	91.7	22	95.7	33	94.3
	Yes	1	8.3	1	4.3	2	5.7
Week 24	No	12	100.0	23	100.0	35	100.0
	Yes	0	0.0	0	0.0	0	0.0
Week 36	No	12	100.0	21	95.5	33	97.1
	Yes	0	0.0	1	4.5	1	2.9
SOC (CANTAB) Problems solved in minimum moves (Standard Score)							
Week 4	No	11	91.7	22	95.7	33	94.3
	Yes	1	8.3	1	4.3	2	5.7
	No	11	91.7	21	91.3	32	91.4

Week 12	Yes	1	8.3	2	8.7	3	8.6
Week 24	No	10	83.3	22	95.7	32	91.4
	Yes	2	16.7	1	4.3	3	8.6
Week 36	No	10	83.3	20	90.9	30	88.2
	Yes	2	16.7	2	9.1	4	11.8
SWM (CANTAB) Between-Errors (Standard Score)							
Week 4	No	11	91.7	22	91.7	33	91.7
	Yes	1	8.3	2	8.3	3	8.3
Week 12	No	12	100.0	22	91.7	34	94.4
	Yes	0	0.0	2	8.3	2	5.6
Week 24	No	11	91.7	22	91.7	33	91.7
	Yes	1	8.3	2	8.3	3	8.3
Week 36	No	11	91.7	21	91.3	32	91.4
	Yes	1	8.3	2	8.7	3	8.6
SWM (CANTAB) Strategy Fields (Standard Score)							
Week 4	No	12	100.0	21	87.5	33	91.7
	Yes	0	0.0	3	12.5	3	8.3
Week 12	No	10	83.3	21	87.5	31	86.1
	Yes	2	16.7	3	12.5	5	13.9
Week 24	No	11	91.7	23	95.8	34	94.4
	Yes	1	8.3	1	4.2	2	5.6
Week 36	No	10	83.3	18	78.3	28	80.0
	Yes	2	16.7	5	21.7	7	20.0
TEA Telephone search whilst counting (Scaled Score)							
Week 4	No	5	45.5	15	62.5	20	57.1
	Yes	6	54.5	9	37.5	15	42.9
Week 12	No	9	81.8	16	66.7	25	71.4
	Yes	2	18.2	8	33.3	10	28.6
Week 24	No	5	45.5	14	58.3	19	54.3
	Yes	6	54.5	10	41.7	16	45.7
Week 36	No	4	36.4	14	63.6	18	54.5
	Yes	7	63.6	8	36.4	15	45.5

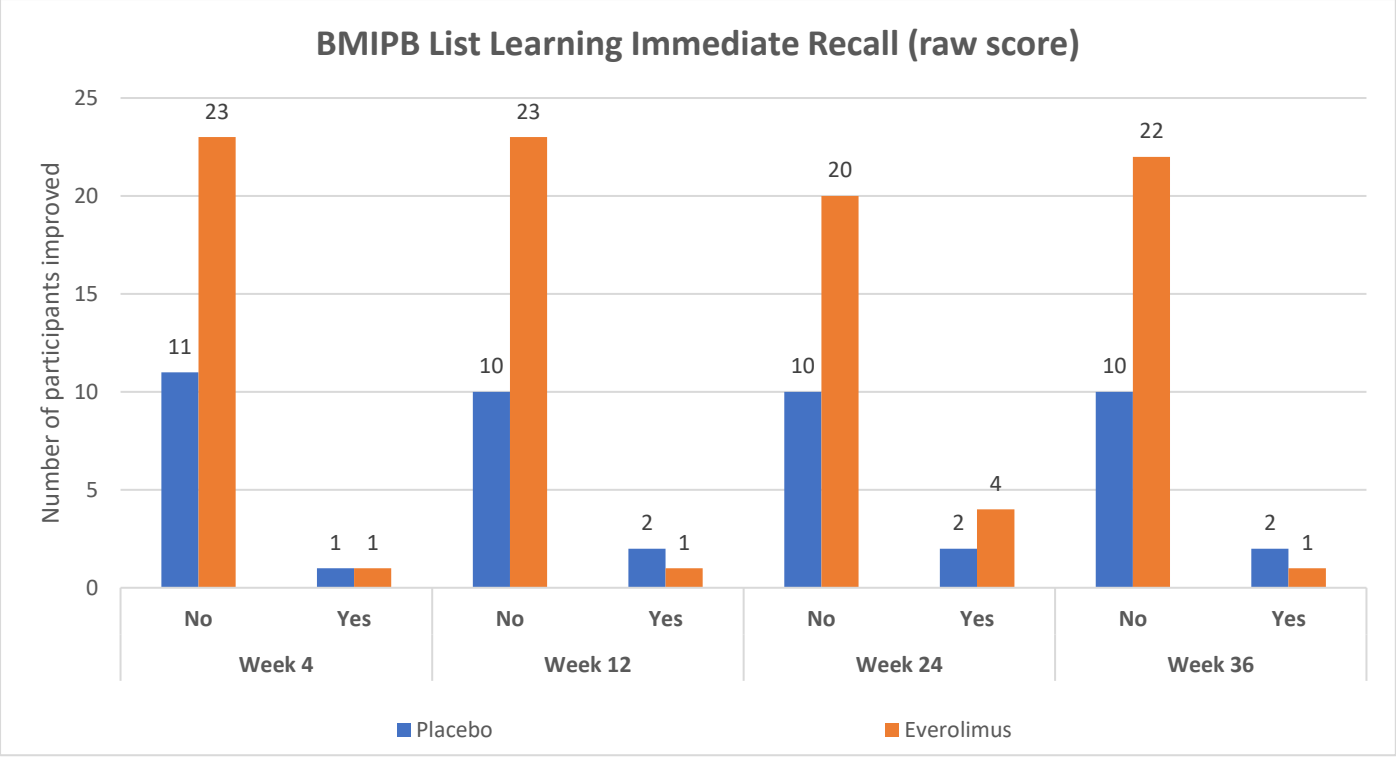


Figure 21 BMIPB List Learning Immediate Recall improvement ≥1SD

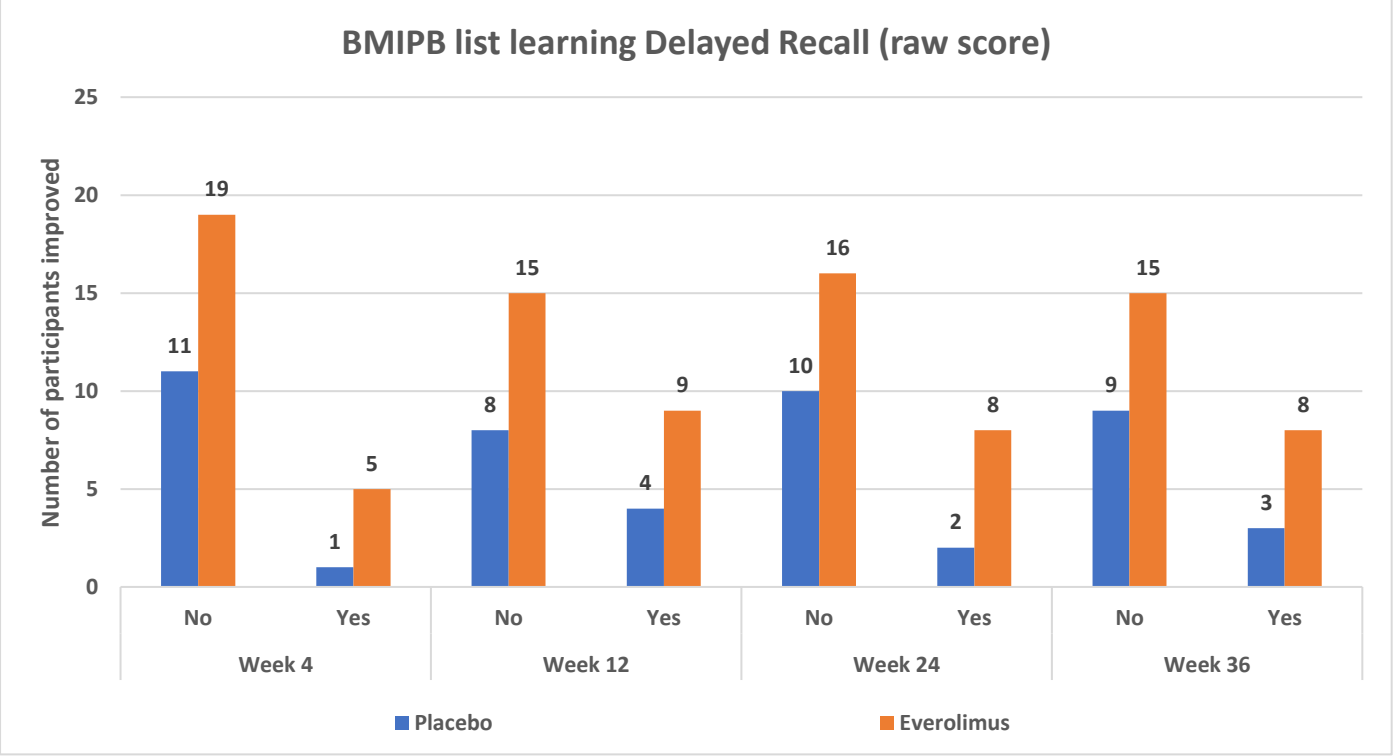


Figure 22 BMIPB list learning Delayed Recall improvement ≥1SD

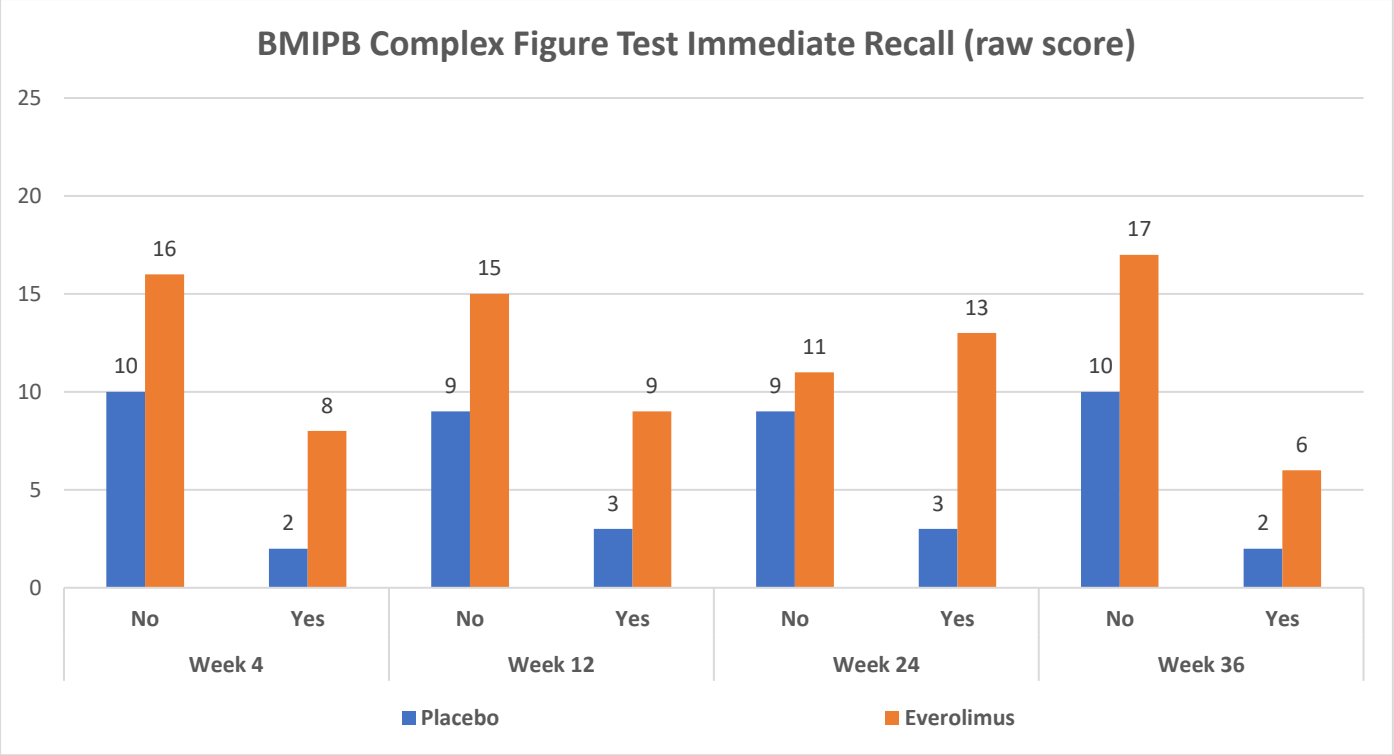


Figure 23 BMIPB Complex Figure Test Immediate Recall improvement ≥1SD

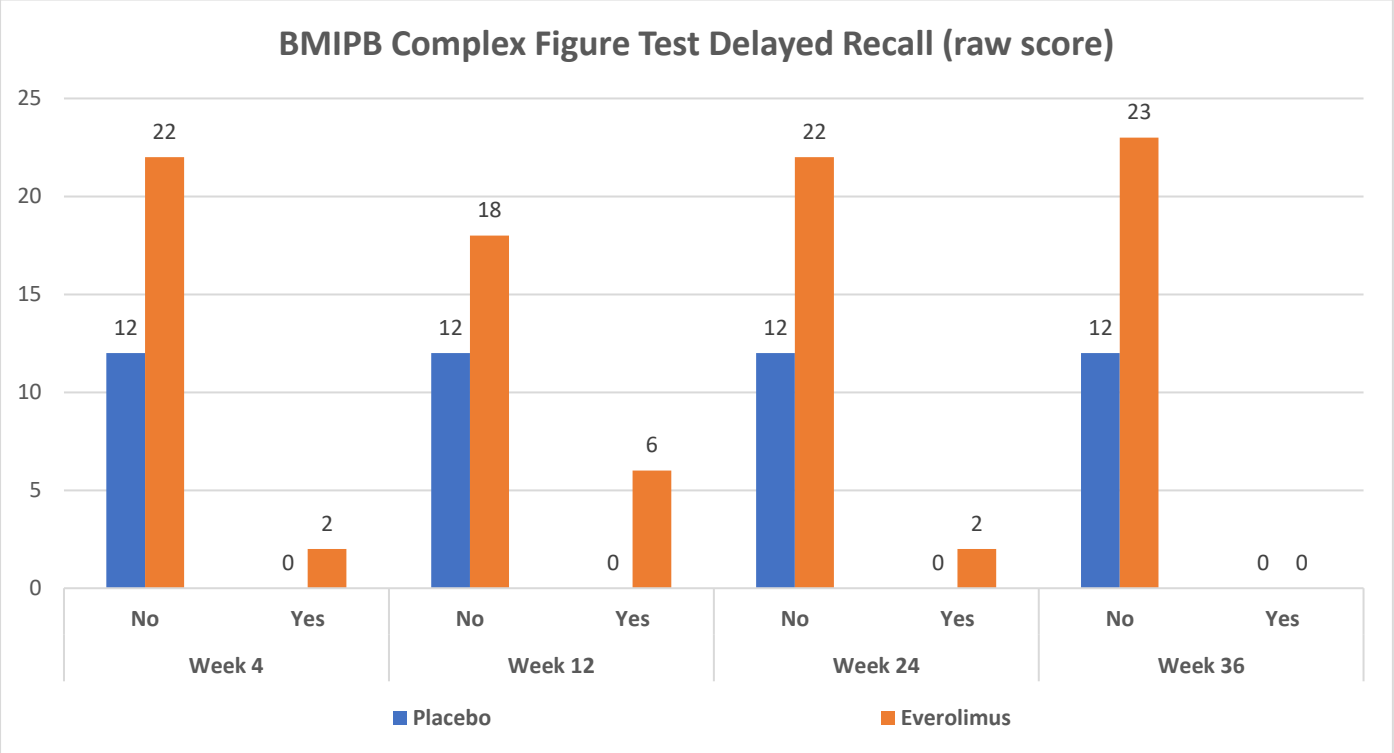


Figure 24 BMIPB Complex Figure Test Delayed Recall improvement ≥1SD

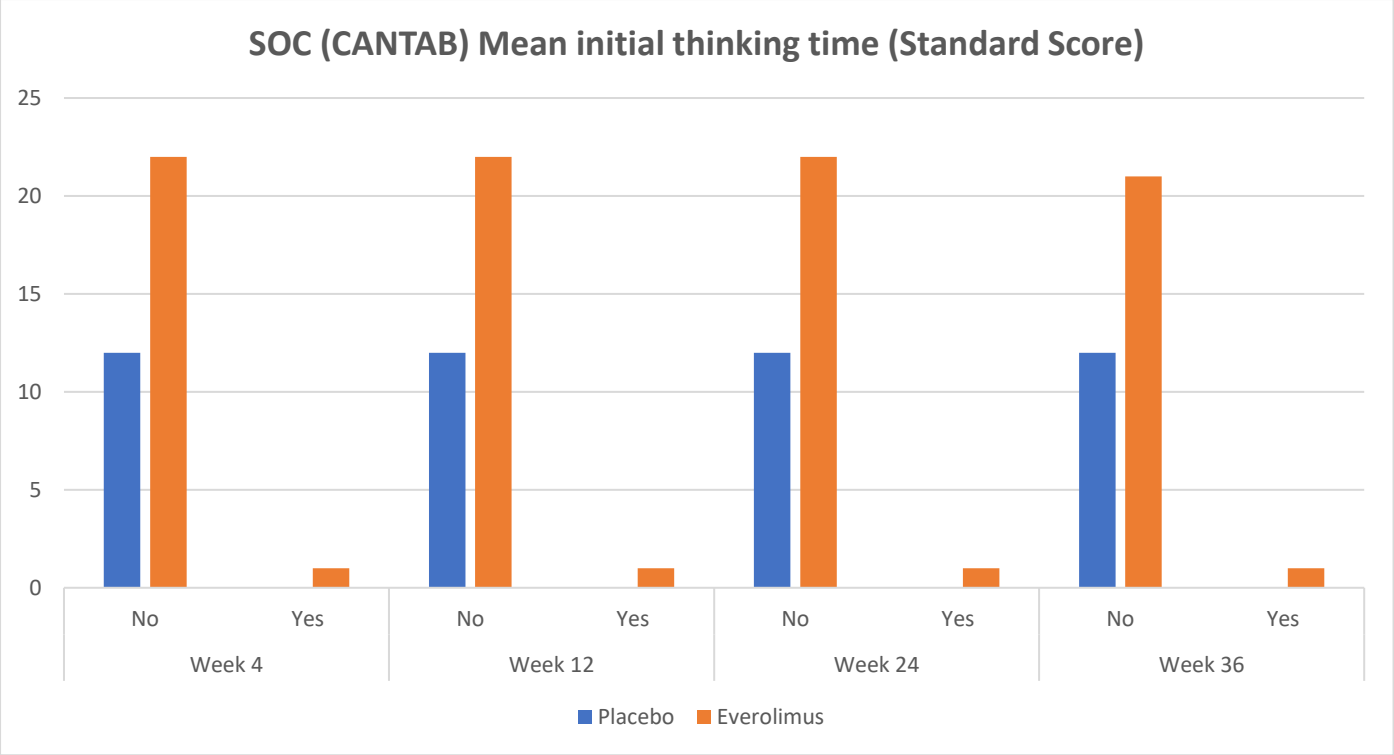


Figure 25 SOC (CANTAB) Mean initial thinking time improvement ≥1SD

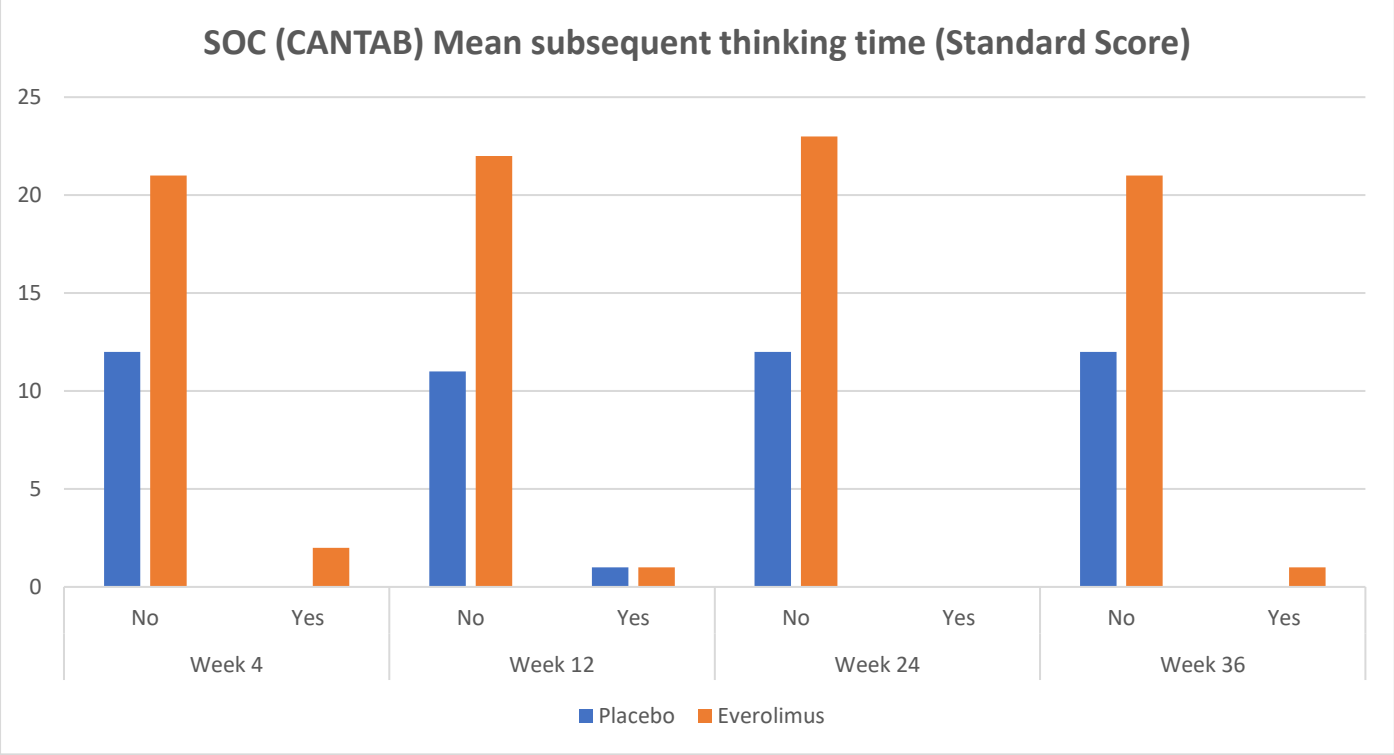


Figure 26 SOC (CANTAB) Mean subsequent thinking time improvement ≥1SD

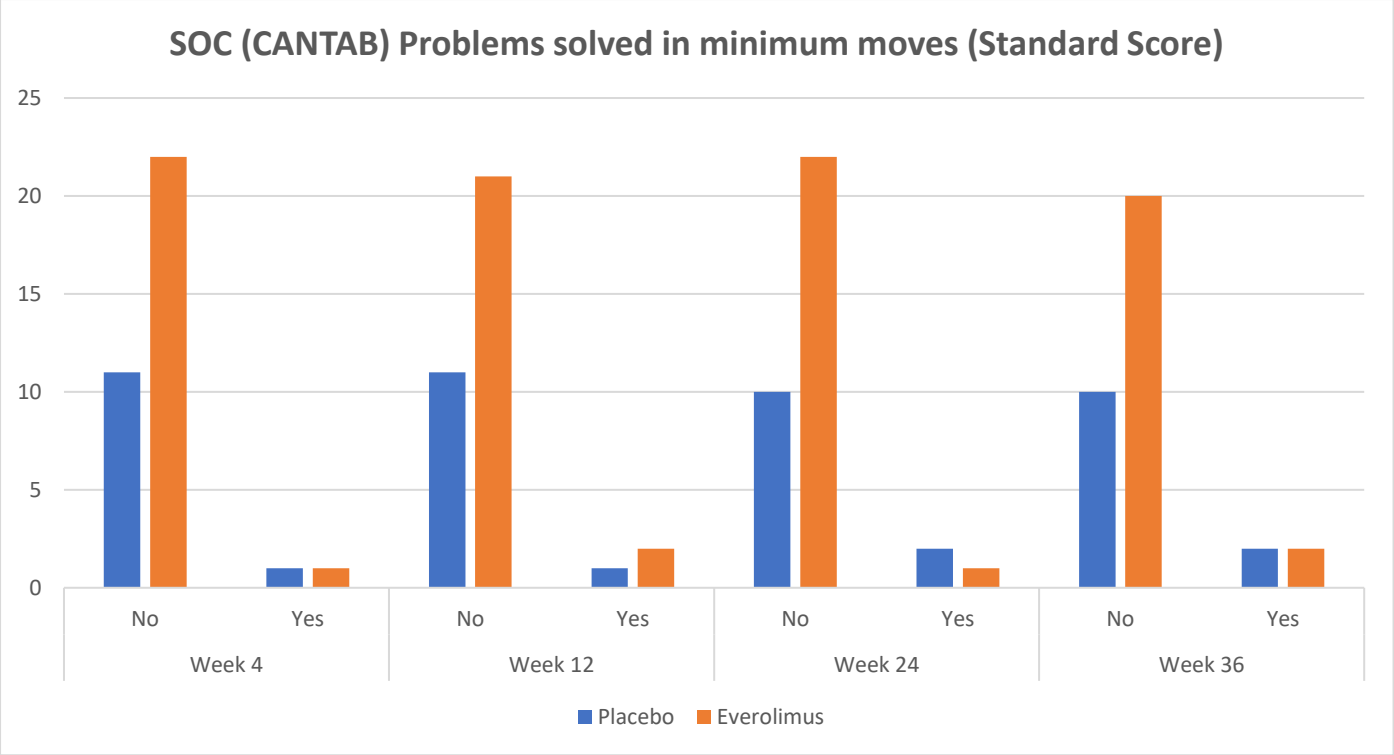


Figure 27 SOC (CANTAB) Problems solved in minimum moves improvement ≥1SD

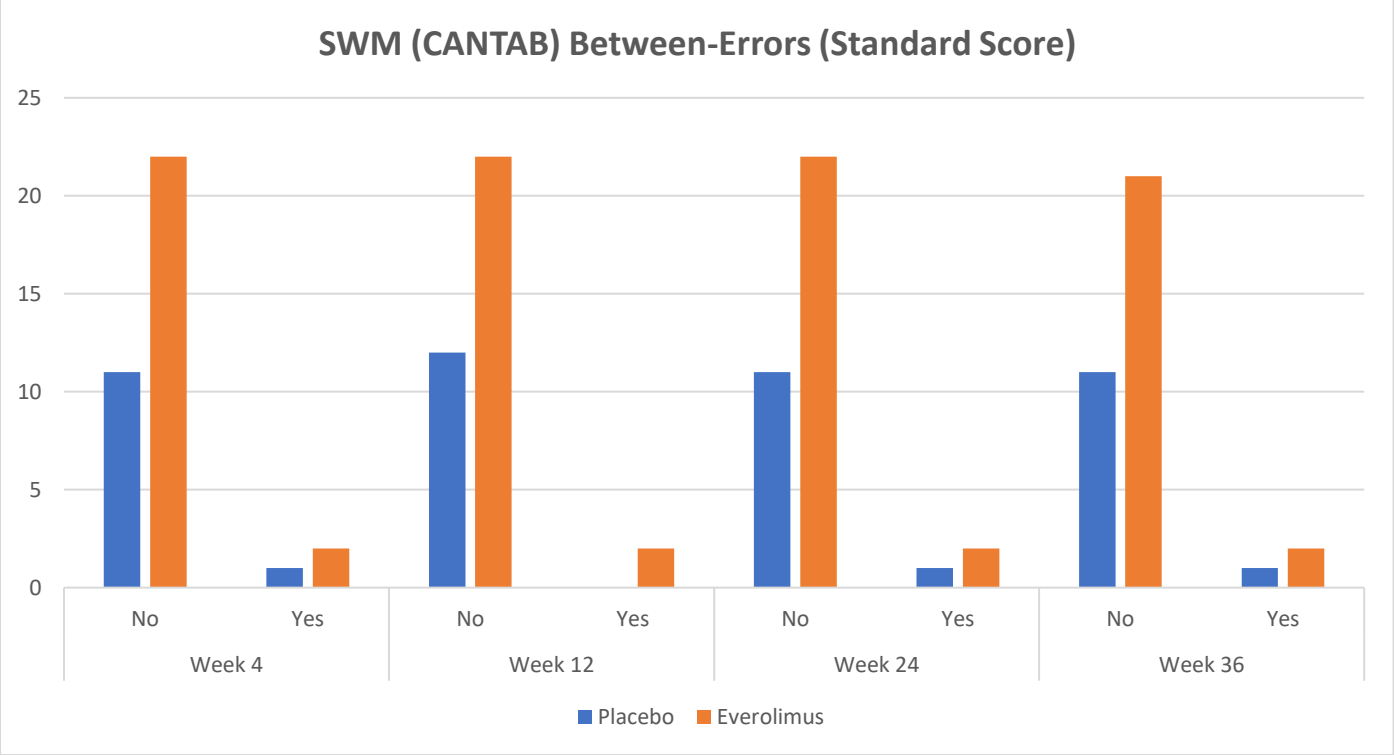


Figure 28 SWM (CANTAB) Between-Errors improvement ≥1SD

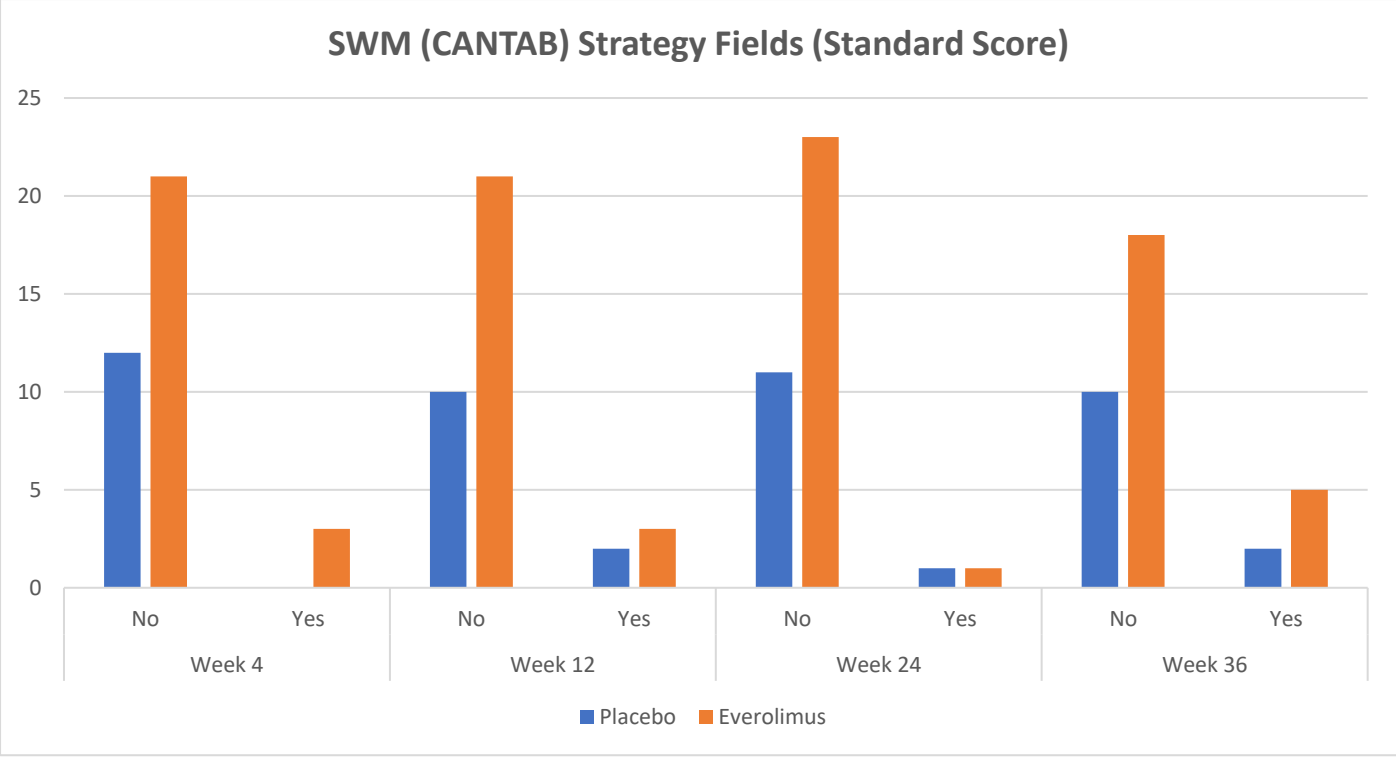


Figure 29 SWM (CANTAB) Strategy Fields improvement ≥1SD

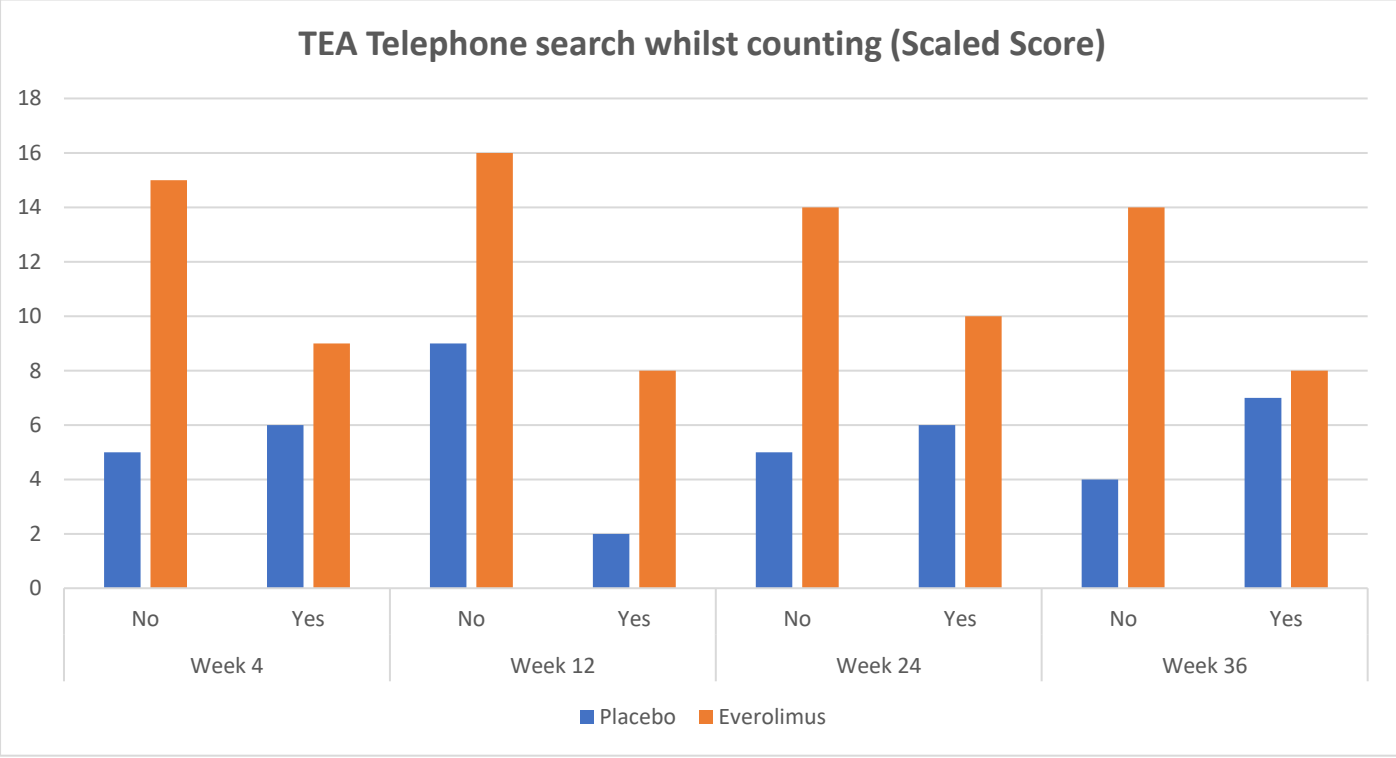


Figure 30 TEA Telephone search whilst counting improvement ≥1SD

2.17.5 Primary Outcome Assessments: Best response time point

Tables 17 provide descriptive and ordinal logistic regression results for differences between arms for the best response time point (Week 4, Week 12, Week 24 or Week 36). In the event of matched response at multiple time points, the first time point is selected. Ordinal Logistic regression (McCullagh 1980) is applicable for datasets where data occurs in an order such as Week 4,12, 24 and 36 (dependent variable) and tests its relationship with each independent variable (placebo group & everolimus group). The proportional odds model (Brant Test) for ordinal logistic regression portrayed, extends the binary logistic model's use to situations where the response variable takes on ordered categorical values (Brant 1990).

These analyses were done to detect any early response to everolimus (week 4 & 12) as well as retained changes at week 36, if any, after the intervention ceased at week 24. The analysis also provides the pattern of change over time in the neuropsychological measures by trial arm.

Red markings show that ordinal regression is not appropriate for these data. Yellow markings indicate significance at $\alpha = 0.05$. BMIPB Complex Figure Test Delayed Recall is significant (however the numbers are small) in all circumstances, with the best response earlier in the Everolimus arm (12 weeks) and later (24 -36 weeks) in the Placebo arm. Although this test is a measure of visuospatial or non-verbal free recall memory, the subject has to do sequencing in the background to accurately reproduce the shape of the object. This test is a difficult composite test not only of memory but also of executive function and therefore it remains difficult to draw clinically relevant conclusions from this result. In addition the mean raw score for placebo group is inexplicably reduced at 12 weeks (table 15 and figure 14) making these result doubtful.

Table 17 Primary Outcomes: Best response time point from baseline

		Arm				Total		Odds Ratio	95% CI	p-value	LR test *1	Brant test p-value *2
		Placebo		Everolimus								
		n	%	n	%	n	%					
BMIPB List Learning Immediate Recall	Week 4	2	16.7	6	25.0	8	22.2	1.145	0.313 to 4.181	0.838	0.467	0.564
	Week 12	5	41.7	5	20.8	10	27.8					
	Week 24	3	25.0	9	37.5	12	33.3					
	Week 36	2	16.7	4	16.7	6	16.7					
BMIPB List Learning Delayed Recall	Week 4	4	33.3	3	12.5	7	19.4	1.589	0.426 to 5.926	0.490	0.170	0.365
	Week 12	3	25.0	9	37.5	12	33.3					
	Week 24	2	16.7	6	25.0	8	22.2					
	Week 36	3	25.0	6	25.0	9	25.0					
BMIPB Complex Figure Test Immediate Recall	Week 4	2	16.7	4	16.7	6	16.7	0.762	0.207 to 2.814	0.684	0.425	0.633
	Week 12	3	25.0	7	29.2	10	27.8					
	Week 24	4	33.3	10	41.7	14	38.9					
	Week 36	3	25.0	3	12.5	6	16.7					
BMIPB Complex	Week 4	0	0.0	1	4.2	1	2.8	0.153	0.035 to 0.662	0.012	0.459	N/A
	Week 12	1	8.3	12	50.0	13	36.1					
	Week 24	7	58.3	8	33.3	15	41.7					

Figure Test Delayed Recall	Week 36	4	33.3	3	12.5	7	19.4					
SOC (CANTAB) Mean initial thinking time	Week 4	5	41.7	4	17.4	9	25.7	1.141	0.300 to 4.336	0.846	0.003	0.001
	Week 12	0	0.0	9	39.1	9	25.7					
	Week 24	3	25.0	4	17.4	7	20.0					
	Week 36	4	33.3	6	26.1	10	28.6					
SOC (CANTAB) Mean subsequent thinking time	Week 4	3	25.0	6	26.1	9	25.7	1.431	0.396 to 5.166	0.585	0.432	0.001
	Week 12	6	50.0	7	30.4	13	37.1					
	Week 24	0	0.0	4	17.4	4	11.4					
	Week 36	3	25.0	6	26.1	9	25.7					
SOC (CANTAB) Problems solved in minimum moves	Week 4	1	8.3	8	34.8	9	25.7	0.391	0.105 to 1.447	0.159	0.676	0.614
	Week 12	6	50.0	9	39.1	15	42.9					
	Week 24	4	33.3	3	13.0	7	20.0					
	Week 36	1	8.3	3	13.0	4	11.4					
SWM (CANTAB) Between-Errors	Week 4	5	41.7	8	33.3	13	36.1	1.169	0.328 to 4.160	0.810	0.084	<0.001
	Week 12	0	0.0	6	25.0	6	16.7					
	Week 24	5	41.7	2	8.3	7	19.4					
	Week 36	2	16.7	8	33.3	10	27.8					

SWM (CANTAB) Strategy Fields	Week 4	2	16.7	8	33.3	10	27.8	0.336	0.091 to 1.234	0.100	0.134	<0.001
	Week 12	3	25.0	10	41.7	13	36.1					
	Week 24	4	33.3	1	4.2	5	13.9					
	Week 36	3	25.0	5	20.8	8	22.2					
TEA Telephone search whilst counting	Week 4	5	45.5	10	41.7	15	42.9	0.730	0.158 to 3.376	0.687	0.233	0.013
	Week 12	2	18.2	6	25.0	8	22.9					
	Week 24	0	0.0	5	20.8	5	14.3					
	Week 36	4	36.4	3	12.5	7	20.0					

*[1] LR Test for proportionality of odds p-value Approximate likelihood-ratio test of proportionality of odds across response categories

*[2]Brant Test of Parallel Regression Assumption

2.17.6 Secondary outcomes

The secondary outcomes seek to address wider aspects of neurocognitive function, seizures, quality of life, and daily life functioning in patients with tuberous sclerosis. The analyses exclude the two patients (PID3 and 63) who withdrew completely from the study. However, it includes one participant who had partial withdrawal (PID-27). This participant stopped the study medication after the week 12 visit but continued to have study assessments and is included on an “intention to treat” basis. The total participants analysed were therefore n = 36 with placebo = 12 vs everolimus = 24.

The majority of patients completed the Cambridge Neuropsychological Test Automated Battery (CANTAB) tests and tasks at baseline and 6 months (Table 18). The average 6 month scores were comparable by treatment, and placebo arms and a low rate of improvement was observed for each. A one-sided chi-squared test indicated that the proportion of patients in the everolimus group who improved at six months by at least one SD was significantly less than hypothesised, and no difference in the proportion showing improvement at 6 months was shown between arms (where tested).

A similar picture was seen in the Quality of Life in Epilepsy (QUOLIE) subscales and overall score (Table 19). No significant improvement was seen in the everolimus arm, and no difference between arms was seen.

Symptoms Checklist 90 (SCL-90R) and its sub-scales were not completed fully by most participants. The Global Severity Index and the Phobic Anxiety sub-scales had some of the lowest responses (Table 20). There no evidence of improvement by six months with no effect seen within the Everolimus cohort or between the two arms.

At 6 months, 8 out of 12 (67%) and 18 out of 24 (75%) patients (placebo and everolimus arms respectively), reported experiencing no seizures over the previous six months. The remaining 9 patients (4 and 5 respectively) completed the Liverpool Seizure Scale (LSS) and had a median (25^{th} to 75^{th} centile) score of 17 (14 to 23.75) and 13 (8 to 24.75), respectively. Only one of these patients observed an improvement in LSS of 1 SD over the 6 months. One participant experienced a deterioration in seizure frequency, leading to admission to hospital; details are provided in the SAE record. No participant needed a change in antiepilepsy medication.

Table 18 Secondary Outcomes – CANTAB

Results are reported as Mean (SD) unless otherwise stated

Test Task	Arm	N[1]	Baseline	6 months	No 1SD* improvement N(%)	1SD* improvement N(%)	1-sided chi-square test, p-value
CANTAB - Rapid Visual Information Processing (RVP)							
Standard score	Placebo	11	2.44 (1.02)	1.44 (1.01)	11 (100.0)	0 (0.0)	-
	Everolimus	22	1.57 (1.55)	1.45 (1.01)	21 (95.5)	1 (4.5)	18.18, <0.001
CANTAB – Spatial Span (SSP)							
Span length (Standard score)	Placebo	12	0.91 (0.56)	1.23 (0.84)	10 (83.3)	2 (16.7)	-
	Everolimus	23	0.75 (0.57)	0.85 (0.57)	21 (91.3)	2 (8.7)	15.70, <0.001
CANTAB - Intra-Extra Dimensional Set Shift (IDED)							
Number of stages completed (Standard score)	Placebo	12	0.93 (1.02)	0.90 (0.58)	10 (83.3)	2 (16.7)	-
	Everolimus	24	0.71 (0.31)	0.71 (0.31)	24 (100.0)	0 (0.0)	#
Total errors (Standard score)	Placebo	12	1.13 (0.77)	0.95 (0.65)	12 (100.0)	0 (0.0)	-
	Everolimus	24	0.87 (0.59)	0.97 (0.65)	22 (91.7)	2 (8.3)	#
Total number of errors adjusted (Standard score)	Placebo	12	0.85 (1.07)	0.85 (0.45)	11 (91.7)	1 (8.3)	-
	Everolimus	24	0.59 (0.24)	0.67 (0.21)	24 (100.0)	0 (0.0)	#

** Improvement in 1 Standard deviation on standard score> 1; ** adjusted for age (<50 years, ≥50 years), gender, IQ level (60-79, 80+), Antiepilepsy drugs (AED). # numbers too small for modelling*

[1] For all these analyses n=24 everolimus vs n=9 placebo will always be the numbers used , unless there are missing outcome data.

Table 19 Secondary Outcomes – Quality of Life in Epilepsy (QOLIE-31)

Results are reported as Median (25th to 75th centiles) unless otherwise stated

T-scores	Arm	N[1]	Baseline	6 months	No 1SD* improvement N(%)	1SD* improvement N(%)	1-sided chi-square test, p-value
Seizure Worry	Placebo	11	61 (43 to 66)	58 (41 to 66)	10 (90.9)	1 (9.1)	-
	Everolimus	22	60 (44.25 to 60)	63.50 (46 to 66)	20 (90.9)	2 (9.1)	14.73, <0.001
Overall Quality of Life	Placebo	11	50 (46 to 59)	46 (43 to 59)	10 (90.9)	1 (9.1)	-
	Everolimus	22	47 (38 to 54)	44.5 (37 to 56)	21 (95.5)	1 (4.5)	18.18, <0.001
Emotional Well-Being	Placebo	11	57 (38 to 64)	49 (40 to 54)	9 (81.8)	2 (18.2)	-
	Everolimus	22	47.5 (37.5 to 57)	50 (39.5 to 57)	20 (90.9)	2 (9.1)	14.73, <0.001
Energy/Fatigue	Placebo	11	55 (50 to 59)	55 (47 to 62)	10 (90.9)	1 (9.1)	-
	Everolimus	22	48.5 (43 to 55)	48.5 (41.75 to 57)	20 (90.9)	2 (9.1)	14.73, <0.001
Cognitive Functioning	Placebo	11	56 (43 to 61)	54 (46 to 59)	9 (81.8)	2 (18.2)	-
	Everolimus	22	51 (42.75 to 58.5)	55.5 (45.75 to 60)	18 (81.8)	4 (18.2)	8.91, 0.003
Medication Effects	Placebo	11	61 (46 to 64)	61 (51 to 64)	10 (90.9)	1 (9.1)	-
	Everolimus	22	55.5 (49 to 64)	58.5 (53 to 64)	19 (86.4)	3 (13.6)	11.64, 0.001
Social Functioning	Placebo	11	54 (48 to 59)	57 (55 to 62)	10 (90.9)	1 (9.1)	-
	Everolimus	22	52.5 (42 to 57)	56 (46.75 to 60)	18 (81.8)	4 (18.2)	8.91, 0.003
Overall Score	Placebo	11	59 (48 to 61)	55 (51 to 60)	11 (100.0)	0 (0.0)	-
	Everolimus	22	52 (44.5 to 55.25)	53 (45 to 58.25)	19 (86.4)	3 (13.6)	11.64, 0.001

** Improvement in 1 Standard deviation on T score> 10; **adjusted for age (<50 years, ≥50 years), gender, IQ level (60-79, 80+), Anti-epilepsy drugs (AED). # numbers too small for modelling*

[1] For all these analyses n=24 everolimus vs n=9 placebo will always be the numbers used , unless there are missing outcome data.

Table 20 Dimensions and Global Scores on the Symptoms Checklist 90 (SCL-90R)

Results are reported as Median (25th to 75th centiles) unless otherwise stated

	Arm	N[1]	Baseline	6 months	No 1SD* improvement N(%)	1SD* Improvement N(%)	1-sided chi-square test
Test scores							
Somatisation	Placebo	12	46 (37 to 67.5)	54 (37 to 65.25)	10 (83.3)	2 (16.7)	
	Everolimus	22	59 (49 to 68)	52 (40 to 60.5)	19 (86.4)	3 (13.6)	11.64, 0.001
Obsessive-Compulsive	Placebo	11	61 (49 to 70)	52 (45 to 58)	10 (90.9)	1 (9.1)	
	Everolimus	18	60.5 (49.75 to 69)	52.5 (46.5 to 56.5)	17 (94.4)	1 (5.6)	14.22, <0.001
Interpersonal Sensitivity	Placebo	10	62 (59.75 to 68.25)	50 (45.75 to 62.25)	10 (100.0)	0 (0.0)	
	Everolimus	19	56 (41 to 62)	56 (41 to 61)	17 (89.5)	2 (10.5)	11.84, 0.001
Depression	Placebo	11	57 (46 to 65)	52 (42 to 63)	10 (90.9)	1 (9.1)	
	Everolimus	20	56.5 (50.75 to 70.50)	56.5 (47 to 63.5)	19 (95.0)	1 (5.0)	16.20, <0.001
Anxiety	Placebo	11	51 ()37 to 63)	49 (40 to 60)	8 (72.7)	3 (27.3)	
	Everolimus	20	57 (41.25 to 64)	52 (40 to 58.5)	18 (90.0)	2 (10.0)	12.80, <0.001
Hostility	Placebo	11	54 (41 to 62)	49 (40 to 58)	10 (90.9)	1 (9.1)	
	Everolimus	20	57.5 (40 to 66.5)	56.5 (40.75 to 64.25)	19 (86.4)	3 (13.6)	11.64, 0.001

	Arm	N[1]	Baseline	6 months	No 1SD* improvement N(%)	1SD* Improvement N(%)	1-sided chi-square test
Phobic Anxiety	Placebo	9	55 (44 to 65.5)	55 (44 to 61.5)	8 (88.9)	1 (11.1)	
	Everolimus	18	60.50 (47 to 63.25)	48 (44 to 60.25)	16 (88.9)	2 (11.1)	10.89, 0.001
Paranoid Ideation	Placebo	11	54 (49 to 64)	41 (41 to 62)	11 (100.0)	0 (0.0)	
	Everolimus	22	58.5 (41 to 67.25)	51 (41 to 64.50)	21 (95.5)	1 (4.5)	18.18, ,0.001
Psychoticism	Placebo	10	47.5 (44 to 63)	44 (44 to 65.25)	9 (90.0)	1 (10.0)	
	Everolimus	21	59 (44 to 67)	54 (44 to 62.5)	20 (95.2)	1 (4.8)	17.19, <0.001
Global Severity Index (GSI)	Placebo	8	65 (57 to 69)	52.5 (47.75 to 65.25)	8 (100.0)	0 (0.0)	
	Everolimus	17	62 (56 to 73)	58 (52.5 to 63.5)	16 (94.1)	1 (5.9)	13.24, <0.001
Positive Symptom Distress Index (PSDI)	Placebo	12	63.5 (57.5 to 69.5)	57.5 (42.25 to 69.5)	12 (100.0)	0 (0.0)	
	Everolimus	18	57 (50 to 64.75)	51 (41.5 to 62.75)	18 (100.0)	0 (0.0)	#
Positive Symptom Total (PST)	Placebo	9	58 (53 to 66.5)	52 (48 to 66)	9 (100.0)	0 (0.0)	
	Everolimus	19	63 (55 to 68)	60 (53 to 65)	19 (100.0)	0 (0.0)	#

** Improvement in 1 Standard deviation on T score> 10; **adjusted for age (<50 years, ≥50 years), gender, IQ level (60-79, 80+), Antiepilepsy drugs (AED)*
 # numbers too small for modelling.[1] For all these analyses n=24 everolimus vs n=9 placebo will always be the numbers used , unless there are missing outcome data

2.17.7 Safety aspects of the TRON study

Adverse events (AE) are defined as the appearance or worsening of any undesirable sign, symptom, or medical condition occurring after starting the study drug or placebo, even if the event is not considered related to the study drug. All AEs were collected, recorded and reported according to GCP guidelines and the requirements of the Medicines for Human Use (Clinical Trials) Regulations 2004. The severity was graded according to the National cancer institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.03(NCI 2010).

Once detected, AEs were treated appropriately and monitored until resolution unless the Investigator deemed this unnecessary and documented accordingly. Treatment could include interruption, modification or discontinuation of study drug and changes in the frequency or nature of assessments, hospitalisation, or any other medically required intervention described in section 1.16.7.1.

2.17.7.1 AE details

A list of the most common AEs reported by the study participants is depicted in table 21. An AE reported by participants on multiple occasions is counted only once in that adverse event category in table 21 (normalised data) in order to present the data in a form comparable to previous studies of everolimus treatment in TSC patients (Bissler et al. 2013). All recruited participants (n=38) were included in this table (placebo = 12 and everolimus = 26).

The most common AEs reported were mouth ulcers and infections, consistent with everolimus' known safety profile in patients with TSC. Known common side effects of the IMP are listed in the Summary of Product Characteristics document (SmPC), which was the Reference Safety Information (RSI) for the trial (Novartis 2019). AEs, where nature and severity were consistent with the information set out in the SmPC for everolimus, were considered to be "expected".

Thirty study participants reported a total of 129 AEs. AEs were Grade 1 or 2 (less severe) in most (125/129 episodes). The study drug was withheld and then restarted for 42 AE episodes in the everolimus group and 9 episodes in the placebo group. No episode of AE led to permanent discontinuation of the study drug. Participants in the everolimus group reported a total of 33 episodes of (Grade 1 or 2) stomatitis (mouth ulcers), while there was one report of mouth ulcers in the placebo group.

The normalised data indicated that infections occurred in 75% (18/24) of participants in the everolimus group and 41.6% (5/12) of the placebo group. Most of the infections reported were upper respiratory tract infections. However, there were two reports of otitis externa (one participant) and one episode each of

urinary tract infection, cellulitis, tooth infection and paronychia needing treatment in the everolimus group. Both groups had one participant treated for chest infection. All these AE were grade 1 or 2. Skin rash or fungal infections were reported in 8 participants in the everolimus group. In contrast, one episode was reported in the placebo group. Four episodes of gastroenteritis were reported in the everolimus group as compared to one episode in the placebo group.

There was no reported grade 4 AE. There were three grade 3 AE in the everolimus group with one episode each of a seizure, low neutrophil count and low serum phosphate levels. One grade 3 episode of seizure was reported in the placebo group.

In other frequent AEs, musculoskeletal pain was reported on 4 occasions in the everolimus group only. Lab AEs comprised a high everolimus level (one), low phosphate levels (two), high CPK (one) and proteinuria (two) in the everolimus group. In contrast, there was no lab AE report in the placebo group. Other AE reported in the everolimus group were one episode of intentional weight loss (>10% body weight) and an altered sense of taste.

The Fisher Exact test was used to assess whether the incidences of AEs were different in the everolimus and placebo groups. Only stomatitis/mouth ulcers (P = 0.0116) and “other” AEs (P = 0.003) reached significance at the 0.05 level, both being more frequent in the everolimus group. After Bonferroni correction for the multiple AE comparisons, the differences were non-significant.

Table 21 List of AE reported by study participants(normalised)

	Placebo (n=12)		Everolimus(n=26)	
Participants reporting AE	n = 6		n = 24	
	Grade1/2 (total=15)	Grade3/4 (total=1)	Grade1/2 (total =77)	Grade3/4 (total=3)
Stomatitis/ mouth ulcers	1		14	
Infections[1]	5		18	
Skin rash/infections	1		8	
Diarrhoea/gastroenteritis	1		4	
Suicidal Ideation	1		3	
Seizures	0	1	2	1
Headache	2		2	
Toothache	0		2	
Neutropenia	0		1	1
Agitation	1		1	
Other AE	4		22	1

[1] Includes any reported flu, upper respiratory tract infection, laryngitis, urinary tract infection, otitis externa, tooth infection and paronychia.

2.17.7.2 Serious Adverse Events (SAE)

Five Serious adverse events (SAE) were recorded during the TRON study, as summarised in table 22. One participant (placebo group) reported a worsening of seizures leading to hospital admission. The reason for the deterioration of the participant's epilepsy could not be ascertained, as all the assessments were normal. Antiepilepsy management remained unchanged. The study drug was withheld during the hospital admission and restarted after a fortnight. At the time of SAE, the study drug was considered 'unlikely' to be the cause of the SAE.

Table 22 SAE details for TRON study participants

Main Diagnosis	Study Group	Everolimus Dose	Study Drug restarted
Seizure	Placebo	N.A.	Yes
Suicidal ideation	Placebo	N.A	Yes
Suicidal Ideation	Everolimus	10mg	Yes
Suicidal Ideation	Everolimus	(after the 24-week visit)	-
Suicidal ideation	Everolimus	(after the 24-week visit)	-

Suicidal ideation was reported in four cases during the trial. These events were considered as SAE due to the seriousness of potential outcomes. However, there was no indication of suicidality, planning or intent in these episodes. One of the affected participants was in the placebo group, while the other three were in the everolimus group. Interestingly, two of these participants had the SAE *after* they had stopped the study drug (after the 24week visit). The reported episodes appeared related to external factors in participants' lives. Safety netting was provided by involvement of local crisis teams and/or GPs at the time. In keeping with the TRON protocol, for those taking trial medication, this was withheld until the SAE resolved

2.17.7.3 Drug Levels of Everolimus

Table 23 illustrates the median everolimus levels after 6 months, the median everolimus trough levels were 7.5 ng/mL (25^{th} to 75^{th} centiles = 5 to 10ng/mL). Nine patients did not change dose levels over the six months, while ten changed once and five changed twice.

Table 23 Trough everolimus blood levels of TRON study participants

	Weeks/months of treatment				
Everolimus levels (ng/mL)	2 weeks	4 weeks	6 weeks	3 months	6 months
Mean	6.7	7.0	7.3	8.2	7.9
SD	2.4	2.4	3.0	3.8	4.0
Median	5	5	5	7.5	7.5
25^{th} to 75^{th} centiles	5 to 10	5 to 10	5 to 10	5 to 10	5 to 10

2.17.8 Discussion

Memory and executive function impairments are part of TSC associated neuropsychiatric disorders (TAND), an umbrella term for a range of neuropsychiatric manifestations of tuberous sclerosis (figure 31). The pattern of TAND issues varies significantly among patients with TSC. They are prevalent even in individuals with TSC who have normal intellectual abilities, particularly in the neuropsychological domain (Prather and de Vries 2004). Despite their prevalence and importance, TAND issues are rarely assessed and treated (Curatolo et al. 2015). The TAND checklist was developed to improve detection of these problems in the clinic. (de Vries et al. 2015)

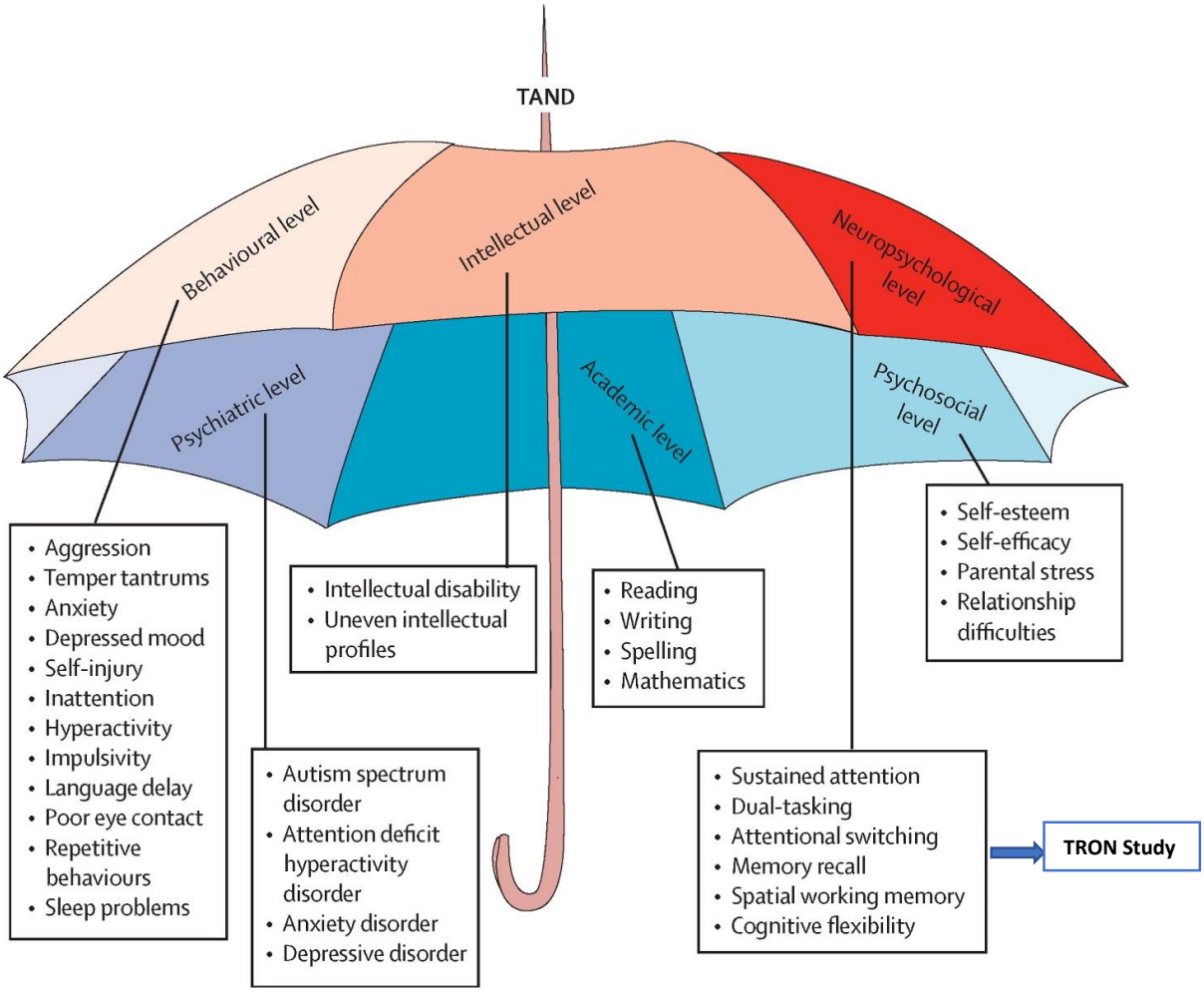


FIGURE 31 TAND is used as an Umbrella term to capture a range of Neuropsychiatric presentations associated with TSC (Curatolo et al. 2015)

The TRON study, investigated a subgroup of TSC adults with infrequent seizures, who were able to complete detailed neurocognitive assessments. The TRON study evaluated changes in the neuropsychological domain of TAND in participants treated with everolimus or placebo for six months.

2.17.8.1 Trial design

The trial was designed based upon the available evidence from preclinical and clinical studies investigating neurocognitive deficits in TSC. When planning the TRON study, there was evidence of reversal of spatial learning deficits by sirolimus (mTOR inhibitor) treatment in heterozygous *Tsc2* mutant mice (Ehninger et al. 2008a). The TESSTAL study of TSC patients taking sirolimus therapy for renal AMLs also provided anecdotal evidence of improvement in recall memory (Davies et al. 2011b).

2.17.8.2 Revision of eligibility criteria

Initially, the study protocol included as an eligibility criterion a ≥ 2SD deficit in the neuropsychological test scores of any one or more of the ten primary outcome measures. It became evident early in the trial that potential participants with deficits of ≥ 2SD had limited ability to function independently, leading to a failure to complete the assessments. Consistent with this observation, it also became clear that very few participants with IQ>60 had deficits of this magnitude, creating a recruitment problem. Additionally, this eligibility criterion was judged to exclude a large number of individuals who might benefit were the treatment found to be effective. For these reasons, following agreement from the data monitoring and steering committees, the eligibility criterion was relaxed to allow participation by patients with a 1.5 SD or greater deficit in any one or more of the primary neurocognitive outcome measures.

2.17.8.3 Choice of outcome criteria

An acceptable criterion for "meaningful change or improvement" in TAND aspects of TSC is yet to be defined, in contrast to solid tumours such as renal AMLs (Bissler et al. 2013) or manifestations such as epilepsy (French et al. 2016).

At the time of trial design, it was uncertain to what extent repeated neuropsychological assessments might result in a learning effect in the selected group of trial participants. The estimate of 15% improvement due to learning effects was based on similar assessments in other studies and the previous experience of the (child and adolescent) psychiatrist and psychologist in the trial team who had previous clinical and research experience with the assessment tools being used.

The study design assumed that learning or practice effects were expected to occur in only a minority of participants. However, the *learning effect* of repeated assessment on neuropsychological test scores was significantly underestimated: 15% of the placebo arm were anticipated to show an improvement in one or more primary outcome measure, 44.4% were observed to do so. A much larger improvement in the everolimus arm would be needed to provide convincing evidence of meaningful change. An effect size of 35% in the everolimus arm, i.e. 20% above the expected learning effect size, was agreed as an appropriate

threshold and the power calculations for study participant numbers were based on these assumptions. However, the values selected proved to be significant underestimates of the observed proportions of participants showing improvement in outcome measures during the trial.

2.17.8.4 Possible explanations of the learning effects

The marked learning effect observed may reflect the TAND profile of TSC individuals, specifically the frequency of coexisting anxiety and autistic traits. Most participants appeared anxious during their early visits, including initial screening and baseline assessments, perhaps reflecting the challenges of travel, novel interaction with the trials team and an unfamiliar hospital environment, as well as the stress of the unfamiliar psychological tests. Participants' anxieties appeared to resolve over time as they got used to the trial team, the environment and the assessments. Anxiety may have led to lower initial assessments scores that were not reflective of their real abilities. However, against this hypothesis, the scores for anxiety (symptom checklist: SCL-90) were not significantly different between the early and late visits.

2.17.8.5 Statistical approach

The statistical analysis plan was to determine the frequency of a one SD improvement in *any* one or more of the ten neurocognitive tests in the primary outcome analysis across the groups of participants in the placebo and intervention arms. An alternative approach could have been to evaluate the longitudinal changes in each individual's neurocognitive measures with time. This approach might provide insight into individual performance changes rather than assessing the frequency of change across a group. The individual approach may also be more clinically relevant than the frequentist inference applied in the TRON study.

The extent of the learning effects observed in the trial suggest that the chosen threshold of improvement of 1 SD used to identify a responder was too low. The value was selected as it was considered to equate to a clinically significant improvement. Alternatively, a higher threshold or improvement in a larger number of measures could be considered when defining a responder.

2.17.8.6 Reliability of test-retest in neuropsychological assessments

Serial assessments are common in neuropsychological practice and provide additive value in numerous clinical and forensic settings. Practice effects seen on serial assessments could be related to age, disease process or regression to the mean (Heilbronner et al. 2010) (Beglinger et al. 2005). The use of alternate forms for a neuropsychological test may stabilise the practice effects seen with serial neuropsychological testing. The placebo group data in a clinical trial enables separation of practice effects from drug signal changes (Beglinger et al. 2005). The practice of using selected subtests can be justified by reference to

research literature employing these measures in order to be efficient and respectful to the patient's time and resources (AACN 2007; Baron 2018).

The reliability of neuropsychological assessments over multiple timepoints in TSC patients has not been tested robustly in previous studies. Instead, previous studies sought to determine the prevalence of neuropsychological problems in TSC and reported measurements at a single time point (Jambaqué et al. 1991; Harrison et al. 1999a; de Vries et al. 2009; Toldo et al. 2019). The variability of the assessment scores we observed could reflect temporal variation in a test-retest setting in the TSC population, independent of any intervention or changes in external factors (such as anxiety) that may impact performance.

The assessor, room environment, hotel, and travel arrangements for participants were unchanged across their study visits to minimise external variables. However, unexpected changes (room availability, overnight disruption in the hotel or problems with a taxi) nevertheless occurred, potentially leading to an unpredictable impact on the test performance.

It is also possible that a trial participant’s neuropsychological assessment scores could vary on a day-to-day or hour to hour basis, depending on external factors such as mood, sleep, hunger, social, and work circumstances. It is possible that such factors impact performance of individuals in the TSC population to a greater extent than in the general population. A larger study population would help to overcome the influence of variation in performance relating to factors other than the trial medication.

2.17.8.7 The TEA-subtest issue

A TEA subtest, *Telephone Search whilst Counting* –TSwC, was one of the TRON study's ten primary outcome measures. Seven participants were eligible to participate in the TRON trial based exclusively on the TSwC, as they did not have a deficit of 1.5SD or more in any other primary outcome measure at screening. However, the scores from the TSwC varied greatly over time, even in the placebo group. This result prompted further consideration of the results from this assessment and it became evident that the two versions of TSwC tests used gave test scores that were not comparable.

The TEA test has poor test-retest reliability between its different versions. Estimates of concordance coefficients based on data of the standardisation sample (n=118) with a 1-week interval between administration of Versions A and B of the TEA subtests ranged from .59 (TSwC) to .86 (MS and TS). Salvia and Ysseldyke (1998) recommend that reliability coefficients should be at least .90 for individual diagnostic

purposes (Salvia 1998). The TEA test does perform better with construct validity based on combined subtests. Studies have shown that it has clinical utility in discriminating clinical populations from control groups. This has been reported in multiple cohorts of patients with traumatic brain injury (Chan 2000; Bate et al. 2001), schizophrenia (Oram et al. 2005) and stroke (Chen et al. 2013). However, it is likely that in general, the reliability of the TEA subtests on their own are suspect (Stinnett 2007), which became evident in the TRON trial.

The exclusion of 7 participants' data from the primary analysis reduced the statistical power of the study and lowered the confidence in its results.

2.17.8.8 Age group

The TRON study assessed longitudinal change in neuropsychological test scores in adults. This was based on the reported recovery of memory deficits in adult TSC transgenic mouse studies using rapamycin (Ehninger et al. 2008a) and a previous clinical trial in TSC, where the TESSTAL trial used similar assessment tools but for secondary rather than primary outcome measures (Davies et al. 2011b).

The adult human brain remains amenable to change as demonstrated by various experimental paradigms such as brain plasticity (structural) changes with musical training, juggling, learning golf, learning a second language and exercise (Herholz and Zatorre 2012) (Cotman and Berchtold 2002; Draganski et al. 2004; Bezzola et al. 2011; Stein et al. 2012). These studies have been performed in healthy controls who would not have the developmental brain structural abnormalities that are seen in TSC. The TRON study participants' results showed an improvement in some brain function scores in both the placebo and everolimus group. However, whether treatment effects might be more prominent in children remains unclear.

2.17.8.9 Recruitment

The TRON study was initially planned to recruit over three years, but this was extended to 6 years due to slow recruitment. The motivation of potential participants to volunteer for the study may have been lower than for clinical trials in TSC for different TSC-associated problems such as epilepsy, renal AMLs and SEGAs. Difficulties with recruitment has been reported in other clinical trials of mTOR inhibitor therapy for neurocognitive functioning in TSC from Boston (Krueger et al. 2017) and Rotterdam (Overwater et al. 2019), although these other trials were conducted in children and young adults where the motivation of parents is likely to have been more important than that of the participants themselves.

Low motivation to participate among the target trial population of relatively mildly affected adults living in the community could be due to acceptance of longstanding neurocognitive difficulties by this patient group. Memory and executive skills problems may be more of a concern to TSC individuals' families and employers than to the potential participants themselves. Participation in the TRON trial involved multiple visits to one of the clinical trial centres, which meant the disruption of routine, travelling and finding one's way to and around a new place, all of which are challenging if there are underlying planning issues symptomatic of memory and executive function problems, or issues with autism or autistic spectrum disorder that are common in TSC.

It is apparent that the profile of participants that the TRON study was trying to recruit was likely to make recruitment challenging. Recognising this, logistical support was provided to all potential participants by the trial team who undertook arranging, booking and paying for all the return travel and hotel requirements form home to the study centres for participants and an accompanying family member or carer, and reimbursing incidental costs. The research clinician (AS) was also available by phone 24 hours a day, or by email, to discuss with potential participants or their families/carers any concerns about participation.

The study team tried to mitigate these recruitment issues by presenting the TRON trial to clinical teams in TSC clinics, to the UK clinical genetics community, to the UK NIHR Genetics Clinical Research Network and to patient and family group forums with the help of the Tuberous Sclerosis Association. An explanatory video including a participants experience was recorded and made available on YouTube (https://www.youtube.com/watch?v=7-X1VfYCCQI).

2.17.8.10 Choice of the everolimus dose and drug levels

At the time of the TRON study planning, evidence emerged that tumours related to TSC respond to lower doses of mTOR inhibitors than are typically used for immunosuppression in transplant patients or for the treatment of sporadic cancers. We were also concerned that the risk-benefit balance for participants should minimize risks of significant everolimus-related adverse events. A starting dose of 5 mg was selected for the TRON study compared to the standard adult starting dose of 10 mg per day for non-TSC indications. Since then, the result of Exist-3 study has shown that in TSC-associated epilepsy (another manifestation of perturbed brain function) a higher dose of everolimus is associated with a higher response rate (French et al. 2016). The best response in EXIST-3 study was noted in the group with an everolimus target level of 8-15 ng/ml as compared to a target everolimus level of 3-10ng/ml in the TRON study.

The differential outcomes of dose responses in tumours and epilepsy in TSC patients might be explained by different underlying genetic changes in these different manifestations of TSC. The tumour manifestations of TSC (renal AMLs and SEGA) have biallelic mutations of TSC genes due to somatic loss of heterozygosity or second point mutations in the tumour cells (Caban et al. 2016). Such null cells appear to be very sensitive to mTOR inhibition and hence the tumurs are likely to respond to low doses of everolimus. In contrast, the vast majority of CNS neurons, and even many of the cells in cortical tubers, are heterozygous for the inherited *TSC1* or *TSC2* mutation but have an intact second allele. Structural and functional changes in heterozygous neurons are well documented and may contribute to the TAND manifestations of TSC in mice (Tavazoie et al. 2005). Heterozygous TSC+/- are less sensitive than TSC null cells to mTOR inhibitors and higher drug concentrations may be required to significantly impact mTOR signalling. The dose response seen in the Exist-3 trial of everolimus treatment for refractory epilepsy in TSC would be consistent with this hypothesis. Furthermore, there is uncertainty about how, in humans, CNS levels of everolimus relate to blood levels.

2.17.8.11 Strengths and weaknesses of the TRON study

Despite the recruitment problems described earlier, the TRON study had good retention of participants. Of 38 recruits, two withdrew early in the study period, while another withdrew mid-way due to the disruption caused by study visits on the participant's education. The low attrition after the earliest visits may reflect participants' reduced anxiety after getting used to the clinical trial team, visit schedule, and travel to the study centres. The additional perceived benefit of being reviewed regularly by a TSC specialist clinical team could also be contributory to the better than expected retention of participants.

The TRON trial also has good data capture for the primary outcome measures; however, a significant amount of missing data was apparent for the secondary outcome measures. This is likely because primary outcome measures were captured by face-to-face assessment with the participants at study visits. In contrast, many tests for secondary outcomes were questionnaires and forms that were filled by a parent, spouse or carer at home (e.g., Vineland tests). This observation may help design of future trials.

Some aspects of trial design were a weakness in the TRON study. Its exploratory nature and the desire to capture as much potentially useful data as possible led to a design that included multiple primary outcome measures in the neurocognitive domains of memory and executive function. However, an improvement of 1SD in any one of 10 different test scores was used to define a responder. Therefore, a responder might improve by this extent in just one measure while deteriorating in others. Clearly, such an outcome would be unlikely to represent genuine clinical benefit. This weakness reflects the ambition for the trial to identify any

biological response to mTOR inhibition in order to establish in principle the potential for this approach in treating neurocognitive problems in TSC. Any future trials should consider ways in which composite scores might be derived to better reflect overall benefits or harms of interventions.

In the selection of outcome measures, the issues associated with use of different forms of the TEA test should have been understood at the trial design stage. At a minimum, the same versions of the test should have been administered at baseline and 6 months to enable like-for-like comparison. Alternatively, a different measure should have been selected. Although the study protocol was developed by a team including a research psychologist with extensive experience of TSC and an academic child and adolescent psychiatrist with high level experience in TSC and subject to external review by Novartis pharmaceuticals and the Tuberous Sclerosis Association prior to being funded, wider discussion with research psychologists would likely have improved the study design.

2.17.8.12 Comparison with previous studies

A clinical trial evaluating neurocognition in children and young adults with TSC was ongoing at the time of initial approval of the TRON study (NCT01289912) and has since reported its findings (Krueger et al. 2017). The randomised controlled trial investigated 47 participants with TSC between 6 and 21 years of age treated with everolimus or placebo for six months. It was an exploratory study investigating the change in scores from baseline for multiple cognitive and behavioural domains. The study authors did not find significant differences between the placebo and everolimus arms. Although the TRON study evaluated similar outcomes in adults, a significant difference was that participants in the TRON study were only eligible to take part if they had a deficit of 1.5 SD or more in at least one of the primary neurocognitive measures, while participants in the study undertaken by Krueger et al., did not have to have any neurocognitive deficit to be eligible. Apart from the question of risks and benefits for participants, it is questionable whether treatment could be expected to improve performance that was already normal.

A second trial of mTOR inhibition for TAND problems was undertaken in the Netherlands in children aged 4-17 years with TSC who had an IQ <80 or learning disability or special schooling or autism (Overwater et al. 2019). Exclusion criteria comprised developmental age <3.5years, >1 seizure per week, severe liver or kidney dysfunction, or prior treatment with an mTOR inhibitor. The study evaluated change in IQ as its primary outcome variable, with secondary outcomes being markers of autism, neuropsychological functioning, and behavioral problems. The participants received everolimus or placebo for 12 months. The study attempted to recruit 60 children. However, recruitment was stopped prematurely, with 32 children randomised due to slow recruitment. The study team concluded that everolimus treatment in the children in the study resulted in no significant improvement in full-scale IQ or secondary outcome measures, including

performance IQ, verbal IQ, and scores achieved in ADOS and CANTAB tests.

The negative results may reflect the study period of 12 months being a short interval to assess a composite measure of intellectual ability such as IQ, as the study team acknowledged. Additionally, the study was powered to its original target of 60 recruits and, having only recruited some 50% of the target, could have failed to detect a signal if one was present, although the study team reported that the reduced participant numbers should not have affected the results.

The TRON study did identify improvements in neurocognitive function in participants, using the analytical approach set out in its statistical analysis plan. However, improvements were seen in both the placebo and everolimus groups. After excluding participants eligible only on the TEA test, the numbers of participants in each arm were below the numbers used in the trial power calculations and the differences between the placebo and everolimus groups were small. TRON study is the first randomised controlled trial to investigate neurocognitive changes in TSC adults receiving everolimus treatment. The selection of some outcome measures and a large learning effect in the placebo arm have compromised the trial. The primary aim to determine effect sizes in the placebo and everolimus arm was achieved, and insights have been gained which would be useful in planning any future and larger trial to investigate everolimus in the treatment of neurocognitive problems in TSC.

2.17.9 Licensing of everolimus (mTORi) in other TSC complications

While the TRON study was ongoing, everolimus has been granted licenses for treatment of SEGA, renal AMLs, and epilepsy associated with TSC, based on phase 3 clinical trials demonstrating efficacy for these indications (Bissler et al. 2013; Franz et al. 2013; French et al. 2016).

Everolimus gained FDA approval on 26th April 2012 to treat large renal AMLs, with EMA approval on 20th September 2012. This decision was based on the EXIST-2 study results, which found that 42% of patients on everolimus experienced a reduction AMLs size versus 0% of patients in the placebo arm (Bissler et al. 2013) (Bissler et al. 2017).

EMA also approved everolimus on 15th December 2016 as an adjunctive treatment of refractory seizures associated with tuberous sclerosis complex (aged 2 years and older), while FDA approval was granted on 10th April 2018, again based on the EXIST-3 trial results (French et al. 2016). Additionally, another mTOR inhibitor, sirolimus, has been licensed for the treatment of pulmonary LAM (McCormack et al. 2011), while topical sirolimus has demonstrated efficacy in managing facial AFBs in a phase 3 RCT (Koenig et al. 2018).

The more extensive trials that led to treatment licenses show that mTOR inhibition is efficacious for treating many important manifestations of TSC. However, larger trials for TAND manifestations of TSC have not yet been undertaken. It is unlikely that Novartis will pursue everolimus as a potential treatment for TAND manifestations of TSC because of patent expiry. Sirolimus' patent held by Pfizer had already expired prior to the TRON trial being initiated. Therefore, further investigations with these mTOR inhibitors in TSC would likely require public sector funding which will be challenging to justify in the absence of a commercial partner to share the considerable costs involved in larger scale trials.

2.17.10 Conclusions and Future directions

The TRON trial was a phase 2 study that identified a modest positive signal for improvement of some aspects of neurocognition in adults with TSC treated with everolimus for 6 months. The trial did not identify any major or unexpected safety issues. The trial adds to the existing and more persuasive evidence of a biological effect provided by preclinical studies and clinical trials that have found everolimus an efficacious management option for other TSC manifestations.

The findings from TRON and the two other reported studies investigating mTOR inhibitors for the treatment of TAND issues provide insights that would help in the design of further clinical studies to evaluate mTOR treatment for the TAND issues that impact important aspects of TSC individuals' everyday lives, including education, employment, and long-term relationships. However, the three small trials that have been undertaken so far have not produced the promising evidence that was seen in phase 2 trials for SEGA, AML and seizures. The complexity of neurocognitive phenotypes and the difficulty in measuring them present a formidable challenge in this regard.

Future trials should consider in particular the participants' needs, drug dosage to be used, and selection of the primary outcome measure. A multicenter trial based around established TSC clinic teams would be helpful in recruitment. While phase III trials for AML, SEGA and epilepsy have been international trials, neurocognitive measures are often only validated in one or a few languages and much effort would need to be put into the design of an international trial in this area.

*Participants' needs: A busy clinical research facility that focu*ses on cancer trials may not be the best place for a study visit for an anxious participant with TSC. A strategy of providing door to door travel bookings as provided in the TRON study should be adopted.

The assessments should be standardised, easily administered (minimally variable), and clinically meaningful. There are no agreed criteria to define a clinically significant improvement for use in clinical trials with neurocognitive outcomes, in contrast to epilepsy and tumour trials. A reduction of seizure

frequency by half (50%) is accepted as a clinically significant outcome and criteria such as the RECIST criteria are widely accepted for tumour trials. A pragmatic outcome measure profile, which satisfies the clinical neuropsychology community's scientific rigour and whose change can be assessed in a clinical trial, should be developed for further studies. The statistical plan should also consider evaluating the changes in individual participant scores rather than group responses.

Recruitment will always be challenging when studying a subgroup of patients with a rare disease. The difficulty of drawing significance from neuropsychological studies with small numbers of participants has been known for some time (Tversky and Kahneman 1971). These difficulties could be managed by aggregation methods based on shared data elements, i.e., merging data from multiple studies, linking data to previous studies, and collaboration between study teams. We should also consider if the traditional (frequentist framework) statistical models and the quest for a significant p-value is the right model for research outcomes in rare diseases. Novel trial methodologies, such as Bayesian methods, could be used for small sample studies (McNeish 2016). However, it currently remains unclear if any significant result in a clinical trial based on Bayesian analysis would be acceptable to regulatory authorities when applying for human use licensing agreements.

3 TRON Diffusion Tensor Imaging (DTI) study

3.1 The rationale of the study

The TRON study was planned to evaluate aspects of neurocognition (part of TAND spectrum) by neuropsychological tests; it also presented an opportunity to assess the effect of mTOR inhibition on white matter tracts for TRON participants.

DTI imaging offers a modality for assessing the white matter tracts integrity in TSC. These have abnormal DTI indices, despite appearing normal on standard MRI scans (Makki et al. 2007b). DTI indices have been used as a novel marker for autism, which is part of the TAND spectrum of TSC manifestation (Peters et al. 2012). Abnormal DTI indices have been correlated with neurocognitive deficits in TSC, with some studies reporting improvement in the DTI measures with everolimus treatment (Tillema et al. 2012); however, these studies have not evaluated longitudinal changes in DTI indices in relation to neurocognitive aspects of the TSC phenotype.

3.2 Aims and objectives of the DTI study

Neuroimaging with DTI was used to investigate alterations in white matter DTI indices in individuals with TSC treated with a mammalian target of rapamycin (mTOR) inhibitor everolimus over 6 months in TRON study.

3.2.1 Study objective

This study's primary objective was to compare the change in DTI indices in TRON study participants, after treatment with Everolimus or placebo for six months.

3.3 Timeline, approval and Study sites

The TRON imaging study received ethical approval by a substantial amendment to the TRON study protocol ver 6.0 on 2/2/13. The first TRON imaging study participant was scanned on 21/5/2013 (figure 6), while the last participant scan was performed on 1/6/2018.

3.3.1 DTI scan sites

The initial 10 participants pre- and post-treatment (baseline and 6 month) scans were acquired in CUBRIC 1 site based at Park Place, Cardiff. As CUBRIC relocated to a new imaging centre based at Maindy Road, Cardiff, after CUBRIC 1 closed, the next 12 participants had their (paired) scans performed at the new scanner based at CUBRIC-2. The imaging participants home location stratified to the different scanners is illustrated in figure 32. Four participants had their study centre in Glasgow or Belfast and travelled to CUBRIC for their scans. The participant's home locations are shown in figure 32.

FIGURE 32 TRON imaging Participant locations

Orange Stars- participant location for CUBRIC 1 scan cohort (n=10)

Blue Pins- participant location for CUBRIC 2 scan cohort (n=12)

Purple Star- CUBRIC scanner location

3.4 Inclusion and exclusion criteria

3.4.1 Inclusion criteria

All participants eligible for the TRON study were potentially eligible to take part in the imaging study.

3.4.2 Exclusion criteria

- Ferromagnetic implants, other than those approved for use in MRI scanner.
- Uncontrollable claustrophobia.
- Inability to fit in scanner due to size.

3.5 Screening for eligibility

Information regarding the imaging study was provided to all potential participants attending the screening visit for the TRON trial. If the participants agreed to participate in the imaging study, fully informed written consent was obtained.

3.6 Study design

Recruitment to the TRON imaging study was based on recruitment to the TRON study. We had estimated the statistical power based on obtaining imaging data on 40 TRON study participants (2:1 treatment to control ratio). Allowing for a 20% drop out from original TRON recruitment, we had anticipated obtaining imaging data on 32 participants.

TRON study recruitment was delayed as described in chapter 2. CUBRIC, during that time, moved to a new site with upgrade to the MRI scanner. The participants were therefore scanned on two sites with each participant having their paired scans at the same site and on the same scanner.

3.6.1 Withdrawal and loss to follow-up

Participants had the right to withdraw consent for participation in any aspect of the study at any time. Declining to participate or withdrawing from the imaging study did not affect participation in the TRON trial or their patient care.

If a participant initially consent but subsequently withdrew from the study, a clear distinction was made as to the aspects of the research the participant was withdrawing from. The following possibilities applied:

- Withdrawal from Imaging study but a continuation in TRON study.
- Withdrawal from TRON study treatment and follow-up but consenting to use of data including scans that were already performed.
- Withdrawal from treatment & follow-up and withdrawing consent to use of data.

3.6.2 Allocation and Blinding

Randomisation data for study groups were kept confidential as per the TRON study. The assessor of the scan data was blinded to the treatment allocation group, age or cognitive scores of the participant.

3.7 DTI MRI scan acquisition

3.7.1 Scanning procedure

DTI MR scans took approximately 40 minutes to acquire. Additional brain structural scans were performed using standard T1, T2 fast spin-echo & FLAIR sequences. The entire protocol took less than 1 hour.

CUBRIC 1 Scans (Park Place site): Diffusion-weighted imaging data were acquired using a 3T GE HDx MRI system (General Electric Healthcare) with a twice-refocused spin-echo echo-planar imaging sequence providing complete oblique axial (parallel to the commissural plane) brain coverage. Data acquisition was peripherally gated to the cardiac cycle. Each scan was acquired with a 3T scanner with 60 Directional DTI, Voxel Resolution of 1.7969 X 1.7969 X 2.4 mm and a field of view of 230x230 mm, b-value 1200s/mm^2. The TR (repetition time) was 18750ms and the TE (echo time) was 90.6ms.

CUBRIC 2 Scans (Maindy Road site): Diffusion-weighted Imaging (DTI) data were acquired using a 3T scanner, SIEMENS Prisma (Siemens Healthcare). Data acquisition was peripherally gated to the cardiac cycle. Each scan was obtained with a 3T scanner with 60 Directional DTI, Voxel Resolution of 2.0X2.0X2.0 mm and a field of view of 1152 x 1152mm and b-value of 1200 s/mm^2. The TR was 9500 ms while the TE was 70ms. CSF contamination artefacts were addressed using a multiple tensor variance approach, as per in-house software for pre-processing diffusion MRI data.

3.7.2 Data collection and retention

The DTI MRI data were stored as a confidential record on CUBRIC servers.

An anonymised CD-ROM was sent to the neuroradiology department for clinical reporting.

3.8 Primary outcome measures

3.8.1 White matter Tractography and tract specific measures:

It was planned to investigate the DTI measures of FA & MD values, for the white matter fibre tracts of the fornix (FX), uncinate fasciculus (UF), cingulum (CG) and superior longitudinal fasciculus (SLF) (all three components), in both hemisphere of participants' scans.

At the planning stage, we had envisaged that a single operator blinded to each participant's age or cognitive scores would draw all ROI manually. However, nearer to the study completion, the 'Tractseg' technique became a fast and reliable tool for analysis of the DTI measures (Wasserthal et al. 2018). We have used TractSeg to calculate the individual tract measures of FA and MD of the target tracts.

3.8.2 Evaluation of response

This study's primary outcome was a change in the value of FA and MD across the chosen white matter tracts of the study participants.

3.9 Statistical analysis plan

All randomised participants who had paired DTI scans (baseline and interval) were included in the data analysis. Data were reported descriptively, and results were presented split by the trial arm as well as grouped into the respective scanners locations.

3.9.1 Descriptive analysis

Summary statistics on eligibility, recruitment and withdrawal were collated for both trials arms and were presented. Specifically, for each arm, those participants who were randomised and completed the study protocol and were analysed for the primary outcome were documented.

The change in FA and MD values of the study participants were grouped into respective scanner subgroups to produce forest plots for every target tract.

3.9.2 Analysis of primary outcome

Data were to be presented descriptively by the trial arm and the respective scanner subgroups at baseline and 6 months. The placebo group was recruited for comparison but was not included in the primary analysis; however, the threshold of change was derived from the results of the placebo group values of each tract (as per the scanner group).

This analysis was a 2 stage process. R library was used for all statistical tests (R Core Team 2020).

- Wilcoxon exact signed rank test was performed to compare the FA and MD values at baseline and 6 months after everolimus treatment.
- Next, the direction of change was assessed by applying the one-sided sign test. The maximum absolute value of change (of the respective placebo group) was chosen as the threshold for the group.

3.10 Results

As the imaging data were acquired in two different scanners, the two datasets are not comparable and as such, could not be analysed as a homogenous group (Pierpaoli and Basser 1996a; Tax et al. 2019).

Table 24 describes the TRON imaging study participants' population characteristics (sex & treatment arm) and gives respective numbers across the two scanners. Table 25 details the reasons for exclusions. One participant (PID – 27) opted for a partial withdrawal, due to which only a baseline scan was available. The data from this participant was excluded from the analysis.

TABLE 24 Population characteristics of the TRON imaging study participants

	CUBRIC 1	CUBRIC 2	Total (n)
Male	7	5	12
Female	3	7	10
Everolimus	*7*	*10*	*17*
Placebo	*3*	*2*	*5*
No of participants	*10*	*12*	*22*

TABLE 25 Exclusions from the Imaging Study

Reasons for exclusion	number
Ferromagnetic implant	4
TRON study withdrawal	3
External sites	3
Claustrophobia	2
Scanner unavailable	2
Awaiting Study approval	2
Total	16

Ten participants' imaging data were acquired in the CUBRIC1 while 12 participants had their scans done in the CUBRIC2 scanner. All participants had their paired scans on the same scanner.

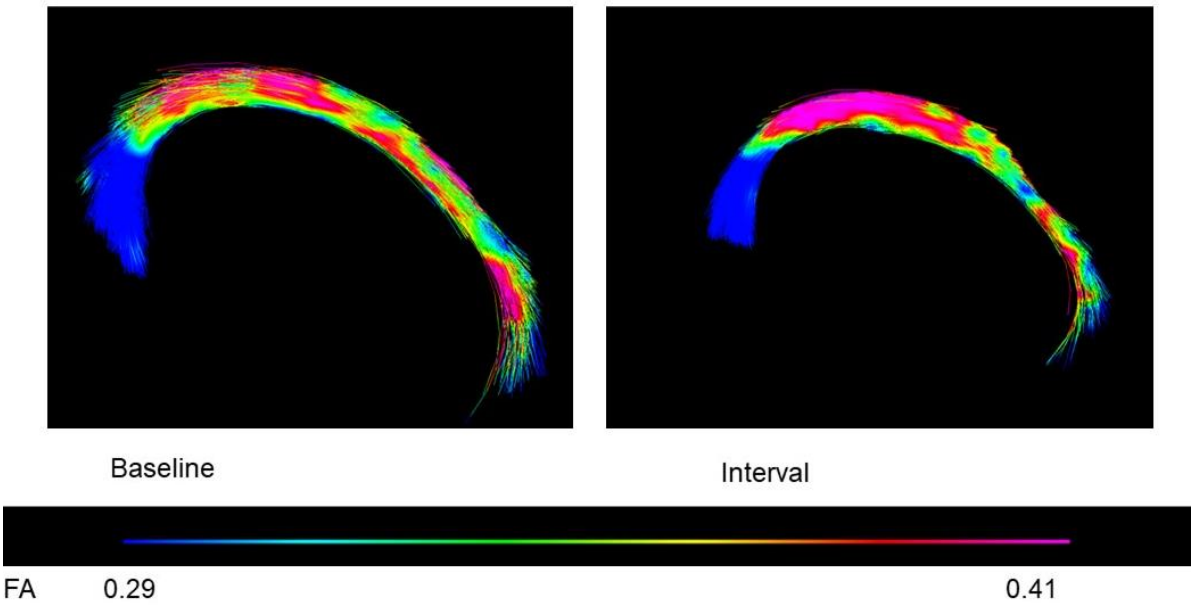


Figure 33 Baseline & Interval image of Fornix of a participant in the everolimus group. Deeper red/purple denote greater FA values. Images scaled as per the range of the values of the subgroup

In figure 33, although increase in FA in some areas of the fornix is apparent, the average values across the whole tract would not be signficant. In addition, the fornix is particularly vulnerable to CSF contamination based partial volume effects, which should be considered before drawing inferences about underlying changes in white matter structures (Metzler-Baddeley et al. 2012).

Forest Plots of change in FA and MD values of all the individual tracts were preprared after calculating the change in values, and their CIs. A sample of the change values of FA and MD for fornix (both sides) across the two scanners is illustrated below (fig 34 – 37). The data for all participants acquired on the same scanner is presented collectively. The remaining tract data are shown in the appendix.

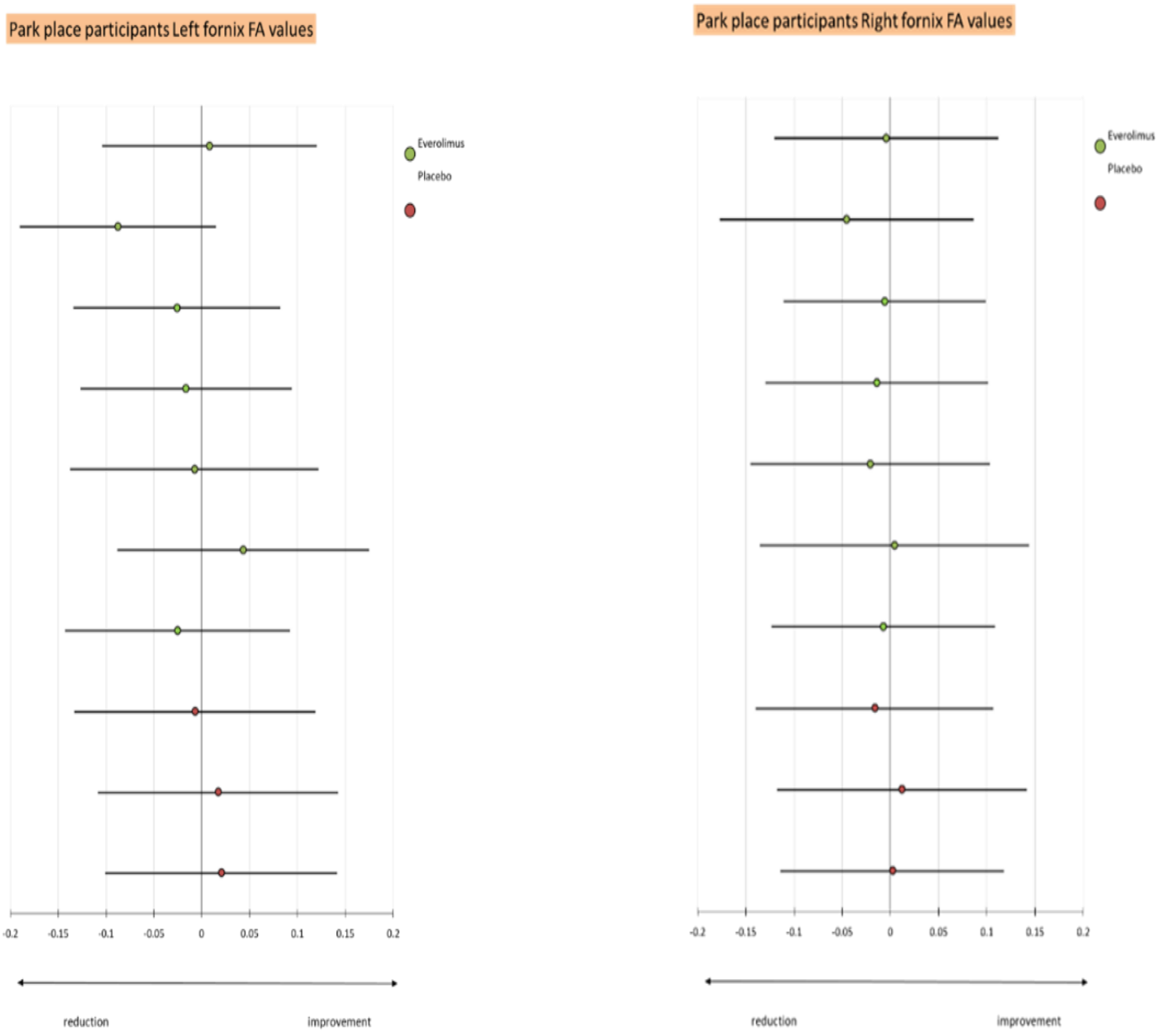
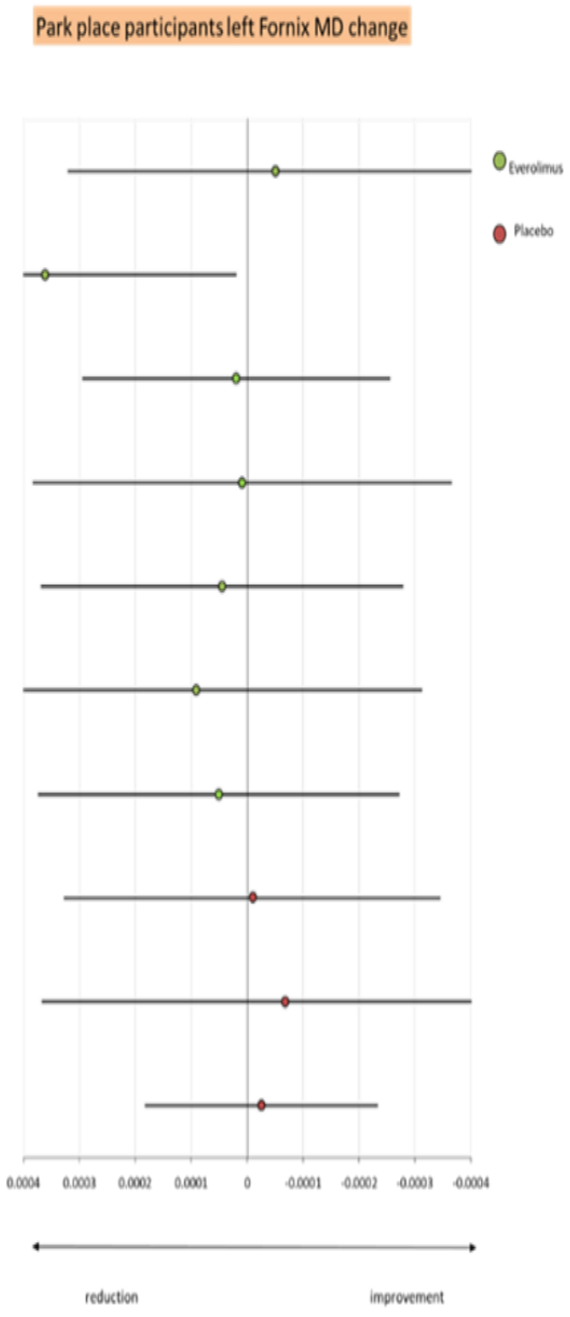


Figure 34 FA value change in Fornix (left & Right) for CUBRIC 1 participants

Figure 34 represents the change in FA values of Fornix in CUBRIC1 scanner participants, while Figure 35 shows a change in MD values for CUBRIC1 scanner participants.



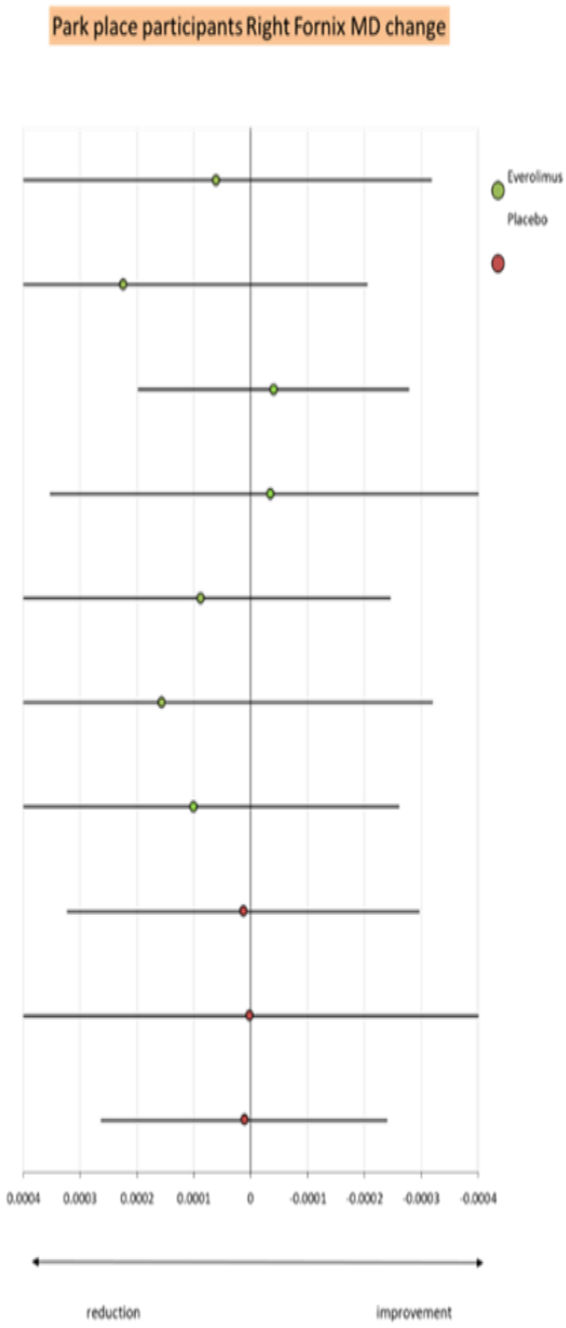
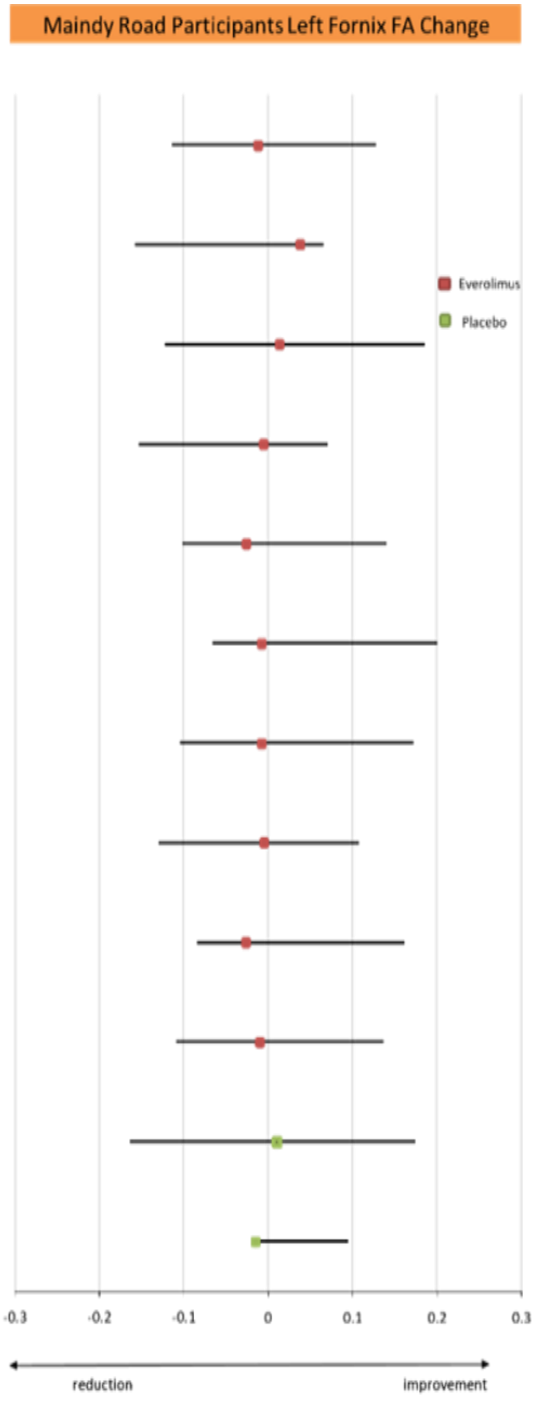


Figure 35 MD value change in Fornix (left & Right) for CUBRIC 1 participants



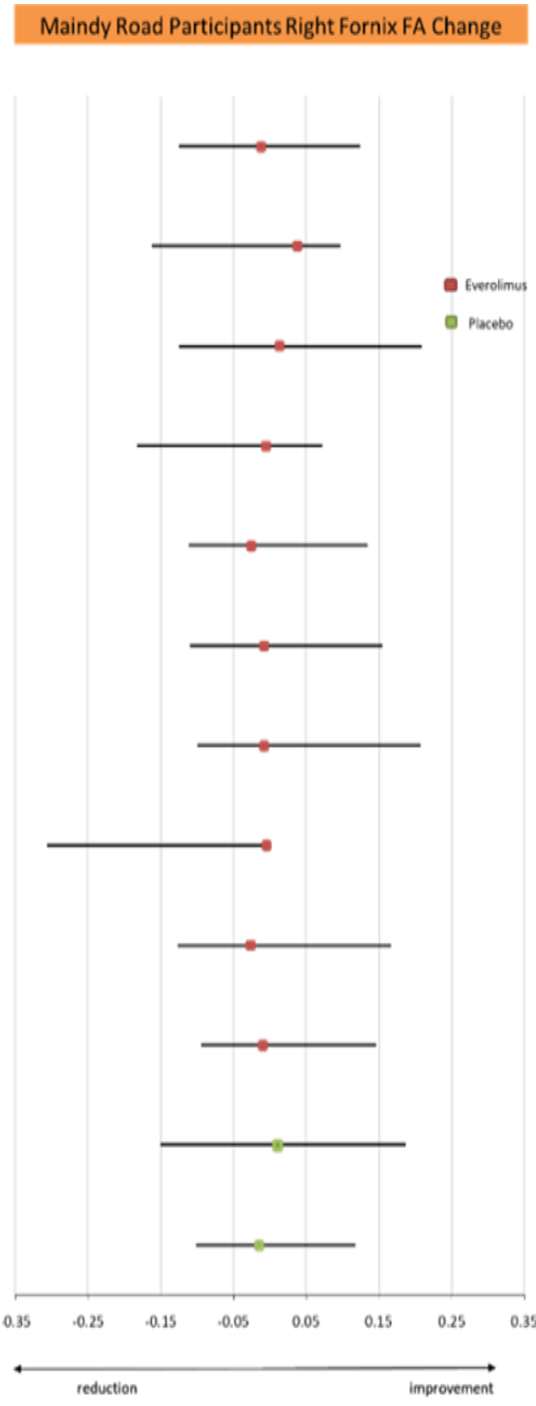
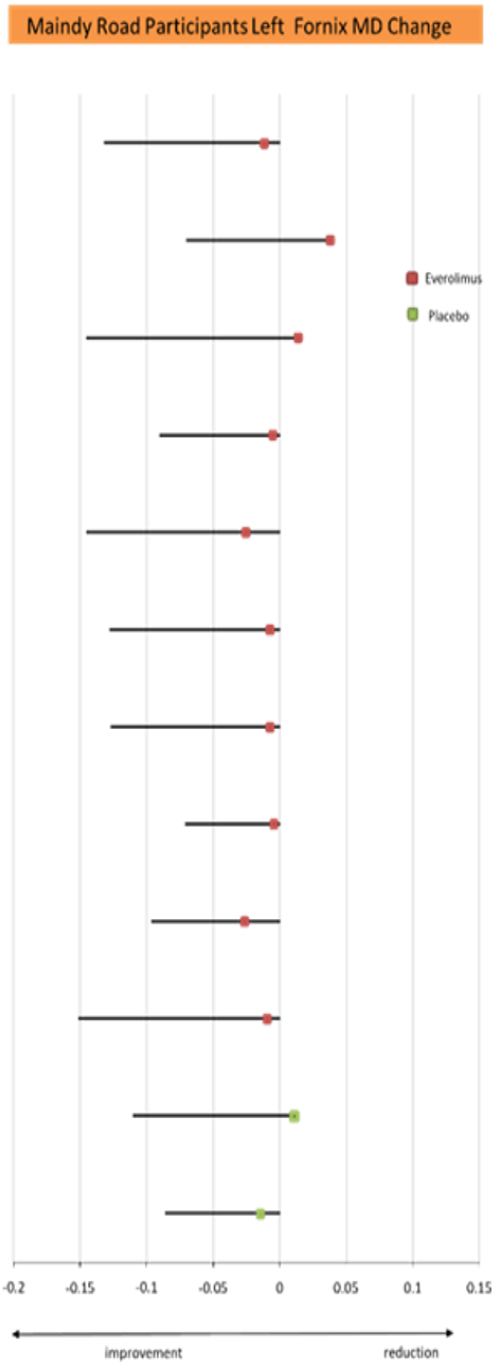


Figure 36 FA value change in Fornix for CUBRIC2 participants

Figures 36 shows the FA value change in the Fornix for CURIC2 scanner participants, and Figure 37 shows change in MD value in CUBRIC2 scanner participants.



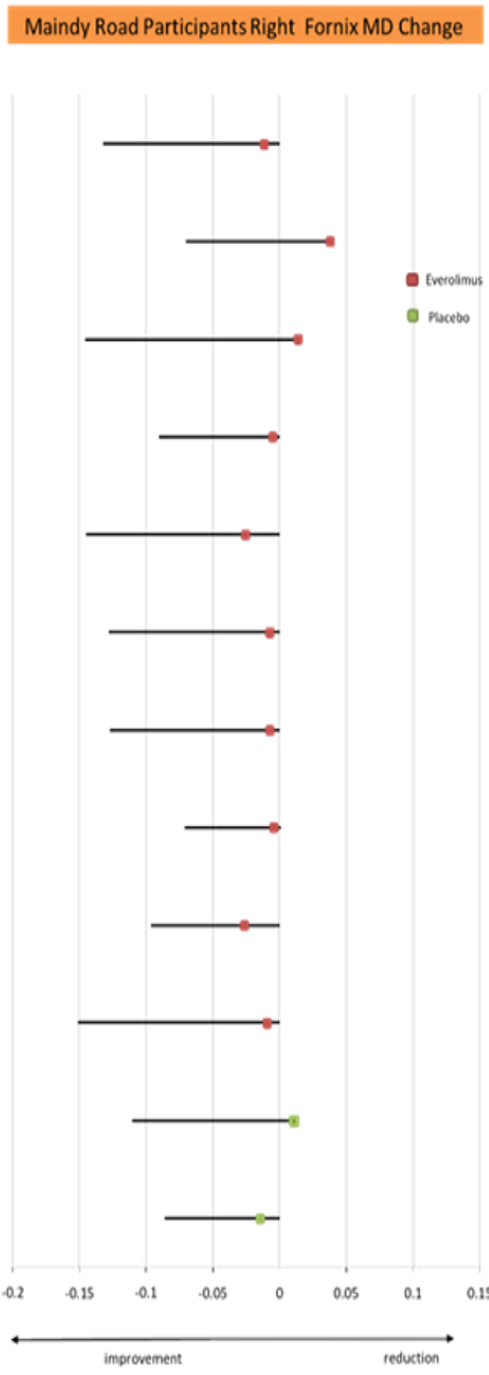


Figure 37 MD value change in Fornix (left & Right) for CUBRIC 2 participants

3.10.1 Primary Outcome analysis

The primary outcome analysis of participants treated with everolimus, FA and MD (as per their respective scanner subgroups) imaging data is presented in Table 26 and Table 27, respectively. The participants' calculated FA and MD values were compared at baseline visit and 6 months (24-week visit) scans using the Wilcoxon signed rank exact test. The significant p-values (≤ 0.05) are highlighted.

TABLE 26: FA change over 6 months in the target tracts of TRON imaging participants (p-value)

FA values	***CUBRIC -2 scans (n=10)***		***CUBRIC-1 scans (n=7)***	
Test	*left*	right	left	right
Fornix (FX)				
p value	0.43	0.92	0.37	0.03
Uncinate Fasciculus (UF)				
p value	0.84	0.76	0.93	0.07
Cingulum (CG)				
p value	0.69	0.37	0.03	0.078
SLF-1				
p value	0.69	0.69	0.10	0.68
SLF-2				
p value	0.69	0.76	0.046	0.10
SLF-3				
p value	0.23	0.62	0.29	0.078

SLF- Superior longitudinal Fasciculus

TABLE 27: MD change over 6 months in the target tracts of TRON imaging participants (p-value)

MD values	***CUBRIC -2 scans (n=10)***		***CUBRIC-1 scans (n=7)***	
Test	*left*	right	left	right
Fornix (FX)				
p value	0.55	0.76	0.10	0.07
Uncinate Fasciculus (UF)				
p value	0.62	0.49	0.46	0.015
Cingulum (CG)				
p value	0.62	0.76	0.015	0.15
SLF-1				
p value	0.69	1	0.015	0.03
SLF-2				
p value	0.49	1	0.07	0.03
SLF-3				
p value	0.76	0.52	0.15	0.03

SLF- Superior longitudinal Fasciculus

Nine of twenty-four FA and/or MD indices of the tracts derived from CUBRIC-1 scans showed significant change pre-and post-everolimus treatment (details below). However, none of the changes in the tracts of participants scanned in the CUBRIC-2 cohort reached statistical significance.

The FA values at six months, when compared to baseline, for the right fornix, left cingulum, and left superior longitudinal fasciculus (SLF) -1, reached a significant statistical difference (p-value <0.05). The MD values of the right uncinate fasciculus, left cingulum, both (right & left) SLF-1, right SLF-2, and right SLF-3 also had a statistically significant difference for the CUBRIC-1 scanner cohort. No tract parameters in the CUBRIC -2 participants had a statistically significant change.

However, on assessing the direction of change by applying the one-sided sign test, no statistical significance could be reached. On inspection of data from the CUBRIC 1 cohort, there was a reduction in FA values in 6 of 7 participants tracts where FA values had shown a significant p-value. The MD values had increased in all 7 participants for right UF, Left CG and Left SLF-1. It had also increased in 6 of 7 participants in right SLF1, right SLF2 and right SLF3.

3.11 Discussion

The TRON imaging trial was an exploratory study, which looked at the longitudinal effects of everolimus on DTI indices in the study population. Due to the study cohort's small size, all the scans were performed at CUBRIC, Cardiff, to minimise the scan-rescan variability to standardise scan quality. The project was delayed due to low recruitment in the TRON study, which lead to additional study centres opening in Glasgow and Belfast. The participants from these study centres found that travel to Cardiff challenging, further contributing to reduced recruitment.

CUBRIC had a scanner update during the study timeline, leading to participants' scans being acquired in two scanners. The longitudinal change in DTI indices of FA and MD of 12 individual tracts are presented descriptively. It was decided not to combine the two scanners' data for reasons discussed in the next section.

Some of the tracts in the everolimus group scanned at CUBRIC -1 showed a statistically significant change. These results were not replicated in the patient cohort using the CUBRIC2 scanner. Although scanner protocols were identical, the CUBRIC-1 scanner had a lower signal to noise ratio (SNR). In contrast, the CUBRIC-2 scanner has a higher SNR, making it more likely to pick up a change in values which may arise due to test re-test variability in DTI indices in the same individual. Despite these limitations, all significant changes in the everolimus treatment cohort were observed in the subgroup of patients scanned in CUBRIC1.

Data inspection revealed that in most of the CUBRIC1 cohort tracts demonstrating a significant difference, the FA had reduced, while the MD values increased after six months of treatment. This result is contrary to a previous report from a TSC children cohort (Tillema et al. 2012; Peters et al. 2019). An increase in FA and reduction in the MD values were observed with everolimus treatment. The assessed WM tracts in this cohort were (large tracts) not the same as the TRON imaging study. FA and MD values are affected by complex white matter architecture in smaller tracts due to CSF contamination, as well as due to the fibre populations having more than one dominant direction or crossing fibres, as discussed in section 3.12 (Douaud et al. 2011; Metzler-Baddeley et al. 2012; Vos et al. 2012).

Meaningful subgroup analysis was not possible due to the small numbers involved. As the two cohorts of patients did not overlap, the results obtained could reflect a difference in the participants themselves, given the small group sizes. Factors such as age, the severity of neuropsychological problems, drug dosage/levels during the study period need to be investigated further. These are preliminary findings, which

appear inconsistent across a small number of participants. As the results could not be replicated in the upgraded scanner, they must be treated with caution. Further work is required to confirm the impact of the everolimus on DTI measures in the target WM tracts of TSC individuals.

3.11.1 DTI data variation across scanners

Although the image acquisition protocol was matched in both scanners, there was a significant potential for inter-scanner variability. This disparity could be caused by various factors including, but not limited to, separate head coils used, different gradients (40 mT/m vs 80 mT/m), and changes in the system calibration - e.g., voxel resolution (1.79mm vs 2.0mm), the field of view (230x230mm vs 1152x1152mm), TR (18750vs 9500ms) & TE (90.6ms vs 70ms).

Besides, there are differences in preprocessing algorithms used to reconstruct the raw data, which can cause changes in the images acquired and the estimated diffusion measures such as FA & MD. Thus, aggregating data sets from different scanners is challenging due to the inherent differences in the acquired images (Mirzaalian et al. 2016). Specifically, the inter-scanner variability in FA and MD values is not uniform over the entire brain but is tissue-specific and region-specific (Mirzaalian et al. 2015). Inter-scanner variability in FA can be up to 5% in major white matter tracts and between 10 and 15% in gray matter areas (Vollmar et al. 2010).

3.11.2 Data Harmonising techniques

Harmonising data from different scanners could potentially increase the statistical power and sensitivity of studies, with apparent benefits in trials and multicentre research, particularly in rare diseases or with difficult-to-recruit participants (Tax et al. 2019), such as tuberous sclerosis. The techniques deployed include using physical phantoms to detect scanner-specific variability and changes to correct for such variability with harmonisation approaches (Prohl et al. 2019). Such phantoms are incapable of fully capturing the complexity of biological tissue and regional differences associated with this complexity. Furthermore, it can be non-trivial to translate the differences observed in physical phantoms to in-vivo acquisitions.

On the other hand, data harmonisation algorithms rely on complex computational tools to reduce the cross-scanner and cross-protocol variability to a similar level as scan-rescan variability using the same scanner protocol (Tax et al. 2019; Ning et al. 2020). However, these are beyond this study's scope and were not used for harmonisation of data acquired with the two different scanners.
The number of placebo group participants was small (5 out of 22 – 3 in CUBRIC 1 scanner; 2 in CUBRIC 2 scanner) compared to the treatment group as not all the TRON participants took part in the imaging study.

If larger, the placebo group data could inform the scan to scan variability measures and provide a threshold for significant change; however, it remains impractical in the current study due to small numbers.

3.11.3 Comparison with previous studies

Previous studies of the longitudinal effects of everolimus in 28 TSC patients reported a significant change in FA (increased) & MD values (reduced) of the corpus callosum, internal capsule and geniculocalcarine tracts at 12-18 months. These DTI measures were reported again at 3.5years of follow-up with similar results (Tillema et al. 2012; Peters et al. 2019). The trough everolimus levels were comparable with the TRON study. The studies' results were combined from scans using 6 different scanners, both 1.5T and 3T, in contrast to the TRON study methodology. We were not able to replicate these results in the TRON imaging study. This may be due to the different age group (approximately half were aged<10years in the studies of Tillema et al. and Peters et al.), tracts studied, length of treatment (6m in TRON study vs >12 m), DTI processing approach (manual ROI in the previous studies vs TBSS in TRON) and scan methodology.

Another small study evaluating evolving diffusion MRI measures in 17 subjects with TSC (mean age, 7.2 ± 4.4 years) used regions of interest analysis to assess the internal capsule/corona radiata, cingulum, and corpus callosum. Mean change in Apparent Diffusion Coefficient (ADC) was significantly smaller in boys in the left internal capsule, right and left cingulum bundles, and corpus callosum, with no such effect in FA values. Epilepsy was a significant predictor of mean change in ADC in the left internal capsule but no other white matter tracts. Autism spectrum disorder was not predictive of diffusion changes in any studied pathways (Baumer et al. 2015). This was an interesting study looking at the longitudinal evolution of diffusion measures in young children. The authors did not give the reasons for their choice of tracts that were studied; most tracts were not comparable to the TRON imaging study. The study did evaluate longitudinal changes in FA values in the cingulum, similar to TRON, which showed no change over time. The study was limited by small numbers and the results showing a gender difference are yet to be replicated.

One further study evaluated the corpus callosum's DTI metrics in children with TSC, (non-TSC) autism, and healthy controls and reported lower FA values for the corpus callosum in TSC children than children with autism without TSC which were themselves lower than the control group. This was a retrospective study with many participants in eachgroup (Baumer et al. 2018). The study authors observed that increasing neurological comorbidity (such as Intellectual disability, epilepsy & autism) in TSC was associated with a lower FA value of corpus callosum, demonstrating an additive effect. This study is not comparable to the TRON imaging study due to the age group (children versus adults population), study design (retrospective versus prospective) as well as being an observational study rather than having an intervention (everolimus) as in the TRON imaging study.

The multicentre TACERN study group has published a series of studies looking at DTI measures in white matter tracts in the study cohort of TSC infants with autistic spectrum disorder. The DTI indices were acquired where routine MRI scans were undertaken for clinical surveillance of TSC infants on seven scanners, as available at study sites. Mean FA and MD values were computed for 17 white matter regions. Expert hand-drawn region of Interest (ROI) analysis was performed within white matter fibre bundles. ROI analysed included the left and right posterior limb of the internal capsule, anterior limb of the internal capsule, and corpus callosum. Developmental measures such as the Mullen scale of early learning (MSEL) and autism diagnostic observation schedule (ADOS) were administered at 24 months of age, along with developmental quotient assessment (DQ). Measures of epileptic seizure frequency, details of medication, and seizure type were also noted. The study reported reduced FA values in the regions of arcuate fasciculi, corpus callosum, cingulum, sagittal striatum and anterior limb of the internal capsule in the cohort with TSC and autism when compared to TSC without autism. The authors suggest lower FA values in the corpus callosum as an indicator of developing autism. However, reduced FA values might simply indicate more severely affected TSC children who are more likely to develop autism. This is one of the few studies looking at the longitudinal evolution of DTI measures in young children with TSC and is not comparable to the TRON imaging study as the participants' ages are different. Also, the evolution of DTI parameters is not assessed after an intervention (such as everolimus). Instead, it reflects the passage of time (Prohl et al. 2019).

3.12 Conclusion

The TRON imaging study, to our knowledge, is the first study to investigate longitudinal evolution of DTI measures in a subgroup of adults with TSC, treated with everolimus, who have neurocognitive problems with memory or executive functions. It is also one of the first studies to evaluate DTI metrics in white matter fibres of fornix, cingulum, uncinate fasciculus, and superior longitudinal fasciculus (three components), which are likely to be related to neuropsychological deficits seen in the study population.

Previous studies' results indicate that a higher FA value and lower MD value are indicative of a healthier tract. This understanding is based on the study of large white matter tracts such as the corpus callosum, which is frequently evaluated in numerous studies. Abnormal DTI indices of the corpus callosum are then inferred to be a clinical marker for the condition under study, with TSC studies reaching similar conclusions. Callosal DTI metric abnormalities are evident in many preclinical models of brain injury (Yu et al. 2017) and varied clinical conditions, including refractory epilepsy, bipolar disorder, Parkinson's disease, and autism, posing doubts on specificity as a biomarker for a particular disease process (Caligiuri et al. 2016; Aoki et al. 2017; Abramovic et al. 2018; Bledsoe et al. 2018).

FA values can also be affected by the 'crossing fibres' effect, which is especially pertinent for the smaller tracts that we have evaluated in the TRON study such as fornix, while MD values are hard to interpret in white matter regions containing fibres with more than one dominant direction (Douaud et al. 2011; Vos et al. 2012). CSF contamination with partial volume effects could also affect the DTI metrics of fibres close to the ventricle, such as fornix. The TRON study results should be interpreted in light of these limitations of the current DTI analysis (Metzler-Baddeley et al. 2012).

It remains uncertain to what extent an improvement in DTI indices (a biological marker) reflects clinically relevant changes in the study population.

The limitations of doing multicentre studies using different scanner data are also evident. However, there is a necessity for multicentre studies in rare diseases, which needs to be considered when future studies are being designed

Future studies:

Multicentre studies are a solution for meaningful statistical analysis in rare diseases such as TSC. However, this is likely to lead to inconsistency in scan data, as discussed in section 3.11.1. The TRON study's lessons are to plan for variation of the scanner's qualities in any future DTI studies. In recognition of this issue, the TACERN study group use phantoms (human & non-human) to standardise the data (Prohl et al. 2019), but these have limitations. It would be advantageous if future studies not only use similar scan protocols and have comparable scanners across multiple sites but also incorporate data harmonisation based on advanced computational strategies at the planning stage

The choice of tracts is crucial to interrogate brain structure in relation to the neuropsychological profile of TAND and should be the primary focus of a future study. However, most of the DTI studies in TSC have looked at large white matter tracts, whose parameters may not be related to memory, learning, or behaviour. It is difficult to be confident about tract choice in the absence of normative data. Ideally, a whole-brain study in where no *a priori* hypothesis is made regarding anatomical localisation of white matter abnormalities (Douaud et al. 2011) or their correlation with the neuropsychological profile of TSC individuals should be undertaken before an everolimus effect study is performed.

4 Conducting future trials in TAND aspects of TSC – challenges and opportunities

The TRON clinical trial was designed to examine the effect of everolimus on memory and executive function problems seen in TSC individuals. The subjects selected comprised individuals who did not have uncontrolled epilepsy, which would make a reliable assessment of these problems difficult. An imaging study was conducted on TRON participants to examine the longitudinal effects of everolimus on DTI indices of fractional anisotropy (FA) and mean diffusivity (MD).

Participants in the TRON study were randomised 2:1 to take everolimus or placebo. The trial had significant difficulties in recruitment; however with delay the study did recruit the adjusted target of 38 participants. On initial analysis of the results, in the placebo group, it became apparent that results from a TEA subtest (Telephone Search whilst Counting - TSwC) showed a pattern suggesting differences in performance at alternate assessments. Further investigation revealed non-comparability of the two versions of the instrument that were alternated during assessments. A decision was made to take the data of those participant eligible based only on the TEA test out of the final analysis, following discussion with the trial statisticians. The primary analysis results then showed that 44% of the placebo group participants improved any one or more of the primary outcome measures by 1SD, while the improvement in the everolimus group was 70%. The difference in effect size between the two arms of more than 20% fulfilled the threshold agreed for the TRON trial outcome supporting future larger trials.

In the imaging study, the participants' results were grouped according to which of two scanners had been used to obtain images. Interestingly, the analysis shows a change in DTI parameters in some of the tracts in the older (inferior) scanner cohort, which was not replicated in the cohort scanned in the newer (superior) scanner. The apparent change in DTI indices in some adult patients with TSC following six months of treatment with everolimus merits further investigation.

4.1 Clinical trial Design challenges in rare diseases

TSC is one of the estimated 7000 or more rare diseases so far identified. A disease is defined as rare if the prevalence is not more than five per 10,000 population (Hilgers et al. 2016) as per the European Union criteria and the current UK definition. (DOHSC 2013). In the United States, a rare disease is defined as a condition that affects fewer than 200,000 people (almost 6 in 10,000) (FDA 2019) and in Japan fewer than 50,000 people (about 4 per 10,000) (Kempf et al. 2018).

The 7000 different rare diseases affecting approximately 6 % of the global population at some stage in their life (Colledge and Solly 2012; Kempf et al. 2018) but individually, rare disease patient populations are small and geographically dispersed and may include vulnerable groups such as children or people with intellectual disabilities. Consequently, trials in these low prevalence conditions face many challenges, including heterogeneity in the patient population, difficulty in clinical trial recruitment, often a poorly understood natural history of the disease, underlying molecular biology and response to treatments. The main issue setting rare diseases apart from drug development for common diseases is the challenge of generating acceptable evidence from clinical trials involving a small number of participants due to limited recruitment (EMA 2006; Kempf et al. 2018).

4.2 Expediting Drug development in rare diseases

The development of drugs to treat rare diseases raises some unique issues recognised by the pharmaceutical industry and by regulators and governments. Some of the endeavours to address this neglected and often unprofitable public health issue are discussed below.

4.2.1 Orphan Status

'Orphan status' designation is an initiative for drugs in the US, EU, and UK associated with financial incentives for drug development and research (EMA 2019; FDA 2019; MHRA 2021). The rare disease label is vital as it leads to 'orphan status' for a drug or biological product intended to treat a rare disease. However, granting an orphan status does not markedly alter the regulatory requirements or process for obtaining marketing approval.

4.2.2 Time considerations

The International Rare Disease Research Consortium (IRDiRC) guidance calls for timely completion and dissemination of research outcomes, even if the results are not entirely convincing (IRDiRC 2020). In contrast, the Committee for Medicinal Products for Human Use (CHMP) guidelines for clinical trials in small populations suggests a deviation from accepted rules and guidance based on a randomised controlled trial should not be routine (EMA 2006). Deviation from such standards is uncommon and should only be considered when completely unavoidable and needs adequate justification. In contrast, the FDA has a series of programs instituted to facilitate the development of treatments for serious diseases, especially for diseases without any approved therapies, to reduce the development and approval time for such drugs. These programs include fast track designation, breakthrough therapy designation, priority review

designation, and accelerated approval, described in the FDA Guidance for Industry: Expedited Programs for Serious Conditions(FDA 2014).

TRON study's recruitment and completion took more than 6years. The timeframe taken suggests that routine processes may not be feasible for future studies for similar indications.

4.3 Choice of Outcome measures

A biological response (improved neurocognitive scores) was the TRON study's primary scientific curiosity, prompted by findings in preclinical animal studies (Ehninger et al. 2008a). The trial team hoped that a response would be seen in patients, as demonstrated in studies of mTOR inhibitors in TSC complications such as SEGA (Bebin et al. 2011) and AMLs (Bissler et al. 2008; Davies et al. 2011b).

In relation to neurocognitive aspects of TSC, the 'most appropriate' clinical endpoint in terms of response is not agreed upon, in contrast to epilepsy trials (50% reduction in seizure frequency) or tumour size trials. The use of other surrogate markers as substitutes for a clinical endpoint may be considered. However, selecting a surrogate marker, such as DTI indices in TRON participants, as a study endpoint requires it to be reasonably likely in predicting benefit. Studies such as the TACERN multi-centre study may inform the use of DTI biomarkers in the future (Prohl et al. 2019). Demonstrating that such a surrogate endpoint adequately reflects a clinically meaningful endpoint remains currently problematic.

Surrogate markers in themselves cannot serve as final proof of clinical efficacy or long-term benefit. Therefore if the markers are used for regulatory review and approval without being validated, there should be a predetermined plan to supplement such studies with further evidence to demonstrate clinical benefit and meaningful, real-world outcomes for patients (EMA 2006).

4.4 Choice of controls

A placebo group acting as a comparator is ideal for an unbiased estimate of the treatment's effect. Its value was demonstrated in the TRON study, where learning effects were much greater than expected.

True clinical equipoise to "new promising therapies" is challenging to achieve, not only for individual investigators but also for a wider medical community involved in a rare disease management. This positive belief inevitably increases if early results appear promising. "Hope" on the part of the clinicians, patients and families can bias the results in trials lacking appropriate controls (Kaptchuk and Miller 2015). If participants feel they are in the placebo group, they may opt for withdrawal from the study, although it was not a frequent TRON study experience.

Future TAND studies could have internal controls (as in the TRON study) or external controls, which could be historical or concurrent. Epidemiological data and data from patient registers may provide some help. There is an urgent need for rigorously collected natural history and patient registry data for rare diseases. One benefit would be increased availability of data for use as an external control in clinical trials, rather than relying on randomised controls (Day et al. 2018). The TuberOus SClerosis registry to increase disease Awareness' (TOSCA) data is unusable as a comparator for the TRON study due to missing details for TAND (Kingswood et al. 2017).

4.5 Natural history and Patient registry

In future, a data registry could be mandated as a post-licensing arrangement for a rare disease. Regulatory approval should not end the collection of safety data; instead, improving data quality should be prioritised (Day et al. 2018). The Yellow Card scheme in the UK and the Canada Vigilance Program are two of the most successful postmarketing surveillance systems implemented worldwide (Raj et al. 2019). These schemes are primarily concerned with significant adverse drug reaction in the general population and therefore do not routinely monitor the efficacy element. Hence, post regulatory efficacy data (patient registry) as a Phase 4 study could be made mandatory in rare diseases' approval conditions. Recent examples of this approach can be noted in a combined registry for children with Spinal Muscular atrophy (SMA), treated with novel therapeutic agents (SMAReachUK 2015) as mandated in the Nusinersen managed access agreement (NICE 2019). The FDA is endorsing the importance of developing natural history studies by the disease-specific stakeholder organisations (patient/family driven data collection) in association with the National Organisation for Rare Disorders (NORD 2016). In time, these interventions would help recruitment for rare disease clinical trial in a timely fashion.

4.6 Dosage

The credibility of study results may be enhanced if a dose-response relationship is seen clearly or in cases where a chain of events can be identified (for example, drug exposure to pharmacodynamic measures, to a clinical outcome). Where no such clear chain of events exists, clinical trials are much less convincing and data requirements are increased regarding robustness and persuasiveness of study results. EXIST3 results that showed better clinical response with higher trough levels of everolimus were not available at the time the TRON study protocol was developed.

4.7 Patient recruitment

Advancing drug development for rare diseases requires cooperation and collaboration among diverse stakeholders. The Tuberous Sclerosis Association (TSA) supported TRON study recruitment by advertising the trial on their web page and in their newsletters, along with providing funds to support travel and expenses for participants. TSA input, however, did not translate significantly into recruitment. The TSA database does not have detailed clinical information on their membership. Therefore, the approach was not focussed. This situation contrasts with Fragile X studies with multiple phases 2 and phase 3 clinical trials completed (Davenport et al. 2016), where the role of patient association has been crucial.

The most substantial impact in recruitment was the role of clinical teams based in TSC clinics. In these clinical teams, the specialist staff had detailed knowledge of the TSC patients' clinical and personal profiles and were trusted by their patient cohorts to act in their best interests. A face-to-face consultation explaining the study proved to be the most effective recruitment strategy.

4.8 Trial Design

Hierarchies of evidence have been described. The credibility of the evidence in descending order is meta-analyses of RCT followed by individual randomised controlled trials, after that meta-analysis of observational studies followed by individual observational studies, followed by case reports and expert opinions the field (EMA 2006). The CHMP guide for trials with small populations recommends that the trial methodology should not be exceptional in rare disease indications compared to large studies (EMA 2006).

The gold standard for clinical trials remains a randomised clinical trial with endpoints that are clear and meaningful to patients (Odgaard-Jensen et al. 2011). An easily interpretable study that clarifies both the efficacy and the safety profile of a drug allows for straightforward interpretation by patients, regulators, clinicians, and commissioners (Kempf et al. 2018). A variety of strategies are employed to improve the clinical trial efficiency, which is essential for a clinical trial in small populations

4.8.1 Enrichment

Optimisation of patient selection is central to the likelihood of success. Using enrichment strategies to focus on and accelerate drug evaluation by identifying a study population is most likely to provide evidence that the drug is effective. There are three principal enrichment strategies. Practical enrichment is a strategy to decrease heterogeneity of the study population and to reduce background variability, while prognostic enrichment is a strategy to identify high-risk subjects more likely to experience the poor outcomes that are

being measured in the study, and predictive enrichment is a strategy to select subjects more likely to respond to the candidate treatment

Predictive enrichment strategies is needed in scenarios such as: if the beneficial effect of the candidate drug in a small population subset that responds to the treatment may be obscured by the much larger unselected population (Kempf et al. 2018). These subgroups can be identified by a biomarker such as a specific genetic mutation in cystic fibrosis cases (McKone et al. 2014) and greater antiseizure effects of cannabidiol prescribed together with clobazam (Bialer and Perucca 2020).

The TRON study had strict inclusion/exclusion criteria and omitted confounders such as uncontrolled epilepsy, a practical enrichment strategy. However, incorporating prognostic and predictive enrichment strategies was challenging as data (natural history) for the TAND response to everolimus was not available before study initiation. An interim analysis to understand the response characteristics to everolimus could not be attempted due to blinding. Future trials could employ an adaptive design where interim analysis of drug response may help with the prognostic enrichment.

4.8.2 Study Visits

Traditional trial designs requiring the patient to come in person to frequent visits at study sites is resource-prohibitive for participants and study budgets. For many rare diseases, there may be only a few treatment sites in the country or the world, and study participation would place an unreasonable travel burden on the participants. In TAND trials, due to the specific challenges with travel for participants, an alternative arrangement would be a welcome change. A change to virtual neuropsychological assessment could be explored to allow participants increased access, convenience, and cost-saving (Brearly et al. 2017; Wadsworth et al. 2018). This approach has become the 'new normal' with the COVID-19 pandemic (APA 2020).

In future TAND clinical trials, virtual assessments could minimise in-person visits and be used for a preliminary suitability assessment. The potential participants would also benefit from being introduced, virtually, to the clinical trial team, improving their willingness to attend a screening visit.

4.8.3 Methodological approaches to a TAND clinical trial

Innovative study designs have been systematically reviewed to inform clinical trials in rare diseases with a small patent population. The recommendations for study designs for interventions in rare diseases (with reversible treatment effect) include Crossover design, N-of-1 trials or one of the adaptive designs (Gupta et al. 2011).

The first two approaches are suitable for situations where the response to the intervention is relatively quick (within a few weeks). If a required number of participants are available, a crossover design is preferred, while for situation of very low numbers, serial N-of-1 trails would be the recommendation. These approaches are unsuitable for TAND trials due to an expected response being over months rather than weeks.

In clinical trials where the response to the intervention is slow, one of the adaptive designs, which allow researchers to dynamically adapt the design without compromising the trial's validity or integrity, could be utilised. These design adaptation strategies should be planned prospectively to address sample size issues in contrast to ad hoc design changes (Gupta et al. 2011). The various designs described are the response-adaptive randomisation design (intervention is assumed better), ranking and selection design (best option evaluation of several interventions), internal pilot design (estimating the final sample size dependent on pilot phase data), and sequential design (repeated interim analysis to guide sample size). These adaptive designs are dependent on clearly agreed outcome measures, which still need to be established for TAND. This would also be applicable for an internal pilot design, which appears to be most appropriate for the TAND population. These strategies are also dependent on reasonable recruitment, which is not guaranteed for a future TAND trial..

Conventional RCTs with trial efficiency measures planned prospectively, therefore, continue to be the approach of choice for future TAND trial. Consideration to minimise time on placebo and, if feasible to implement a design that ensures that all participants receive active treatment by the end of the study would make the trial more attractive to eligible individuals (Gagne et al. 2014; Whicher et al. 2018)

4.9 Data analysis

At the study conclusion, the treatment effect should ideally be clinically relevant, confidence intervals for that effect should be narrow, and the effect size statistically significant. Well-planned and well-conducted meta-analyses of such trials will provide even more robust evidence. Studies with few patients are often presented as simple (descriptive) analyses as there is not much data, but in this situation a more complex approach may, paradoxically, be necessary (EMA 2006).

4.9.1 Non-parametric methods

A small study population may lead to clustering of results (such as a change in a test score or DTI indices) when it cannot be determined if data are from a normal (or other specified) distribution. Therefore, non-parametric, or 'distribution-free' methods should be used to analyse the results, such as the Wilcoxon signed-rank test employed in the TRON imaging study.

4.9.2 Bayesian methods

Bayesian methods for estimating a treatment effect start with formal incorporation of prior belief regarding the size of the effect and update this belief as new data accumulates. This analytical method is a way to formally combine knowledge from previous data from prior studies or sequential data analysis in a single study. This approach can be applied to conventional and novel designs (Lilford et al. 1995; Spiegelhalter et al. 1999).

The Bayesian approach could measure the magnitude of the treatment effect, which may not be possible by traditional analysis methods in a small sample size (Abrahamyan et al. 2014). The inclusion of prior information remains controversial as it might be based on biased data. There are risks in integrating all evidence into a single analysis rather than a series of individual studies that could be mutually supportive and more likely to generate a favourable response from regulatory authorities. The International rare disease research consortium has recommended a pragmatic adoption of Bayesian approaches notwithstanding the known arguments (Day et al. 2018).

4.10 Conclusions

Every effort should be made to improve clinical trial efficiency with due diligence to identify the population being studied, meaningful outcome measures with (preferably) internal controls or robust external controls. Trials should be designed in collaboration with patient groups, regulatory authorities, biostatisticians, and clinical trialists who have interest, experience, and expertise in rare disease drug development (Kempf et al. 2018).

The two previous TAND studies (Krueger et al. 2017, Overwater et al. 2019) concluded that everolimus treatment resulted in no significant improvement in TAND, while the TRON study has reported some improved neuropsychological test scores in its placebo and everolimus arms, with a greater effect size in the everolimus arm. These results reflect the diverse populations (age and phenotype), different outcome measures (neurocognitive tests) applied, and different everolimus dosages and treatment durations used in these trials. In the TRON study cohort, DTI indices in a limited number of white matter tracts showed a change after everolimus treatment. However, as these findings were seen in only one of two treated patient groups, the results should be interpreted cautiously. These preliminary results would justify further clinical trials to investigate everolimus as a treatment option in TAND.

Experimental therapeutics for TSC manifestations face many obstacles. Despite these challenges, several TSC clinical trials have been completed, resulting in everolimus being licensed in managing SEGA, AMLs and TSC related epilepsy, while sirolimus is licensed for LAM. Further research priorities should include continued improvement of our understanding of TAND natural history and phenotype variation by establishing a detailed patient registry. Continued monitoring of safety and efficacy data after licensing should be mandatory for rare diseases to improve the understanding of natural history, especially if treated with disease-modifying agents.

Clinical trials for TAND in the future should disseminate data for all the assessments in detail, even if the results are not convincing. A clinically meaningful outcome measure in future studies for TAND should be based on this data in consultation with patient groups such as the TSA. There are numerous challenges in performing meaningful clinical trials in rare diseases and contexts, such as TAND, that can be overcome with a carefully planned study design based on prior knowledge of the disease's natural history and innovative trial design and analysis methodology.

References

1993. Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* 75(7), pp. 1305-1315. doi: https://doi.org/10.1016/0092-8674(93)90618-Z

A.K Coughlan, M. O., J.R. Crawford. 2007. BIRT memory and information processing battery (BMIPB) test manual

Brain Injury and Rehabilitation Trust (BIRT), Horsham (2007),

AACN, B. o. D. 2007. American Academy of Clinical Neuropsychology (AACN) Practice Guidelines for Neuropsychological Assessment and Consultation. *The Clinical Neuropsychologist* 21(2), pp. 209-231. doi: 10.1080/13825580601025932

Abraham, R. T. and Eng, C. H. 2008. Mammalian target of rapamycin as a therapeutic target in oncology. *Expert opinion on therapeutic targets* 12(2), pp. 209-222.

Abrahamyan, L. et al. 2014. A new toolkit for conducting clinical trials in rare disorders. *Journal of Population Therapeutics and Clinical Pharmacology* 21(1),

Abramovic, L. et al. 2018. White matter disruptions in patients with bipolar disorder. *European Neuropsychopharmacology* 28(6), pp. 743-751. doi: https://doi.org/10.1016/j.euroneuro.2018.01.001

Afzali, M. et al. 2011. Tract based spatial statistical analysis and voxel based morphometry of diffusion indices in temporal lobe epilepsy. *Computers in Biology and Medicine* 41(12), pp. 1082-1091. doi: https://doi.org/10.1016/j.compbiomed.2011.05.006

Agricola, K. et al. 2013. Nursing Implications for the Lifelong Management of Tuberous Sclerosis Complex. *Journal of Neuroscience Nursing* 45(4), pp. 226-242. doi: 10.1097/JNN.0b013e3182986146

Albert, M. L. 1973. A simple test of visual neglect. *Neurology* 23(6), pp. 658-664. doi: 10.1212/wnl.23.6.658

Alexander, A. and Walker, C. L. 2011. The role of LKB1 and AMPK in cellular responses to stress and damage. *FEBS Letters* 585(7), pp. 952-957. doi: https://doi.org/10.1016/j.febslet.2011.03.010

Altman, N. R. et al. 1988. Tuberous sclerosis: characteristics at CT and MR imaging. *Radiology* 167(2), pp. 527-532. doi: 10.1148/radiology.167.2.3357966

Amin, S. et al. 2019. The UK guidelines for management and surveillance of Tuberous Sclerosis Complex. *QJM* 112(3), pp. 171-182. doi: 10.1093/qjmed/hcy215

Aoki, Y. et al. 2017. Association of White Matter Structure With Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. *JAMA Psychiatry* 74(11), pp. 1120-1128. doi: 10.1001/jamapsychiatry.2017.2573

APA. 2020. Guidance on psychological tele-assessment during the COVID-19 crisis. American Psychological Association

Arfanakis, K. et al. 2002. Independent component analysis applied to diffusion tensor MRI. *Magn Reson Med* 47(2), pp. 354-363. doi: 10.1002/mrm.10046

Arulrajah, S. et al. 2009. Magnetic resonance imaging and diffusion-weighted imaging of normal-appearing white matter in children and young adults with tuberous sclerosis complex. *Neuroradiology* 51(11), pp. 781-786. doi: 10.1007/s00234-009-0563-2

Asano, E. et al. 2001. Autism in tuberous sclerosis complex is related to both cortical and subcortical dysfunction. *Neurology* 57(7), pp. 1269-1277. doi: 10.1212/wnl.57.7.1269

Au, K. S. et al. 2007. Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States. *Genet Med* 9(2), pp. 88-100. doi: 10.1097/gim.0b013e31803068c7

Axelrod, B. N. 2002. Validity of the Wechsler abbreviated scale of intelligence and other very short forms of estimating intellectual functioning. *Assessment* 9(1), pp. 17-23. doi: 10.1177/1073191102009001003

Ay, H. et al. 1998. Posterior leukoencephalopathy without severe hypertension: utility of diffusion-weighted MRI. *Neurology* 51(5), pp. 1369-1376. doi: 10.1212/wnl.51.5.1369

Bader, R. S. et al. 2003. Fetal rhabdomyoma: prenatal diagnosis, clinical outcome, and incidence of associated tuberous sclerosis complex. *J Pediatr* 143(5), pp. 620-624. doi: 10.1067/S0022-3476(03)00494-3

Baker, G. A. et al. 1991. The development of a seizure severity scale as an outcome measure in epilepsy. *Epilepsy Res* 8(3), pp. 245-251. doi: 10.1016/0920-1211(91)90071-m

Baker, G. A. et al. 1998. Liverpool Seizure Severity Scale revisited. *Seizure* 7(3), pp. 201-205. doi: 10.1016/s1059-1311(98)80036-8

Barkovich, A. J. 2005. *Pediatric neuroimaging*. Lippincott Williams & Wilkins.

Baron, I. S. 2018. *Neuropsychological evaluation of the child: Domains, methods, & case studies*. Oxford University Press.

Baskin, H. J., Jr. 2008. The pathogenesis and imaging of the tuberous sclerosis complex. *Pediatr Radiol* 38(9), pp. 936-952. doi: 10.1007/s00247-008-0832-y

Basser, P. J. et al. 1994a. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B* 103(3), pp. 247-254. doi: 10.1006/jmrb.1994.1037

Basser, P. J. et al. 1994b. MR diffusion tensor spectroscopy and imaging. *Biophys J* 66(1), pp. 259-267. doi: 10.1016/S0006-3495(94)80775-1

Basser, P. J. et al. 2000. In vivo fiber tractography using DT-MRI data. *Magn Reson Med* 44(4), pp. 625-632. doi: 10.1002/1522-2594(200010)44:4<625::aid-mrm17>3.0.co;2-o

Bate, A. J. et al. 2001. Performance on the Test of Everyday Attention and standard tests of attention following severe traumatic brain injury. *Clin Neuropsychol* 15(3), pp. 405-422. doi: 10.1076/clin.15.3.405.10279

Baumer, F. M. et al. 2018. Corpus Callosum White Matter Diffusivity Reflects Cumulative Neurological Comorbidity in Tuberous Sclerosis Complex. *Cerebral cortex (New York, N.Y. : 1991)* 28(10), pp. 3665-3672. doi: 10.1093/cercor/bhx247

Baumer, F. M. et al. 2015. Longitudinal changes in diffusion properties in white matter pathways of children with tuberous sclerosis complex. *Pediatr Neurol* 52(6), pp. 615-623. doi: 10.1016/j.pediatrneurol.2015.02.004

Bebin, M. et al. 2011. Everolimus in Subependymal Giant Cell Astrocytomas (SEGA) Associated with Tuberous Sclerosis Complex (TSC): Results of EXIST-1, a Double-Blind Placebo-controlled Phase III Trial. *Eur J Cancer* 47, pp. 4-5. doi: https://doi.org/10.1016/S0959-8049(11)70103-4

Beglinger, L. J. et al. 2005. Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology* 20(4), pp. 517-529.

Behrens, T. E. et al. 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat Neurosci* 6(7), pp. 750-757. doi: 10.1038/nn1075

Benton, A. et al. 1983. *Controlled oral word association test (COWAT)*.

Bezzola, L. et al. 2011. Training-induced neural plasticity in golf novices. *Journal of Neuroscience* 31(35), pp. 12444-12448.

Bialer, M. and Perucca, E. 2020. Does cannabidiol have antiseizure activity independent of its interactions with clobazam? An appraisal of the evidence from randomized controlled trials. *Epilepsia* 61(6), pp. 1082-1089. doi: https://doi.org/10.1111/epi.16542

Bissler, J. J. and Kingswood, J. C. 2004. Renal angiomyolipomata. *Kidney Int* 66(3), pp. 924-934. doi: 10.1111/j.1523-1755.2004.00838.x

Bissler, J. J. et al. 2017. Everolimus long-term use in patients with tuberous sclerosis complex: Four-year update of the EXIST-2 study. *PLoS One* 12(8), p. e0180939. doi: 10.1371/journal.pone.0180939

Bissler, J. J. et al. 2013. Everolimus for angiomyolipoma associated with tuberous sclerosis complex or sporadic lymphangioleiomyomatosis (EXIST-2): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet* 381(9869), pp. 817-824. doi: 10.1016/S0140-6736(12)61767-X

Bissler, J. J. et al. 2008. Sirolimus for angiomyolipoma in tuberous sclerosis complex or lymphangioleiomyomatosis. *N Engl J Med* 358(2), pp. 140-151. doi: 10.1056/NEJMoa063564

Bledsoe, I. O. et al. 2018. White matter abnormalities in the corpus callosum with cognitive impairment in Parkinson disease. *Neurology* 91(24), pp. e2244-e2255. doi: 10.1212/wnl.0000000000006646

Bolton, P. F. and Griffiths, P. D. 1997. Association of tuberous sclerosis of temporal lobes with autism and atypical autism. *Lancet* 349(9049), pp. 392-395. doi: 10.1016/S0140-6736(97)80012-8

Boronat, S. and Barber, I. 2018. Less common manifestations in TSC. *Am J Med Genet C Semin Med Genet* 178(3), pp. 348-354. doi: 10.1002/ajmg.c.31648

Bosi, G. et al. 1996. The natural history of cardiac rhabdomyoma with and without tuberous sclerosis. *Acta Paediatr* 85(8), pp. 928-931.

Bourneville, D. 1899. Idiotie symptomatique de la sclérose tubéreuse ou hypertrophique. *Le Progrès Médical (Paris)(série III)* 10, pp. 241-248.

Braffman, B. H. et al. 1992a. MR imaging of tuberous sclerosis: pathogenesis of this phakomatosis, use of gadopentetate dimeglumine, and literature review. *Radiology* 183(1), pp. 227-238.

Braffman, B. H. et al. 1992b. MR imaging of tuberous sclerosis: pathogenesis of this phakomatosis, use of gadopentetate dimeglumine, and literature review. *Radiology* 183(1), pp. 227-238. doi: 10.1148/radiology.183.1.1549677

Brant, R. 1990. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* 46(4), pp. 1171-1178.

Brearly, T. W. et al. 2017. Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis. *Neuropsychol Rev* 27(2), pp. 174-186. doi: 10.1007/s11065-017-9349-1

Brook-Carter, P. T. et al. 1994. Deletion of the TSC2 and PKD1 genes associated with severe infantile polycystic kidney disease--a contiguous gene syndrome. *Nat Genet* 8(4), pp. 328-332. doi: 10.1038/ng1294-328

Brugarolas, J. et al. 2004. Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. *Genes Dev* 18(23), pp. 2893-2904. doi: 10.1101/gad.1256804

Buccoliero, A. M. et al. 2009. Subependymal giant cell astrocytoma (SEGA): Is it an astrocytoma? Morphological, immunohistochemical and ultrastructural study. *Neuropathology* 29(1), pp. 25-30. doi: 10.1111/j.1440-1789.2008.00934.x

Buchsbaum, M. S. et al. 1998. MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. *Neuroreport* 9(3), pp. 425-430. doi: 10.1097/00001756-199802160-00013

Burzynska, A. Z. et al. 2011. Microstructure of frontoparietal connections predicts cortical responsivity and working memory performance. *Cereb Cortex* 21(10), pp. 2261-2271. doi: 10.1093/cercor/bhq293

Busatto, G. F. et al. 2008. Voxel-based morphometry in Alzheimer's disease. *Expert Rev Neurother* 8(11), pp. 1691-1702. doi: 10.1586/14737175.8.11.1691

Caban, C. et al. 2016. Genetics of tuberous sclerosis complex: implications for clinical practice. *The application of clinical genetics* 10, pp. 1-8. doi: 10.2147/TACG.S90262

Cabrera-López, C. et al. 2012. Assessing the effectiveness of rapamycin on angiomyolipoma in tuberous sclerosis: a two years trial. *Orphanet J Rare Dis* 7, p. 87. Available at: http://europepmc.org/abstract/MED/23140536

https://www.ncbi.nlm.nih.gov/pmc/articles/pmid/23140536/pdf/?tool=EBI

https://www.ncbi.nlm.nih.gov/pmc/articles/pmid/23140536/?tool=EBI

https://doi.org/10.1186/1750-1172-7-87

https://europepmc.org/articles/PMC3519505

https://europepmc.org/articles/PMC3519505?pdf=render [Accessed: 2012/11//]. doi: 10.1186/1750-1172-7-87

Caligiuri, M. E. et al. 2016. Integrity of the corpus callosum in patients with benign temporal lobe epilepsy. *Epilepsia* 57(4), pp. 590-596. doi: https://doi.org/10.1111/epi.13339

Camposano, S. E. et al. 2009. Distinct clinical characteristics of tuberous sclerosis complex patients with no mutation identified. *Ann Hum Genet* 73(2), pp. 141-146. doi: 10.1111/j.1469-1809.2008.00496.x

Capo-Chichi, J.-M. et al. 2013. Disruption of TBC1D7, a subunit of the TSC1-TSC2 protein complex, in intellectual disability and megalencephaly. *J Med Genet* 50(11), pp. 740-744. doi: 10.1136/jmedgenet-2013-101680

Carrozzino, D. et al. 2019. The prevalence of psychological distress in Parkinson's disease patients: The brief symptom inventory (BSI-18) versus the Hopkins symptom checklist (SCL-90-R). *Prog Neuropsychopharmacol Biol Psychiatry* 88, pp. 96-101. doi: 10.1016/j.pnpbp.2018.07.012

Castilho, R. M. et al. 2009. mTOR mediates Wnt-induced epidermal stem cell exhaustion and aging. *Cell Stem Cell* 5(3), pp. 279-289. doi: 10.1016/j.stem.2009.06.017

Catani, M. et al. 2002. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *Neuroimage* 17(1), pp. 77-94. doi: 10.1006/nimg.2002.1136

Chabriat, H. et al. 1999. Clinical severity in CADASIL related to ultrastructural damage in white matter: in vivo study with diffusion tensor MRI. *Stroke* 30(12), pp. 2637-2643. doi: 10.1161/01.str.30.12.2637

Chamberland, M. et al. 2014. Real-time multi-peak tractography for instantaneous connectivity display. *Frontiers in neuroinformatics* 8, p. 59.

Chan, R. C. 2000. Attentional deficits in patients with closed head injury: a further study to the discriminative validity of the test of everyday attention. *Brain Inj* 14(3), pp. 227-236. doi: 10.1080/026990500120709

Chen, H. C. et al. 2013. Test of Everyday Attention in patients with chronic stroke: test-retest reliability and practice effects. *Brain Inj* 27(10), pp. 1148-1154. doi: 10.3109/02699052.2013.775483

Chesnut, S. R. et al. 2017. A meta-analysis of the social communication questionnaire: Screening for autism spectrum disorder. *Autism* 21(8), pp. 920-928. doi: 10.1177/1362361316660065

Chuang, P. and Langone, A. J. 2007. Clobetasol Ameliorates Aphthous Ulceration in Renal Transplant Patients on Sirolimus. *American Journal of Transplantation* 7(3), pp. 714-717. doi: https://doi.org/10.1111/j.1600-6143.2006.01678.x

Clarke, M. J. et al. 2006. Imaging characteristics and growth of subependymal giant cell astrocytomas. *Neurosurg Focus* 20(1), p. E5. doi: 10.3171/foc.2006.20.1.6

Coe, R. 2002. It's the effect size, stupid: What effect size is and why it is important.

Colledge, V. L. and Solly, J. 2012. The rare disease challenge and how to promote a productive rare disease community: Case study of Birt-Hogg-Dubé Symposia. *Orphanet J Rare Dis* 7(1), pp. 1-2.

Consortium, T. E. C. T. S. 1993. Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* 75(7), pp. 1305-1315.

Constantino, J. N., & Gruber, C. P. 2005. *Social Responsiveness Scale (SRS)*. Los Angeles, CA: Western Psychological Services.

Conturo, T. E. et al. 1999. Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci U S A* 96(18), pp. 10422-10427. doi: 10.1073/pnas.96.18.10422

Costello, L. C. et al. 2000. High frequency of pulmonary lymphangioleiomyomatosis in women with tuberous sclerosis complex. *Mayo Clin Proc* 75(6), pp. 591-594. doi: 10.4065/75.6.591

Cotman, C. W. and Berchtold, N. C. 2002. Exercise: a behavioral intervention to enhance brain health and plasticity. *Trends Neurosci* 25(6), pp. 295-301. doi: https://doi.org/10.1016/S0166-2236(02)02143-4

Coughlan, A. K. and Hollows, S. E. 1985. *The Adult Memory and Information Processing Battery (AMIPB): Test Manual*. A.K. Coughlin, Psychology Department, St James' Hospital.

Crino, P. B. 2011. mTOR: A pathogenic signaling pathway in developmental brain malformations. *Trends Mol Med* 17(12), pp. 734-742. doi: 10.1016/j.molmed.2011.07.008

Crino, P. B. 2015. mTOR signaling in epilepsy: insights from malformations of cortical development. *Cold Spring Harb Perspect Med* 5(4), doi: 10.1101/cshperspect.a022442

Crino, P. B. et al. 2006. The tuberous sclerosis complex. *N Engl J Med* 355(13), pp. 1345-1356. doi: 10.1056/NEJMra055323

Crowe, A. et al. 1999. Absorption and intestinal metabolism of SDZ-RAD and rapamycin in rats. *Drug Metabolism and Disposition* 27(5), pp. 627-632.

Curatolo, P. and Moavero, R. 2013. mTOR inhibitors as a new therapeutic option for epilepsy. *Expert Rev Neurother* 13(6), pp. 627-638. doi: 10.1586/ern.13.49

Curatolo, P. et al. 2015. Neurological and neuropsychiatric aspects of tuberous sclerosis complex. *Lancet Neurol* 14(7), pp. 733-745. doi: 10.1016/S1474-4422(15)00069-1

Curatolo, P. et al. 2001. Infantile spasms in tuberous sclerosis complex. *Brain Dev* 23(7), pp. 502-507.

Dabora, S. L. et al. 2001. Mutational analysis in a cohort of 224 tuberous sclerosis patients indicates increased severity of TSC2, compared with TSC1, disease in multiple organs. *Am J Hum Genet* 68(1), pp. 64-80. doi: 10.1086/316951

Davenport, M. H. et al. 2016. Pharmacotherapy for Fragile X Syndrome: Progress to Date. *Drugs* 76(4), pp. 431-445. doi: 10.1007/s40265-016-0542-y

Davies, D. M. 2011. *mTOR inhibition as a therapeutic strategy in tuberous sclerosis or sporadic lymphangioleiomyomatosis.* Cardiff University.

Davies, D. M. et al. 2011a. Sirolimus therapy for angiomyolipoma in tuberous sclerosis and sporadic lymphangioleiomyomatosis: a phase 2 trial. *Clinical cancer research : an official journal of the American Association for Cancer Research* 17(12), pp. 4071-4081. doi: https://dx.doi.org/10.1158/1078-0432.CCR-11-0445

Davies, D. M. et al. 2011b. Sirolimus therapy for angiomyolipoma in tuberous sclerosis and sporadic lymphangioleiomyomatosis: a phase 2 trial. *Clin Cancer Res* 17(12), pp. 4071-4081. doi: 10.1158/1078-0432.CCR-11-0445

Davies, D. M. et al. 2008. Sirolimus therapy in tuberous sclerosis or sporadic lymphangioleiomyomatosis. *N Engl J Med* 358(2), pp. 200-203. doi: 10.1056/NEJMc072500

Davies, M. et al. 2017. Management of everolimus-associated adverse events in patients with tuberous sclerosis complex: a practical guide. *Orphanet J Rare Dis* 12(1), p. 35. doi: 10.1186/s13023-017-0581-9

Day, S. et al. 2018. Recommendations for the design of small population clinical trials. *Orphanet J Rare Dis* 13(1), p. 195. doi: 10.1186/s13023-018-0931-2

de Vries, P. J. 2010. Targeted treatments for cognitive and neurodevelopmental disorders in tuberous sclerosis complex. *Neurotherapeutics* 7(3), pp. 275-282. doi: 10.1016/j.nurt.2010.05.001

de Vries, P. J. et al. 2009. Neuropsychological attention deficits in tuberous sclerosis complex (TSC). *American Journal of Medical Genetics Part A* 149A(3), pp. 387-395. doi: https://doi.org/10.1002/ajmg.a.32690

de Vries, P. J. and Howe, C. J. 2007. The tuberous sclerosis complex proteins--a GRIPP on cognition and neurodevelopment. *Trends Mol Med* 13(8), pp. 319-326. doi: 10.1016/j.molmed.2007.06.003

de Vries, P. J. et al. 2007. The psychopathologies of children and adolescents with tuberous sclerosis complex (TSC): a postal survey of UK families. *Eur Child Adolesc Psychiatry* 16(1), pp. 16-24. doi: 10.1007/s00787-006-0570-3

de Vries, P. J. et al. 2015. Tuberous sclerosis associated neuropsychiatric disorders (TAND) and the TAND Checklist. *Pediatr Neurol* 52(1), pp. 25-35. doi: 10.1016/j.pediatrneurol.2014.10.004

Derogatis, L. R. and Unger, R. 2010. Symptom Checklist-90-Revised.*The Corsini Encyclopedia of Psychology*. pp. 1-2.

Devinsky, O. et al. 1995. Development of the quality of life in epilepsy inventory. *Epilepsia* 36(11), pp. 1089-1104. doi: 10.1111/j.1528-1157.1995.tb00467.x

Devroede, G. et al. 1988. Colonic hamartomas in tuberous sclerosis. *Gastroenterology* 94(1), pp. 182-188.

DeYoung, M. P. et al. 2008. Hypoxia regulates TSC1/2-mTOR signaling and tumor suppression through REDD1-mediated 14-3-3 shuttling. *Genes Dev* 22(2), pp. 239-251. doi: 10.1101/gad.1617608

Di Nardo, A. et al. 2009. Tuberous Sclerosis Complex Activity Is Required to Control Neuronal Stress Responses in an mTOR-Dependent Manner. *The Journal of Neuroscience* 29(18), pp. 5926-5937. doi: 10.1523/jneurosci.0778-09.2009

Dibbens, L. M. et al. 2013. Mutations in DEPDC5 cause familial focal epilepsy with variable foci. *Nat Genet* 45(5), pp. 546-551. doi: 10.1038/ng.2599

Dineen, R. A. et al. 2012. Extra-hippocampal subcortical limbic involvement predicts episodic recall performance in multiple sclerosis. *PLoS One* 7(10), p. e44942. doi: 10.1371/journal.pone.0044942

DOHSC. 2013. The UK strategy for Rare Diseases.

Douaud, G. et al. 2011. DTI measures in crossing-fibre areas: Increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *Neuroimage* 55(3), pp. 880-890. doi: https://doi.org/10.1016/j.neuroimage.2010.12.008

Downes, J. J. et al. 1989. Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: evidence for a specific attentional dysfunction. *Neuropsychologia* 27(11-12), pp. 1329-1343. doi: 10.1016/0028-3932(89)90128-0

Draganski, B. et al. 2004. Changes in grey matter induced by training. *Nature* 427(6972), pp. 311-312. doi: 10.1038/427311a

Duvel, K. et al. 2010. Activation of a metabolic gene regulatory network downstream of mTOR complex 1. *Mol Cell* 39(2), pp. 171-183. doi: 10.1016/j.molcel.2010.06.022

Dworakowska, D. and Grossman, A. B. 2009. Are neuroendocrine tumours a feature of tuberous sclerosis? A systematic review. *Endocr Relat Cancer* 16(1), pp. 45-58. doi: 10.1677/ERC-08-0142

Easton, J. B. and Houghton, P. J. 2006. mTOR and cancer therapy. *Oncogene* 25(48), pp. 6436-6446. doi: 10.1038/sj.onc.1209886

Ehninger, D. et al. 2008a. Reversal of learning deficits in a Tsc2+/- mouse model of tuberous sclerosis. *Nat Med* 14(8), pp. 843-848. doi: 10.1038/nm1788

Ehninger, D. et al. 2008b. Reversal of learning deficits in a Tsc2+/- mouse model of tuberous sclerosis. *Nature medicine* 14(8), pp. 843-848. doi: https://dx.doi.org/10.1038/nm1788

Ehninger, D. and Silva, A. J. 2011. Rapamycin for treating Tuberous sclerosis and Autism spectrum disorders. *Trends Mol Med* 17(2), pp. 78-87. doi: 10.1016/j.molmed.2010.10.002

Eichler, F. S. et al. 2002. Proton MR spectroscopic and diffusion tensor brain MR imaging in X-linked adrenoleukodystrophy: initial experience. *Radiology* 225(1), pp. 245-252. doi: 10.1148/radiol.2251011040

Ekici, M. A. et al. 2011. Surgical timing of the subependymal giant cell astrocytoma (SEGA) with the patients of tuberous sclerosis complex. *Turk Neurosurg* 21(3), pp. 315-324. doi: 10.5137/1019-5149.jtn.4169-11.0

EMA. 2006. Committee for Medicinal Products for Human Use (CHMP): Guideline on Clinical Trials in Small Populations. London.

EMA. 2019. Orphan Medicines Figures 2000 -2019.

Fazio, R. et al. 2012. The original instructions for the Edinburgh Handedness Inventory are misunderstood by a majority of participants. *Laterality* 17(1), pp. 70-77. doi: 10.1080/1357650X.2010.532801

FDA. 2014. Guidance for Industry: Expedited Programs for Serious Conditions – Drugs and Biologics. Maryland.

FDA. 2019. Rare Diseases: Common Issues in Drug Development Guidance for Industry. .

Feldman, M. E. et al. 2009. Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS Biol* 7(2), p. e38. doi: 10.1371/journal.pbio.1000038

Feng, Z. et al. 2005. The coordinate regulation of the p53 and mTOR pathways in cells. *Proc Natl Acad Sci U S A* 102(23), pp. 8204-8209. doi: 10.1073/pnas.0502857102

Filippi, C. G. et al. 2001. Diffusion tensor imaging of patients with HIV and normal-appearing white matter on MR images of the brain. *AJNR Am J Neuroradiol* 22(2), pp. 277-283.

Fleming, T. R. 1982. One-sample multiple testing procedure for phase II clinical trials. *Biometrics* 38(1), pp. 143-151.

Franz, D. N. et al. 2013. Efficacy and safety of everolimus for subependymal giant cell astrocytomas associated with tuberous sclerosis complex (EXIST-1): a multicentre, randomised, placebo-controlled phase 3 trial. *Lancet* 381(9861), pp. 125-132. doi: 10.1016/S0140-6736(12)61134-9

Fray, P. J. and Robbins, T. W. 1996. CANTAB battery: proposed utility in neurotoxicology. *Neurotoxicol Teratol* 18(4), pp. 499-504. doi: 10.1016/0892-0362(96)00027-x

French, J. A. et al. 2016. Adjunctive everolimus therapy for treatment-resistant focal-onset seizures associated with tuberous sclerosis (EXIST-3): a phase 3, randomised, double-blind, placebo-controlled study. *Lancet* 388(10056), pp. 2153-2163. doi: 10.1016/S0140-6736(16)31419-2

Frias, M. A. et al. 2006. mSin1 is necessary for Akt/PKB phosphorylation, and its isoforms define three distinct mTORC2s. *Curr Biol* 16(18), pp. 1865-1870. doi: 10.1016/j.cub.2006.08.001

Froeling, M. et al. 2016. DTI Analysis Methods: Region of Interest Analysis. In: Van Hecke, W. et al. eds. *Diffusion Tensor Imaging: A Practical Handbook*. New York, NY: Springer New York, pp. 175-182.

Furth, R. and Cowper, A. 1956. Albert einstein: Investigations on the theory of brownian movement. New York: Dover.

Gagne, J. J. et al. 2014. Innovative research methods for studying treatments for rare diseases: methodological review. *BMJ : British Medical Journal* 349, p. g6802. doi: 10.1136/bmj.g6802

Galluzzi, P. et al. 2002. Hemimegalencephaly in tuberous sclerosis complex. *J Child Neurol* 17(9), pp. 677-680. doi: 10.1177/088307380201700905

Garcia-Martinez, J. M. et al. 2009. Ku-0063794 is a specific inhibitor of the mammalian target of rapamycin (mTOR). *Biochem J* 421(1), pp. 29-42. doi: 10.1042/BJ20090489

Gau, S. S. and Shang, C. Y. 2010a. Executive functions as endophenotypes in ADHD: evidence from the Cambridge Neuropsychological Test Battery (CANTAB). *J Child Psychol Psychiatry* 51(7), pp. 838-849. doi: 10.1111/j.1469-7610.2010.02215.x

Gau, S. S. and Shang, C. Y. 2010b. Improvement of executive functions in boys with attention deficit hyperactivity disorder: an open-label follow-up study with once-daily atomoxetine. *Int J Neuropsychopharmacol* 13(2), pp. 243-256. doi: 10.1017/S1461145709990836

Gomez, M. 1988. *Criteria for diagnosis. In: GomezMR, ed. Tuberous sclerosis*. 2nd ed. New York: Raven.

Goorden, S. M. et al. 2007. Cognitive deficits in Tsc1+/- mice in the absence of cerebral lesions and seizures. *Ann Neurol* 62(6), pp. 648-655. doi: 10.1002/ana.21317

Gould, S. R. 1991. Hamartomatous rectal polyps are common in tuberous sclerosis. *Ann N Y Acad Sci* 615, pp. 71-80. doi: 10.1111/j.1749-6632.1991.tb37749.x

Gould, S. R. et al. 1990. Rectal polyposis in tuberous sclerosis. *J Ment Defic Res* 34 (Pt 6), pp. 465-473. doi: 10.1111/j.1365-2788.1990.tb01558.x

Granader, Y. E. et al. 2010. The clinical utility of the Social Responsiveness Scale and Social Communication Questionnaire in tuberous sclerosis complex. *Epilepsy Behav* 18(3), pp. 262-266. doi: 10.1016/j.yebeh.2010.04.010

Guinee, D. et al. 1995. Multifocal micronodular pneumocyte hyperplasia: a distinctive pulmonary manifestation of tuberous sclerosis. *Mod Pathol* 8(9), pp. 902-906.

Gupta, S. et al. 2011. A framework for applying unfamiliar trial designs in studies of rare diseases. *Journal of Clinical Epidemiology* 64(10), pp. 1085-1094. doi: https://doi.org/10.1016/j.jclinepi.2010.12.019

Hanyu, H. et al. 1997. Increased water diffusion in cerebral white matter in Alzheimer's disease. *Gerontology* 43(6), pp. 343-351. doi: 10.1159/000213874

Hara, K. et al. 2002. Raptor, a Binding Partner of Target of Rapamycin (TOR), Mediates TOR Action. *Cell* 110(2), pp. 177-189. doi: https://doi.org/10.1016/S0092-8674(02)00833-4

Harrison, J. E. et al. 1999a. Cognitive deficits in normally intelligent patients with tuberous sclerosis. *Am J Med Genet* 88(6), pp. 642-646. doi: https://doi.org/10.1002/(SICI)1096-8628(19991215)88:6<642::AID-AJMG12>3.0.CO;2-O

Harrison, J. E. et al. 1999b. Cognitive deficits in normally intelligent patients with tuberous sclerosis. *American journal of medical genetics* 88(6), pp. 642-646.

Hartman, T. R. et al. 2009. The tuberous sclerosis proteins regulate formation of the primary cilium via a rapamycin-insensitive and polycystin 1-independent pathway. *Hum Mol Genet* 18(1), pp. 151-163. doi: 10.1093/hmg/ddn325

Heilbronner, R. L. et al. 2010. Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges of repeat test administrations in clinical and forensic contexts. *The Clinical Neuropsychologist* 24(8), pp. 1267-1278.

Henson, R. N. et al. 2016. The effects of hippocampal lesions on MRI measures of structural and functional connectivity. *Hippocampus* 26(11), pp. 1447-1463. doi: 10.1002/hipo.22621

Hentges, K. E. et al. 2001. FRAP/mTOR is required for proliferation and patterning during embryonic development in the mouse. *Proceedings of the National Academy of Sciences* 98(24), pp. 13796-13801.

Herholz, Sibylle C. and Zatorre, Robert J. 2012. Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron* 76(3), pp. 486-502. doi: https://doi.org/10.1016/j.neuron.2012.10.011

Hilgers, R.-D. et al. 2016. Directions for new developments on statistical design and analysis of small population group trials. *Orphanet J Rare Dis* 11(1), p. 78. doi: 10.1186/s13023-016-0464-5

Hizawa, K. et al. 1994. Gastrointestinal involvement in tuberous sclerosis. Two case reports. *J Clin Gastroenterol* 19(1), pp. 46-49. doi: 10.1097/00004836-199407000-00012

Hoeffer, C. A. and Klann, E. 2010. mTOR signaling: at the crossroads of plasticity, memory and disease. *Trends Neurosci* 33(2), pp. 67-75.

Honea, R. et al. 2005. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *Am J Psychiatry* 162(12), pp. 2233-2245. doi: 10.1176/appi.ajp.162.12.2233

Huang, J. and Manning, B. D. 2008. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. *Biochem J* 412(2), pp. 179-190. doi: 10.1042/BJ20080281

Huang, J. and Manning, B. D. 2009. A complex interplay between Akt, TSC2 and the two mTOR complexes. *Biochem Soc Trans* 37(Pt 1), pp. 217-222. doi: 10.1042/BST0370217

Hunt, A. 1983. Tuberous sclerosis: a survey of 97 cases. II: Physical findings. *Dev Med Child Neurol* 25(3), pp. 350-352.

Huntsman, R. J. et al. 2006. Tuberous sclerosis with open lipped schizencephaly. *Pediatr Neurol* 34(3), pp. 231-234. doi: 10.1016/j.pediatrneurol.2005.08.012

IRDiRC. 2020. International Rare Disease Research Consortium: Policies and Guidelines.

Jacinto, E. et al. 2006. SIN1/MIP1 maintains rictor-mTOR complex integrity and regulates Akt phosphorylation and substrate specificity. *Cell* 127(1), pp. 125-137. doi: 10.1016/j.cell.2006.08.033

Jacinto, E. et al. 2004. Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. *Nat Cell Biol* 6(11), pp. 1122-1128. doi: 10.1038/ncb1183

Jambaqué, I. et al. 1991. NEUROPSYCHOLOGICAL ASPECTS OF TUBEROUS SCLEROSIS IN RELATION TO EPILEPSY AND MRI FINDINGS. *Developmental Medicine & Child Neurology* 33(8), pp. 698-705. doi: https://doi.org/10.1111/j.1469-8749.1991.tb14947.x

Janowsky, J. S. et al. 1989. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behav Neurosci* 103(3), pp. 548-560. doi: 10.1037//0735-7044.103.3.548

Jansen, F. E. et al. 2008. Cognitive impairment in tuberous sclerosis complex is a multifactorial condition. *Neurology* 70(12), pp. 916-923. doi: 10.1212/01.wnl.0000280579.04974.c0

Jaworski, J. et al. 2005. Control of dendritic arborization by the phosphoinositide-3′-kinase–Akt–mammalian target of rapamycin pathway. *Journal of Neuroscience* 25(49), pp. 11300-11312.

Joinson, C. et al. 2003. Learning disability and epilepsy in an epidemiological sample of individuals with tuberous sclerosis complex. *Psychol Med* 33(2), pp. 335-344.

Jones, A. C. et al. 1999a. Comprehensive mutation analysis of TSC1 and TSC2-and phenotypic correlations in 150 families with tuberous sclerosis. *Am J Hum Genet* 64(5), pp. 1305-1315. doi: 10.1086/302381

Jones, D. K. et al. 2006. Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. *Hum Brain Mapp* 27(3), pp. 230-238. doi: 10.1002/hbm.20179

Jones, D. K. et al. 2005a. A diffusion tensor magnetic resonance imaging study of frontal cortex connections in very-late-onset schizophrenia-like psychosis. *Am J Geriatr Psychiatry* 13(12), pp. 1092-1099. doi: 10.1176/appi.ajgp.13.12.1092

Jones, D. K. et al. 1999b. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. *Magn Reson Med* 42(1), pp. 37-41. doi: 10.1002/(sici)1522-2594(199907)42:1<37::aid-mrm7>3.0.co;2-o

Jones, D. K. et al. 2005b. The effect of filter size on VBM analyses of DT-MRI data. *Neuroimage* 26(2), pp. 546-554. doi: 10.1016/j.neuroimage.2005.02.013

Juh, R. et al. 2012. SU-E-I-73: Gray Matter Atrophy and White Matter Tract Abnormalities by Voxel Wise Correlation Analysis in Patients with Alzheimer's Disease. *Med Phys* 39(6Part5), p. 3641. doi: 10.1118/1.4734790

Kang, Y. J. et al. 2011. The TSC1 and TSC2 tumor suppressors are required for proper ER stress response and protect cells from ER stress-induced apoptosis. *Cell Death & Differentiation* 18(1), pp. 133-144. doi: 10.1038/cdd.2010.82

Kaper, F. et al. 2006. Mutations in the PI3K/PTEN/TSC2 pathway contribute to mammalian target of rapamycin activity and increased translation under hypoxic conditions. *Cancer Res* 66(3), pp. 1561-1569. doi: 10.1158/0008-5472.can-05-3375

Kaptchuk, T. J. and Miller, F. G. 2015. Placebo effects in medicine. *N Engl J Med* 373(1), pp. 8-9.

Karlsgodt, K. H. et al. 2008. Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biol Psychiatry* 63(5), pp. 512-518. doi: 10.1016/j.biopsych.2007.06.017

Kelleher III, R. J. et al. 2004. Translational regulatory mechanisms in persistent forms of synaptic plasticity. *Neuron* 44(1), pp. 59-73.

Kelley, K. and Preacher, K. J. 2012. On effect size. *Psychological methods* 17(2), p. 137.

Kemenoff, L. A. et al. 2002. Frontal Lobe. In: Ramachandran, V.S. ed. *Encyclopedia of the Human Brain*. New York: Academic Press, pp. 317-325.

Kempf, L. et al. 2018. Challenges of developing and conducting clinical trials in rare disorders. *Am J Med Genet A* 176(4), pp. 773-783. doi: 10.1002/ajmg.a.38413

Kim, D. H. et al. 2002. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* 110(2), pp. 163-175. doi: 10.1016/s0092-8674(02)00808-5

Kim, D. H. et al. 2003. GbetaL, a positive regulator of the rapamycin-sensitive pathway required for the nutrient-sensitive interaction between raptor and mTOR. *Mol Cell* 11(4), pp. 895-904. doi: 10.1016/s1097-2765(03)00114-x

Kim, E. K. et al. 2008a. Linker region of Akt1/protein kinase Balpha mediates platelet-derived growth factor-induced translocation and cell migration. *Cell Signal* 20(11), pp. 2030-2037. doi: 10.1016/j.cellsig.2008.07.012

Kim, H. R. et al. 2008b. Bosellia serrata-induced apoptosis is related with ER stress and calcium release. *Genes Nutr* 2(4), pp. 371-374. doi: 10.1007/s12263-007-0072-z

Kim, S. G. et al. 2013. Nutrient regulation of the mTOR complex 1 signaling pathway. *Mol Cells* 35(6), pp. 463-473. doi: 10.1007/s10059-013-0138-2

Kingswood, J. C. et al. 2017. TuberOus SClerosis registry to increase disease Awareness (TOSCA) – baseline data on 2093 patients. *Orphanet J Rare Dis* 12(1), p. 2. doi: 10.1186/s13023-016-0553-5

Kirchner, G. I. et al. 2004. Clinical Pharmacokinetics of Everolimus. *Clinical Pharmacokinetics* 43(2), pp. 83-95. doi: 10.2165/00003088-200443020-00002

Koenig, M. K. et al. 2018. Efficacy and Safety of Topical Rapamycin in Patients With Facial Angiofibromas Secondary to Tuberous Sclerosis Complex: The TREATMENT Randomized Clinical Trial. *JAMA Dermatol* 154(7), pp. 773-780. doi: 10.1001/jamadermatol.2018.0464

Kovats, D. et al. 2017. Factors affecting quality of life in Hungarian adults with epilepsy: A comparison of four psychiatric instruments. *Epilepsy Behav* 74, pp. 45-58. doi: 10.1016/j.yebeh.2017.04.035

Kozlowski, P. et al. 2007. Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations. *Hum Genet* 121(3-4), pp. 389-400. doi: 10.1007/s00439-006-0308-9

Krishnan, M. L. et al. 2010. Diffusion features of white matter in tuberous sclerosis with tractography. *Pediatr Neurol* 42(2), pp. 101-106. doi: 10.1016/j.pediatrneurol.2009.08.001

Krueger, D. A. et al. 2010a. Everolimus for subependymal giant-cell astrocytomas in tuberous sclerosis. *N Engl J Med* 363(19), pp. 1801-1811. doi: 10.1056/NEJMoa1001671

Krueger, D. A. et al. 2010b. Everolimus for Subependymal Giant-Cell Astrocytomas in Tuberous Sclerosis. *New England Journal of Medicine* 363(19), pp. 1801-1811. doi: 10.1056/NEJMoa1001671

Krueger, D. A. et al. 2013. Tuberous sclerosis complex surveillance and management: recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference. *Pediatr Neurol* 49(4), pp. 255-265. doi: 10.1016/j.pediatrneurol.2013.08.002

Krueger, D. A. et al. 2017. Everolimus for treatment of tuberous sclerosis complex-associated neuropsychiatric disorders. *Ann Clin Transl Neurol* 4(12), pp. 877-887. doi: 10.1002/acn3.494

Kubicki, M. et al. 2003. Cingulate fasciculus integrity disruption in schizophrenia: a magnetic resonance diffusion tensor imaging study. *Biol Psychiatry* 54(11), pp. 1171-1180. doi: 10.1016/s0006-3223(03)00419-0

Kumar, V. et al. 2005. Regulation of dendritic morphogenesis by Ras–PI3K–Akt–mTOR and Ras–MAPK signaling pathways. *Journal of Neuroscience* 25(49), pp. 11288-11299.

Kurahashi, H. and Hirose, S. 1993. Autosomal Dominant Nocturnal Frontal Lobe Epilepsy. In: Adam, M.P. et al. eds. *GeneReviews(®)*. Seattle (WA): University of Washington, Seattle

Lacher, M. D. et al. 2010. Rheb activates AMPK and reduces p27Kip1 levels in Tsc2-null cells via mTORC1-independent mechanisms: implications for cell proliferation and tumorigenesis. *Oncogene* 29(50), pp. 6543-6556. doi: 10.1038/onc.2010.393

Land, S. C. and Tee, A. R. 2007. Hypoxia-inducible factor 1alpha is regulated by the mammalian target of rapamycin (mTOR) via an mTOR signaling motif. *J Biol Chem* 282(28), pp. 20534-20543. doi: 10.1074/jbc.M611782200

Laplante, M. and Sabatini, D. M. 2013. Regulation of mTORC1 and its impact on gene expression at a glance. *J Cell Sci* 126(Pt 8), pp. 1713-1719. doi: 10.1242/jcs.125773

Lazorchak, A. S. and Su, B. 2011. Perspectives on the role of mTORC2 in B lymphocyte development, immunity and tumorigenesis. *Protein Cell* 2(7), pp. 523-530. doi: 10.1007/s13238-011-1077-3

Le Bihan, D. 2003. Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci* 4(6), pp. 469-480. doi: 10.1038/nrn1119

Le Bihan, D. et al. 2001. Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging* 13(4), pp. 534-546. doi: 10.1002/jmri.1076

Le Bihan, D. et al. 1991. Imaging of diffusion and microcirculation with gradient sensitization: Design, strategy, and significance. *Journal of magnetic resonance imaging* 1(1), pp. 7-28. doi: 10.1002/jmri.1880010103

Lehericy, S. et al. 2004. Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann Neurol* 55(4), pp. 522-529. doi: 10.1002/ana.20030

Lewis, J. C. et al. 2004. Genotype and psychological phenotype in tuberous sclerosis. *J Med Genet* 41(3), pp. 203-207.

Lilford, R. J. et al. 1995. Clinical trials and rare diseases: a way out of a conundrum. *BMJ* 311(7020), pp. 1621-1625.

Lim, K. O. and Helpern, J. A. 2002. Neuropsychiatric applications of DTI - a review. *NMR Biomed* 15(7-8), pp. 587-593. doi: 10.1002/nbm.789

Liu, L. et al. 2006. Hypoxia-Induced Energy Stress Regulates mRNA Translation and Cell Growth. *Molecular Cell* 21(4), pp. 521-531. doi: https://doi.org/10.1016/j.molcel.2006.01.010

Loewith, R. et al. 2002. Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Mol Cell* 10(3), pp. 457-468. doi: 10.1016/s1097-2765(02)00636-6

Luat, A. F. et al. 2007. Neuroimaging in tuberous sclerosis complex. *Curr Opin Neurol* 20(2), pp. 142-150. doi: 10.1097/WCO.0b013e3280895d93

Luciana, M. 2003. Practitioner review: computerized assessment of neuropsychological function in children: clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB). *J Child Psychol Psychiatry* 44(5), pp. 649-663. doi: 10.1111/1469-7610.00152

Lyczkowski, D. A. et al. 2007. Intrafamilial phenotypic variability in tuberous sclerosis complex. *Journal of Child Neurology* 22(12), pp. 1348-1355. doi: https://dx.doi.org/10.1177/0883073807307093

Ma, L. et al. 2007. Identification of S664 TSC2 Phosphorylation as a Marker for Extracellular Signal-Regulated Kinase–Mediated mTOR Activation in Tuberous Sclerosis and Human Cancer. *Cancer Research* 67(15), pp. 7106-7112. doi: 10.1158/0008-5472.can-06-4798

Makki, M. I. et al. 2007a. Characteristics of abnormal diffusivity in normal-appearing white matter investigated with diffusion tensor MR imaging in tuberous sclerosis complex. *AJNR. American journal of neuroradiology* 28(9), pp. 1662-1667. doi: ajnr.A0642 [pii]

10.3174/ajnr.A0642

Makki, M. I. et al. 2007b. Characteristics of abnormal diffusivity in normal-appearing white matter investigated with diffusion tensor MR imaging in tuberous sclerosis complex. *AJNR Am J Neuroradiol* 28(9), pp. 1662-1667. doi: 10.3174/ajnr.A0642

Manning, B. D. et al. 2002. Identification of the tuberous sclerosis complex-2 tumor suppressor gene product tuberin as a target of the phosphoinositide 3-kinase/akt pathway. *Molecular Cell* 10(1), pp. 151-162.

Martin, S. J. et al. 2000. Synaptic Plasticity and Memory: An Evaluation of the Hypothesis. *Annual Review of Neuroscience* 23(1), pp. 649-711. doi: 10.1146/annurev.neuro.23.1.649

McCormack, F. X. et al. 2011. Efficacy and safety of sirolimus in lymphangioleiomyomatosis. *New England Journal of Medicine* 364(17), pp. 1595-1606.

McCullagh, P. 1980. Regression Models for Ordinal Data. *Journal of the Royal Statistical Society: Series B (Methodological)* 42(2), pp. 109-127. doi: https://doi.org/10.1111/j.2517-6161.1980.tb01109.x

McKone, E. F. et al. 2014. Long-term safety and efficacy of ivacaftor in patients with cystic fibrosis who have the Gly551Asp-CFTR mutation: a phase 3, open-label extension study (PERSIST). *The Lancet Respiratory Medicine* 2(11), pp. 902-910. doi: https://doi.org/10.1016/S2213-2600(14)70218-8

McNeish, D. 2016. On Using Bayesian Methods to Address Small Sample Problems. *Structural Equation Modeling: A Multidisciplinary Journal* 23(5), pp. 750-773. doi: 10.1080/10705511.2016.1186549

Meikle, L. et al. 2008. Response of a Neuronal Model of Tuberous Sclerosis to Mammalian Target of Rapamycin (mTOR) Inhibitors: Effects on mTORC1 and Akt Signaling Lead to Improved Survival and Function. *The Journal of Neuroscience* 28(21), pp. 5422-5432. doi: 10.1523/jneurosci.0955-08.2008

Metzler-Baddeley, C. et al. 2012. How and how not to correct for CSF-contamination in diffusion MRI. *Neuroimage* 59(2), pp. 1394-1403. doi: https://doi.org/10.1016/j.neuroimage.2011.08.043

MHRA. 2021. Orphan registered medicinal products

Mielke, M. M. et al. 2012. Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimers Dement* 8(2), pp. 105-113. doi: 10.1016/j.jalz.2011.05.2416

Miller, F. D. and Kaplan, D. R. 2003. Signaling mechanisms underlying dendrite formation. *Current opinion in neurobiology* 13(3), pp. 391-398.

Milner, B. 1971. Interhemispheric differences in the localization of psychological processes in man. *Br Med Bull* 27(3), pp. 272-277. doi: 10.1093/oxfordjournals.bmb.a070866

Mirzaa, G. M. et al. 2012. Megalencephaly-capillary malformation (MCAP) and megalencephaly-polydactyly-polymicrogyria-hydrocephalus (MPPH) syndromes: Two closely related disorders of brain overgrowth and abnormal brain and body morphogenesis. *American Journal of Medical Genetics, Part A* 158 A(2), pp. 269-291. doi: 10.1002/ajmg.a.34402

Mirzaalian, H. et al. eds. 2015. *Harmonizing Diffusion MRI Data Across Multiple Sites and Scanners*. Cham, Springer International Publishing.

Mirzaalian, H. et al. 2016. Inter-site and inter-scanner diffusion MRI data harmonization. *Neuroimage* 135, pp. 311-323. doi: 10.1016/j.neuroimage.2016.04.041

Mizuguchi, M. and Takashima, S. 2001. Neuropathology of tuberous sclerosis. *Brain Dev* 23(7), pp. 508-515. doi: 10.1016/s0387-7604(01)00304-7

Mori, S. et al. 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol* 45(2), pp. 265-269. doi: 10.1002/1531-8249(199902)45:2<265::aid-ana21>3.0.co;2-3

Moseley, M. E. et al. 1990. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology* 176(2), pp. 439-445. doi: 10.1148/radiology.176.2.2367658

Mukuria, C. et al. 2017. Sensitivity and responsiveness of the EQ-5D-3L in patients with uncontrolled focal seizures: an analysis of Phase III trials of adjunctive brivaracetam. *Qual Life Res* 26(3), pp. 749-759. doi: 10.1007/s11136-016-1483-3

NCI. 2010. Common Terminology Criteriafor Adverse Events (CTCAE) https://www.eortc.be/services/doc/ctc/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf.

Nelson, H. E. and O'Connell, A. 1978. Dementia: The Estimation of Premorbid Intelligence Levels Using the New Adult Reading Test. *Cortex* 14(2), pp. 234-244. doi: https://doi.org/10.1016/S0010-9452(78)80049-5

Ni, H. C. et al. 2013. A head-to-head randomized clinical trial of methylphenidate and atomoxetine treatment for executive function in adults with attention-deficit hyperactivity disorder. *Int J Neuropsychopharmacol* 16(9), pp. 1959-1973. doi: 10.1017/S1461145713000357

NICE. 2019. Managed Access Agreement –nusinersen 5q SMA.

Ning, L. et al. 2020. Cross-scanner and cross-protocol multi-shell diffusion MRI data harmonization: Algorithms and results. *Neuroimage* 221, p. 117128. doi: 10.1016/j.neuroimage.2020.117128

Nobukuni, T. et al. 2005. Amino acids mediate mTOR/raptor signaling through activation of class 3 phosphatidylinositol 3OH-kinase. *Proc Natl Acad Sci U S A* 102(40), pp. 14238-14243. doi: 10.1073/pnas.0506925102

NORD. 2016. NORD-FDA Natural History Study Project.

Northrup, H. et al. 2013. Tuberous sclerosis complex diagnostic criteria update: recommendations of the 2012 Iinternational Tuberous Sclerosis Complex Consensus Conference. *Pediatr Neurol* 49(4), pp. 243-254. doi: 10.1016/j.pediatrneurol.2013.08.001

Novartis. 2019. Votubia (everolimus): Summary of Product Characteristics. *Summary of Product Characteristics*,

O'Callaghan, F. J. et al. 2004. The relation of infantile spasms, tubers, and intelligence in tuberous sclerosis complex. *Arch Dis Child* 89(6), pp. 530-533. doi: 10.1136/adc.2003.026815

Odgaard‐Jensen, J. et al. 2011. Randomisation to protect against selection bias in healthcare trials. *Cochrane database of systematic reviews* (4),

Oldfield, R. C. 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9(1), pp. 97-113. doi: 10.1016/0028-3932(71)90067-4

Oram, J. et al. 2005. Executive control of working memory in schizophrenia. *Psychiatry Res* 135(2), pp. 81-90. doi: 10.1016/j.psychres.2005.03.002

Osborne, L. R. 2010. Caveat mTOR: aberrant signaling disrupts corticogenesis. *The Journal of Clinical Investigation* 120(5), pp. 1392-1395. doi: 10.1172/JCI43030

Oshiro, N. et al. 2004. Dissociation of raptor from mTOR is a mechanism of rapamycin-induced inhibition of mTOR function. *Genes Cells* 9(4), pp. 359-366. doi: 10.1111/j.1356-9597.2004.00727.x

Overwater, I. E. et al. 2019. A randomized controlled trial with everolimus for IQ and autism in tuberous sclerosis complex. *Neurology* 93(2), pp. e200-e209. doi: 10.1212/WNL.0000000000007749

Owen, A. M. et al. 1990. Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia* 28(10), pp. 1021-1034. doi: 10.1016/0028-3932(90)90137-d

Owen, A. M. et al. 1997. Spatial and non-spatial working memory at different stages of Parkinson's disease. *Neuropsychologia* 35(4), pp. 519-532. doi: 10.1016/s0028-3932(96)00101-7

Owen, A. M. et al. 1991. Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia* 29(10), pp. 993-1006. doi: 10.1016/0028-3932(91)90063-e

Ozcan, U. et al. 2008. Loss of the Tuberous Sclerosis Complex Tumor Suppressors Triggers the Unfolded Protein Response to Regulate Insulin Signaling and Apoptosis. *Molecular Cell* 29(5), pp. 541-551. doi: https://doi.org/10.1016/j.molcel.2007.12.023

Parkhitko, A. A. et al. 2014. Kinase mTOR: regulation and role in maintenance of cellular homeostasis, tumor development, and aging. *Biochemistry (Mosc)* 79(2), pp. 88-101. doi: 10.1134/S0006297914020023

Parloff, M. B. et al. 1954. Comfort, effectiveness, and self-awareness as criteria of improvement in psychotherapy. *Am J Psychiatry* 111(5), pp. 343-352. doi: 10.1176/ajp.111.5.343

Pearce, L. R. et al. 2007. Identification of Protor as a novel Rictor-binding component of mTOR complex-2. *Biochem J* 405(3), pp. 513-522. doi: 10.1042/BJ20070540

Peng, S. S. et al. 2004. Cerebral diffusion tensor images in children with tuberous sclerosis: a preliminary report. *Pediatric radiology* 34(5), pp. 387-392. doi: 10.1007/s00247-004-1162-3

Peters, J. M. et al. 2019. Longitudinal Effects of Everolimus on White Matter Diffusion in Tuberous Sclerosis Complex. *Pediatr Neurol* 90, pp. 24-30. doi: 10.1016/j.pediatrneurol.2018.10.005

Peters, J. M. et al. 2012. Loss of white matter microstructural integrity is associated with adverse neurological outcome in tuberous sclerosis complex. *Acad Radiol* 19(1), pp. 17-25. doi: 10.1016/j.acra.2011.08.016

Petrides, M. and Milner, B. 1982. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* 20(3), pp. 249-262. doi: 10.1016/0028-3932(82)90100-2

Pierpaoli, C. and Basser, P. J. 1996a. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med* 36(6), pp. 893-906. doi: 10.1002/mrm.1910360612

Pierpaoli, C. and Basser, P. J. 1996b. Toward a quantitative assessment of diffusion anisotropy. *Magnetic resonance in Medicine* 36(6), pp. 893-906.

Poupon, C. et al. 2000. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *Neuroimage* 12(2), pp. 184-195. doi: 10.1006/nimg.2000.0607

Prather, P. and de Vries, P. J. 2004. Behavioral and cognitive aspects of tuberous sclerosis complex. *J Child Neurol* 19(9), pp. 666-674. doi: 10.1177/08830738040190090601

Prohl, A. K. et al. 2019. Reproducibility of Structural and Diffusion Tensor Imaging in the TACERN Multi-Center Study. *Front Integr Neurosci* 13, p. 24. doi: 10.3389/fnint.2019.00024

Psychological, C. and PsychCorp. 1999. *Wechsler Abbreviated Scale of Intelligence (WASItm) Complete Kit*. San Antonio: San Antonio : Psychological Corporation.

R Core Team. 2020. A language and environment for statistical computing. Vienna, Austria.: R Foundation for Statistical Computing.

Rae, C. L. et al. 2012. White matter pathology in Parkinson's disease: The effect of imaging protocol differences and relevance to executive function. *Neuroimage* 62(3), pp. 1675-1684. doi: https://doi.org/10.1016/j.neuroimage.2012.06.012

Rahman, S. et al. 1999. Comparative cognitive neuropsychological studies of frontal lobe function: implications for therapeutic strategies in frontal variant frontotemporal dementia. *Dement Geriatr Cogn Disord* 10 Suppl 1, pp. 15-28. doi: 10.1159/000051207

Raj, N. et al. 2019. Postmarket surveillance: a review on key aspects and measures on the effective functioning in the context of the United Kingdom and Canada. *Therapeutic advances in drug safety* 10, pp. 2042098619865413-2042098619865413. doi: 10.1177/2042098619865413

Randell, E. et al. 2016. The use of everolimus in the treatment of neurocognitive problems in tuberous sclerosis (TRON): study protocol for a randomised controlled trial. *Trials* 17, p. 398. doi: 10.1186/s13063-016-1446-6

Raznahan, A. et al. 2006. Psychopathology in tuberous sclerosis: an overview and findings in a population-based sample of adults with tuberous sclerosis. *J Intellect Disabil Res* 50(Pt 8), pp. 561-569. doi: 10.1111/j.1365-2788.2006.00828.x

Roach, E. S. et al. 1992. Report of the Diagnostic Criteria Committee of the National Tuberous Sclerosis Association. *J Child Neurol* 7(2), pp. 221-224. doi: 10.1177/088307389200700219

Roach, E. S. and Sparagana, S. P. 2004. Diagnosis of tuberous sclerosis complex. *J Child Neurol* 19(9), pp. 643-649. doi: 10.1177/08830738040190090301

Robbins, T. W. et al. 1994. Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia* 5(5), pp. 266-281. doi: 10.1159/000106735

Robertson, I. H. et al. 1996. The structure of normal human attention: The Test of Everyday Attention. *J Int Neuropsychol Soc* 2(6), pp. 525-534. doi: 10.1017/s1355617700001697

Rose, D. 2003. The NS-SEC Explained. *A Researcher's Guide to the National Statistics Socio-economic Classification*. London: SAGE Publications, Ltd.

Rosner, M. et al. 2008. The mTOR pathway and its role in human genetic diseases. *Mutat Res* 659(3), pp. 284-292. doi: 10.1016/j.mrrev.2008.06.001

Roth, J. et al. 2013. Subependymal Giant Cell Astrocytoma: Diagnosis, Screening, and Treatment. Recommendations From the International Tuberous Sclerosis Complex Consensus Conference 2012. *Pediatr Neurol* 49(6), pp. 439-444. doi: https://doi.org/10.1016/j.pediatrneurol.2013.08.017

Roux, P. P. et al. 2004. Tumor-promoting phorbol esters and activated Ras inactivate the tuberous sclerosis tumor suppressor complex via p90 ribosomal S6 kinase. *Proc Natl Acad Sci U S A* 101(37), pp. 13489-13494. doi: 10.1073/pnas.0405659101

Rowley, S. A. et al. 2001. Ophthalmic manifestations of tuberous sclerosis: a population based study. *Br J Ophthalmol* 85(4), pp. 420-423. doi: 10.1136/bjo.85.4.420

Rugg-Gunn, F. J. et al. 2001. Diffusion tensor imaging of cryptogenic and acquired partial epilepsies. *Brain* 124(Pt 3), pp. 627-636. doi: 10.1093/brain/124.3.627

Rutter M, B. A., Lord C. 2003. *Social Communication Questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.

Ryan, J. J. et al. 2003. Exploratory factor analysis of the Wechsler Abbreviated Scale of Intelligence (WASI) in adult standardization and clinical samples. *Appl Neuropsychol* 10(4), pp. 252-256. doi: 10.1207/s15324826an1004_8

Sahakian, B. et al. 1989. The effects of nicotine on attention, information processing, and short-term memory in patients with dementia of the Alzheimer type. *Br J Psychiatry* 154, pp. 797-800. doi: 10.1192/bjp.154.6.797

Sahakian, B. J. and Owen, A. M. 1992. Computerized assessment in neuropsychiatry using CANTAB: discussion paper. *J R Soc Med* 85(7), pp. 399-402.

Sahakian, B. J. et al. 1993. Further analysis of the cognitive effects of tetrahydroaminoacridine (THA) in Alzheimer's disease: assessment of attentional and mnemonic function using CANTAB. *Psychopharmacology (Berl)* 110(4), pp. 395-401. doi: 10.1007/bf02244644

Salvia, J., & Ysseldyke, J. E. 1998. *Assessment*. Boston, MA: Houghton Mifflin Company.

Sampson, J. R. et al. 1997. Renal cystic disease in tuberous sclerosis: role of the polycystic kidney disease 1 gene. *Am J Hum Genet* 61(4), pp. 843-851. doi: 10.1086/514888

Sancak, O. et al. 2005. Mutational analysis of the TSC1 and TSC2 genes in a diagnostic setting: genotype--phenotype correlations and comparison of diagnostic DNA techniques in Tuberous Sclerosis Complex. *Eur J Hum Genet* 13(6), pp. 731-741. doi: 10.1038/sj.ejhg.5201402

Sancak, Y. et al. 2008. The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* 320(5882), pp. 1496-1501. doi: 10.1126/science.1157535

Sandoval, J. 2014. Test review of Wechsler Abbreviated Scale of Intelligence—Second Edition.*In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), The nineteenth mental measurements yearbook.* http://marketplace.unl.edu/buros/.

Sarbassov, D. D. et al. 2004. Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. *Curr Biol* 14(14), pp. 1296-1302. doi: 10.1016/j.cub.2004.06.054

Sarbassov, D. D. et al. 2006. Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. *Molecular Cell* 22(2), pp. 159-168.

Sato, A. et al. 2012. Rapamycin reverses impaired social interaction in mouse models of tuberous sclerosis complex. *Nature Communications* 3(1), p. 1292. doi: 10.1038/ncomms2295

Saucedo, L. J. et al. 2003. Rheb promotes cell growth as a component of the insulin/TOR signalling network. *Nat Cell Biol* 5(6), pp. 566-571. doi: 10.1038/ncb996

Saxena, A. and Sampson, J. R. 2014. Phenotypes associated with inherited and developmental somatic mutations in genes encoding mTOR pathway components. *Semin Cell Dev Biol* 36, pp. 140-146. doi: 10.1016/j.semcdb.2014.09.018

Saxena, A. and Sampson, J. R. 2015. Epilepsy in Tuberous Sclerosis: Phenotypes, Mechanisms, and Treatments. *Semin Neurol* 35(3), pp. 269-276. doi: 10.1055/s-0035-1552616

Scott-Lennox, J. et al. 2001. Reliability, validity and responsiveness of a revised scoring system for the Liverpool Seizure Severity Scale. *Epilepsy Res* 44(1), pp. 53-63. doi: 10.1016/s0920-1211(01)00186-3

Selvaraj, S. et al. 2012. Grey matter differences in bipolar disorder: a meta-analysis of voxel-based morphometry studies. *Bipolar Disord* 14(2), pp. 135-145. doi: 10.1111/j.1399-5618.2012.01000.x

Shallice, T. 1982. Specific impairments of planning. *Philos Trans R Soc Lond B Biol Sci* 298(1089), pp. 199-209. doi: 10.1098/rstb.1982.0082

Shaw, R. J. et al. 2004. The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. *Proc Natl Acad Sci U S A* 101(10), pp. 3329-3335. doi: 10.1073/pnas.0308061100

Simon, T. J. et al. 2005. Volumetric, connective, and morphologic changes in the brains of children with chromosome 22q11.2 deletion syndrome: an integrative study. *Neuroimage* 25(1), pp. 169-180. doi: 10.1016/j.neuroimage.2004.11.018

Smalley, S. L. 1998. Autism and tuberous sclerosis. *J Autism Dev Disord* 28(5), pp. 407-414. doi: 10.1023/a:1026052421693

SMAReachUK. 2015. The SMA REACH UK project: A new initiative in collaboration with existing UK SMA registries.

Smith, S. M. 2002. Fast robust automated brain extraction. *Hum Brain Mapp* 17(3), pp. 143-155. doi: 10.1002/hbm.10062

Smith, S. M. et al. 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31(4), pp. 1487-1505. doi: 10.1016/j.neuroimage.2006.02.024

Soares, J. M. et al. 2013. A hitchhiker's guide to diffusion tensor imaging. *Front Neurosci* 7, p. 31. doi: 10.3389/fnins.2013.00031

Sofer, A. et al. 2005. Regulation of mTOR and cell growth in response to energy stress by REDD1. *Mol Cell Biol* 25(14), pp. 5834-5845. doi: 10.1128/MCB.25.14.5834-5845.2005

Soliman, G. A. et al. 2010. mTOR Ser-2481 autophosphorylation monitors mTORC-specific catalytic activity and clarifies rapamycin mechanism of action. *J Biol Chem* 285(11), pp. 7866-7879. doi: 10.1074/jbc.M109.096222

South, M. et al. 2017. Symptom overlap on the srs-2 adult self-report between adults with asd and adults with high anxiety. *Autism Res* 10(7), pp. 1215-1220. doi: 10.1002/aur.1764

Sparrow, S. S. et al. 2005. *Vineland-II : Vineland adaptive behavior scales : survey forms manual*. Circle Pines, Minn.: AGS Publishing.

Spiegelhalter, D. J. et al. 1999. An introduction to Bayesian methods in health technology assessment. *BMJ* 319(7208), pp. 508-512.

Spreen, O. and Strauss, E. 1998. *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press.

Stein, M. et al. 2012. Structural plasticity in the language system related to increased second language proficiency. *Cortex* 48(4), pp. 458-465. doi: https://doi.org/10.1016/j.cortex.2010.10.007

Stejskal, E. O. and Tanner, J. E. 1965. Spin diffusion measurements: spin echoes in the presence of a time‐dependent field gradient. *The journal of chemical physics* 42(1), pp. 288-292.

Stinnett, T. A. 2007. (Test review of The Test of Everyday Attention). *In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.),*

The seventeenth mental measurements yearbook.,

Sugranyes, G. et al. 2012. Multimodal analyses identify linked functional and white matter abnormalities within the working memory network in schizophrenia. *Schizophr Res* 138(2-3), pp. 136-142. doi: 10.1016/j.schres.2012.03.011

Tang, S. J. et al. 2002. A rapamycin-sensitive signaling pathway contributes to long-term synaptic plasticity in the hippocampus. *Proceedings of the National Academy of Sciences* 99(1), pp. 467-472. doi: 10.1073/pnas.012605299

Tartaglia, M. C. et al. 2012. Executive dysfunction in frontotemporal dementia is related to abnormalities in frontal white matter tracts. *J Neurol* 259(6), pp. 1071-1080. doi: 10.1007/s00415-011-6300-x

Tavazoie, S. F. et al. 2005. Regulation of neuronal morphology and function by the tumor suppressors Tsc1 and Tsc2. *Nature Neuroscience* 8(12), pp. 1727-1734. doi: 10.1038/nn1566

Tax, C. M. et al. 2019. Cross-scanner and cross-protocol diffusion MRI data harmonisation: A benchmark database and evaluation of algorithms. *Neuroimage* 195, pp. 285-299. doi: 10.1016/j.neuroimage.2019.01.077

Tee, A. R. et al. 2003. Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. *Curr Biol* 13(15), pp. 1259-1268. doi: 10.1016/s0960-9822(03)00506-2

Tello, R. et al. 1998. Meta analysis of the relationship between tuberous sclerosis complex and renal cell carcinoma. *Eur J Radiol* 27(2), pp. 131-138. doi: 10.1016/s0720-048x(97)00037-5

Thiele, E. A. 2004. Managing epilepsy in tuberous sclerosis complex. *J Child Neurol* 19(9), pp. 680-686.

Thoreen, C. C. et al. 2009. An ATP-competitive Mammalian Target of Rapamycin Inhibitor Reveals Rapamycin-resistant Functions of mTORC1*. *Journal of Biological Chemistry* 284(12), pp. 8023-8032. doi: https://doi.org/10.1074/jbc.M900301200

Tierney, K. M. et al. 2011. Neuropsychological Attention Skills and Related Behaviours in Adults with Tuberous Sclerosis Complex. *Behavior Genetics* 41(3), pp. 437-444. doi: 10.1007/s10519-010-9423-4

Tillema, J. M. et al. 2012. Everolimus alters white matter diffusion in tuberous sclerosis complex. *Neurology* 78(8), pp. 526-531. doi: 10.1212/WNL.0b013e318247ca8d

Toldo, I. et al. 2019. Tuberous sclerosis-associated neuropsychiatric disorders: a paediatric cohort study. *Dev Med Child Neurol* 61(2), pp. 168-173. doi: 10.1111/dmcn.14055

Torske, T. et al. 2017. Metacognitive Aspects of Executive Function Are Highly Associated with Social Functioning on Parent-Rated Measures in Children with Autism Spectrum Disorder. *Front Behav Neurosci* 11, p. 258. doi: 10.3389/fnbeh.2017.00258

Tversky, A. and Kahneman, D. 1971. Belief in the law of small numbers. *Psychological bulletin* 76(2), p. 105.

Tyburczy, M. E. et al. 2015. Mosaic and Intronic Mutations in TSC1/TSC2 Explain the Majority of TSC Patients with No Mutation Identified by Conventional Testing. *PLoS Genet* 11(11), p. e1005637. doi: 10.1371/journal.pgen.1005637

Tyburczy, M. E. et al. 2010. Novel proteins regulated by mTOR in subependymal giant cell astrocytomas of patients with tuberous sclerosis complex and new therapeutic implications. *Am J Pathol* 176(4), pp. 1878-1890. doi: 10.2353/ajpath.2010.090950

Tye, C. et al. 2018. Secular changes in severity of intellectual disability in tuberous sclerosis complex: A reflection of improved identification and treatment of epileptic spasms? *Epilepsia Open* 3(2), pp. 276-280. doi: 10.1002/epi4.12111

van Slegtenhorst, M. et al. 1997. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science* 277(5327), pp. 805-808.

Verhoef, S. et al. 1999. Malignant pancreatic tumour within the spectrum of tuberous sclerosis complex in childhood. *Eur J Pediatr* 158(4), pp. 284-287. doi: 10.1007/s004310051073

Vignoli, A. et al. 2013. Epilepsy in TSC: certain etiology does not mean certain prognosis. *Epilepsia* 54(12), pp. 2134-2142. doi: 10.1111/epi.12430

Vogt, H. 1908. Zur Diagnostik der tuberosen Sklerose. *Z. Erforsch. Behandl. jugendl. Schwachsinns.* 2, pp. 1-12.

Vollmar, C. et al. 2010. Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners. *Neuroimage* 51(4), pp. 1384-1394. doi: 10.1016/j.neuroimage.2010.03.046

Vos, S. B. et al. 2012. The influence of complex white matter architecture on the mean diffusivity in diffusion tensor MRI of the human brain. *Neuroimage* 59(3), pp. 2208-2216. doi: https://doi.org/10.1016/j.neuroimage.2011.09.086

Wadsworth, H. E. et al. 2018. Validity of teleneuropsychological assessment in older patients with cognitive disorders. *Archives of Clinical Neuropsychology* 33(8), pp. 1040-1045.

Wasserthal, J. et al. 2018. TractSeg - Fast and accurate white matter tract segmentation. *Neuroimage* 183, pp. 239-253. doi: 10.1016/j.neuroimage.2018.07.070

Webb, D. W. et al. 1991. Cranial magnetic resonance imaging in patients with tuberous sclerosis and normal intellect. *Arch Dis Child* 66(12), pp. 1375-1377.

Weber, A. M. et al. 2000. Autism and the cerebellum: evidence from tuberous sclerosis. *J Autism Dev Disord* 30(6), pp. 511-517. doi: 10.1023/a:1005679108529

Wechsler, D. 2011. WASI -II : Wechsler abbreviated scale of intelligence. *WASI*. San Antonio, Tex.: San Antonio, Tex. : Psychological Corporation.

Wesnes, K. and Warburton, D. M. 1984. Effects of scopolamine and nicotine on human rapid information processing performance. *Psychopharmacology (Berl)* 82(3), pp. 147-150. doi: 10.1007/BF00427761

Whicher, D. et al. 2018. An overview of the impact of rare disease characteristics on research methodology. *Orphanet J Rare Dis* 13(1), pp. 14-14. doi: 10.1186/s13023-017-0755-5

Wilkinson, L. 1999. Statistical methods in psychology journals: Guidelines and explanations. *American psychologist* 54(8), p. 594.

Wong, M. 2010. Mammalian target of rapamycin (mTOR) inhibition as a potential antiepileptogenic therapy: From tuberous sclerosis to common acquired epilepsies. *Epilepsia* 51(1), pp. 27-36. doi: 10.1111/j.1528-1167.2009.02341.x

Woo, S.-Y. et al. 2007. PRR5, a Novel Component of mTOR Complex 2, Regulates Platelet-derived Growth Factor Receptor β Expression and Signaling*. *Journal of Biological Chemistry* 282(35), pp. 25604-25612. doi: https://doi.org/10.1074/jbc.M704343200

Yang, H. et al. 2006. DNA damage-induced protein 14-3-3 sigma inhibits protein kinase B/Akt activation and suppresses Akt-activated cancer. *Cancer Res* 66(6), pp. 3096-3105. doi: 10.1158/0008-5472.CAN-05-3620

Yecies, J. L. and Manning, B. D. 2011. mTOR links oncogenic signaling to tumor cell metabolism. *J Mol Med (Berl)* 89(3), pp. 221-228. doi: 10.1007/s00109-011-0726-6

Yilmaz, Ö. H. et al. 2006. Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* 441(7092), pp. 475-482. doi: 10.1038/nature04703

Yu, F. et al. 2017. Repetitive Model of Mild Traumatic Brain Injury Produces Cortical Abnormalities Detectable by Magnetic Resonance Diffusion Imaging, Histopathology, and Behavior. *Journal of neurotrauma* 34(7), pp. 1364-1381. doi: 10.1089/neu.2016.4569

Zhang, H. et al. 2003. Loss of Tsc1/Tsc2 activates mTOR and disrupts PI3K-Akt signaling through downregulation of PDGFR. *J Clin Invest* 112(8), pp. 1223-1233. doi: 10.1172/JCI17222

Zhou, J. et al. 2009. mTOR supports long-term self-renewal and suppresses mesoderm and endoderm activities of human embryonic stem cells. *Proc Natl Acad Sci U S A* 106(19), pp. 7840-7845. doi: 10.1073/pnas.0901854106

Appendices

Appendix 1: Best response point analysis – Ordinal regression analysis

Table 17 provide descriptive and ordinal logistic regression results for differences between arms for the best response time point (Week 4, Week 12, Week 24 or Week 36). In the event of matched response at multiple time points, the first time point is selected. Figures 38 to 47 depicting Best response time from baseline as per the data in table 17

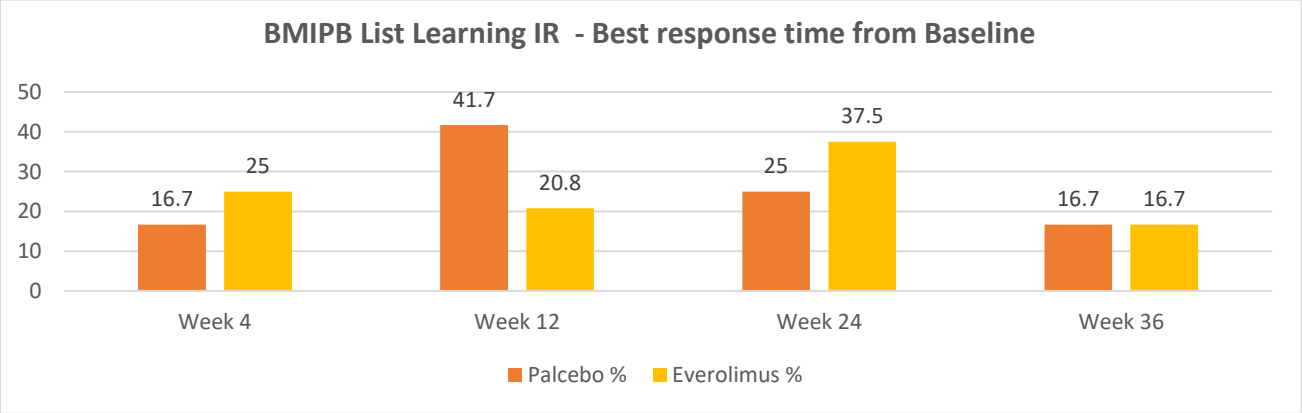


Figure 38 BMIPB List Learning IR - Best response time from Baseline

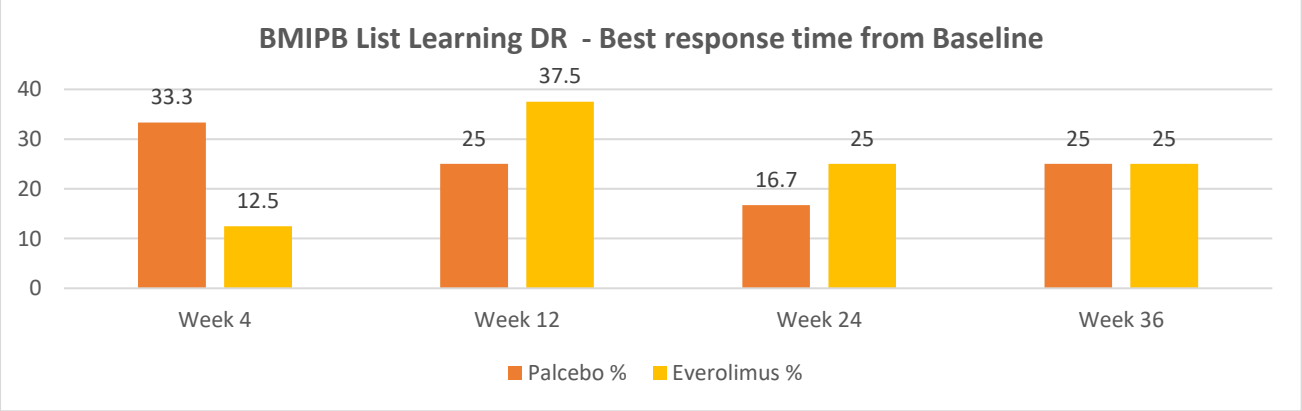


Figure 39 BMIPB List Learning DR - Best response time from Baseline

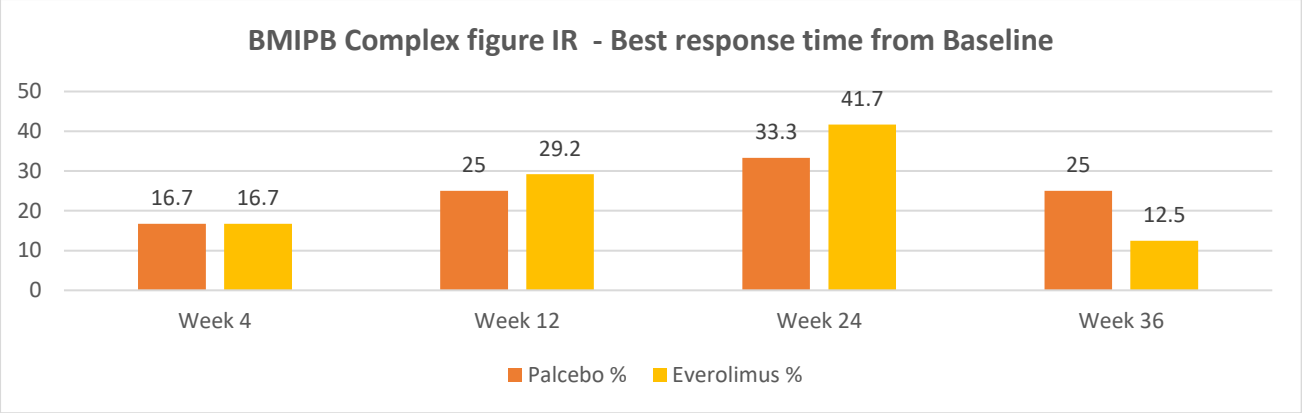


Figure 40 BMIPB Complex figure IR - Best response time from Baseline

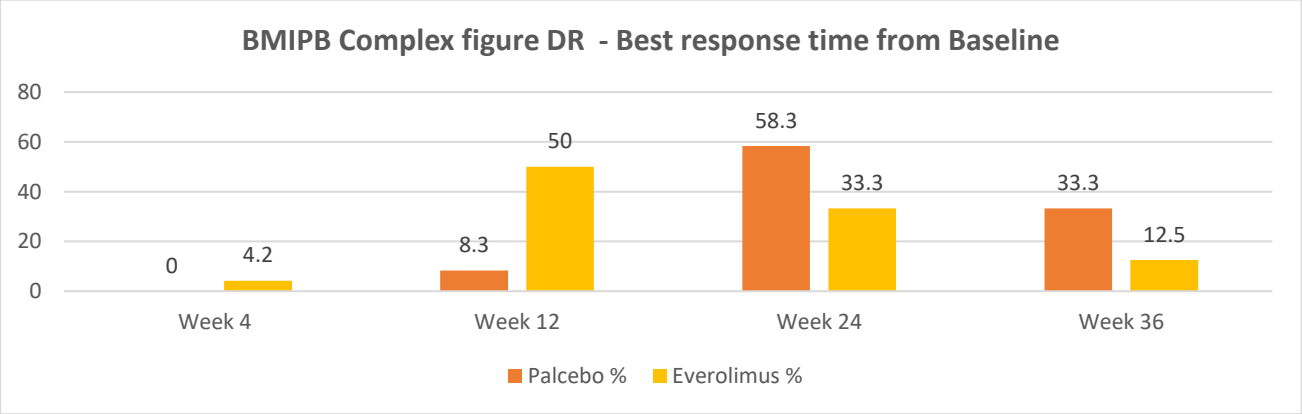


Figure 41 BMIPB Complex figure DR - Best response time from Baseline

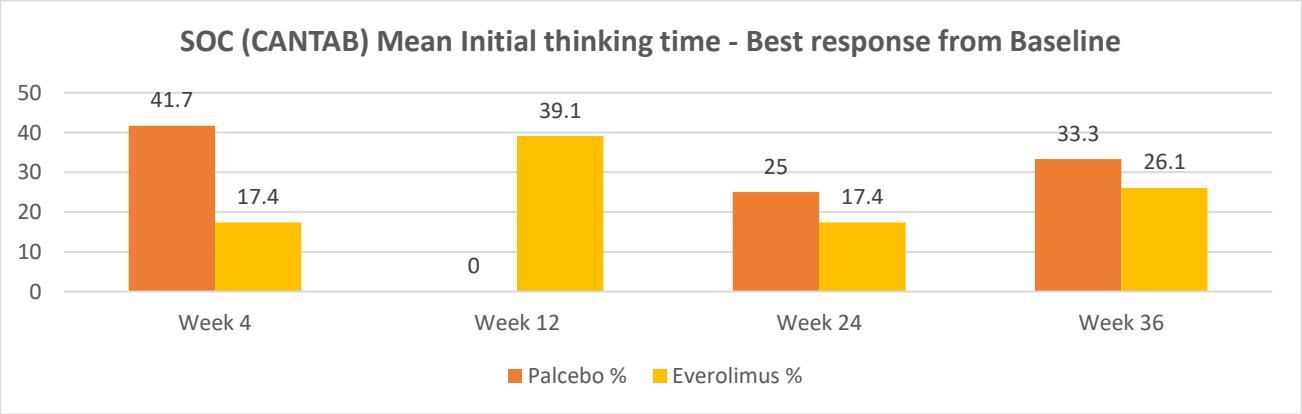


Figure 42 SOC (CANTAB) Mean Initial thinking time - Best response from Baseline

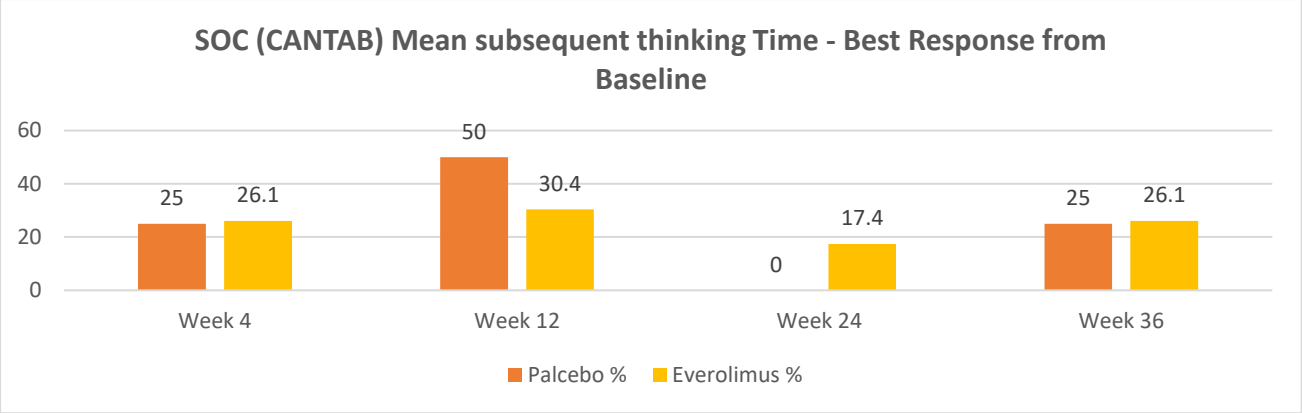


Figure 43 SOC (CANTAB) Mean subsequent thinking Time - Best Response from Baseline

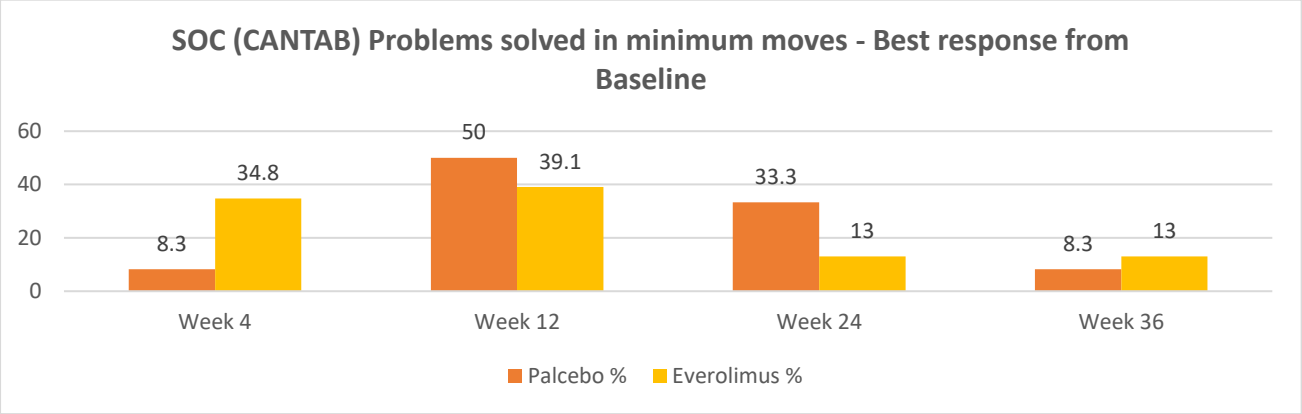


Figure 44 SOC (CANTAB) Problems solved in minimum moves - Best response from Baseline

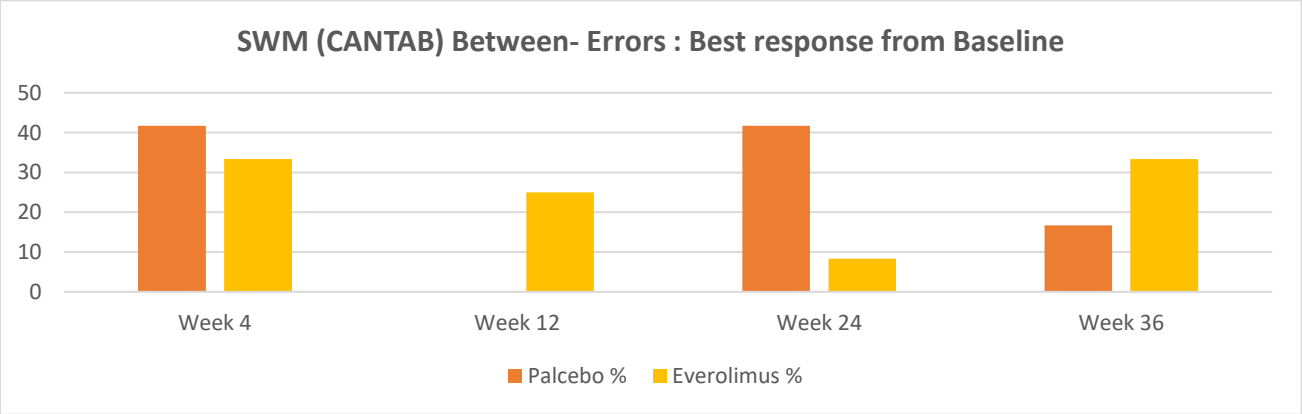


Figure 45 SWM (CANTAB) Between- Errors : Best response from Baseline

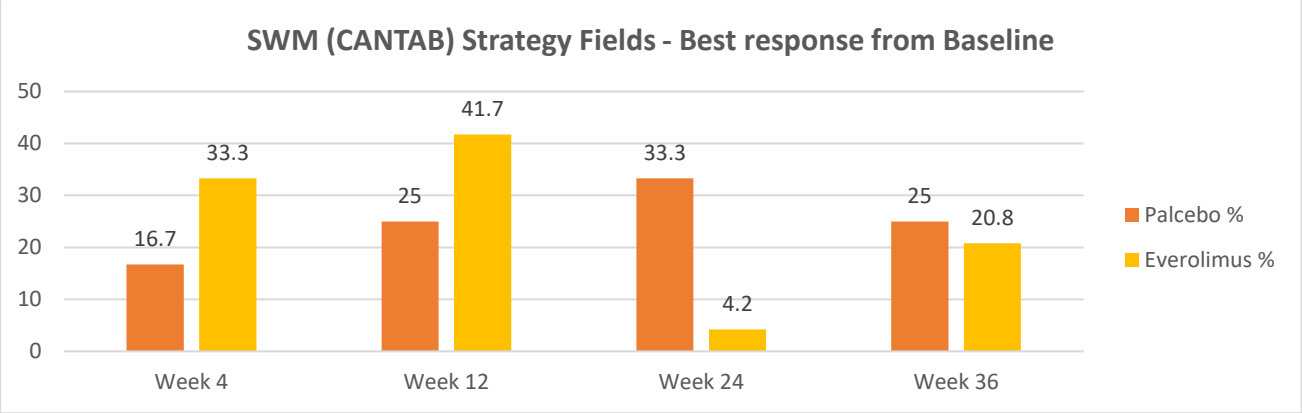


Figure 46 SWM (CANTAB) Strategy Fields - Best response from Baseline

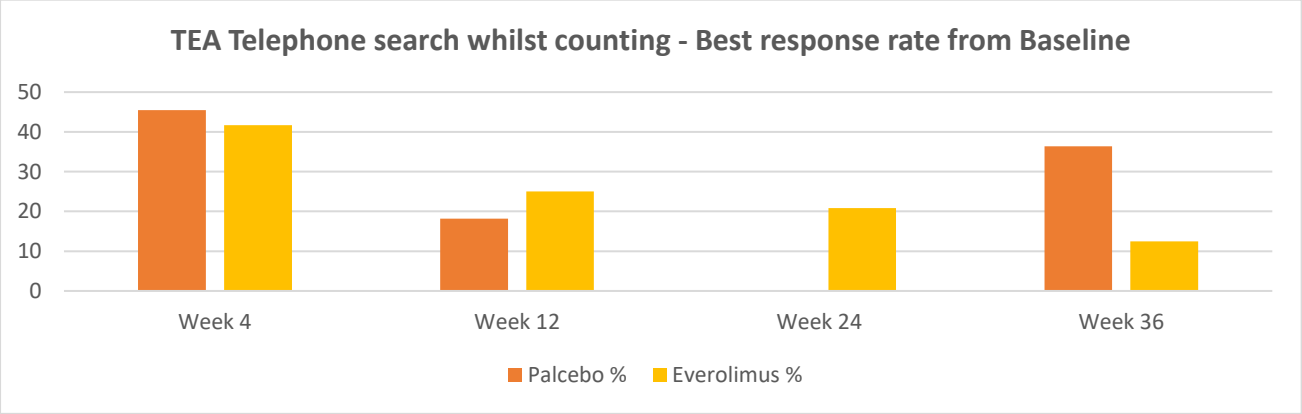


Figure 47 TEA Telephone search whilst counting - Best response rate from baseline

Table 28 provide descriptive and ordinal logistic regression results for differences between arms for the best response time point (Week 4, Week 12, Week 24 or Week 36) from screening visit for the whole population, while table 29 describes the regression analysis from baseline visit data excluding the participants recruited exclusively on TEA test results.

Red markings show that ordinal regression is not appropriate for these data. Yellow markings indicate significance at α = 0.05.

Table 28 Best response time point from Screening visit

		Arm										
		Placebo		Everolimus		Total					LR test for proportionality of odds p-value*[1]	Brant test p-value *[2]
		n	%	n	%	n	%	Odds Ratio	95% CI	p-value		
BMIPB List Learning Immediate Recall	Week 4	2	16.7	6	25.0	8	22.2	1.145	0.313 to 4.181	0.838	0.467	0.564
	Week 12	5	41.7	5	20.8	10	27.8					
	Week 24	3	25.0	9	37.5	12	33.3					
	Week 36	2	16.7	4	16.7	6	16.7					
BMIPB List Learning Delayed Recall	Week 4	4	33.3	3	12.5	7	19.4	1.589	0.426 to 5.926	0.490	0.170	0.365
	Week 12	3	25.0	9	37.5	12	33.3					
	Week 24	2	16.7	6	25.0	8	22.2					
	Week 36	3	25.0	6	25.0	9	25.0					

BMIPB Complex Figure Test Immediate Recall	Week 4	2	16.7	4	16.7	6	16.7	0.762	0.207 to 2.814	0.684	0.425	0.633
	Week 12	3	25.0	7	29.2	10	27.8					
	Week 24	4	33.3	10	41.7	14	38.9					
	Week 36	3	25.0	3	12.5	6	16.7					
BMIPB Complex Figure Test Delayed Recall	Week 4	0	0.0	1	4.2	1	2.8	0.153	0.035 to 0.662	0.012	0.459	N/A
	Week 12	1	8.3	12	50.0	13	36.1					
	Week 24	7	58.3	8	33.3	15	41.7					
	Week 36	4	33.3	3	12.5	7	19.4					
SOC (CANTAB) Mean initial thinking time	Week 4	5	41.7	4	16.7	9	25.0	1.253	0.334 to 4.703	0.738	0.003	<0.001
	Week 12	0	0.0	9	37.5	9	25.0					
	Week 24	3	25.0	4	16.7	7	19.4					
	Week 36	4	33.3	7	29.2	11	30.6					
SOC (CANTAB) Mean subsequent thinking time	Week 4	3	25.0	6	25.0	9	25.0	1.446	0.399 to 5.246	0.575	0.433	0.001
	Week 12	6	50.0	8	33.3	14	38.9					
	Week 24	0	0.0	4	16.7	4	11.1					
	Week 36	3	25.0	6	25.0	9	25.0					
SOC (CANTAB) Problems solved in minimum moves	Week 4	1	8.3	9	37.5	10	27.8	0.359	0.097 to 1.322	0.123	0.658	0.608
	Week 12	6	50.0	9	37.5	15	41.7					
	Week 24	4	33.3	3	12.5	7	19.4					
	Week 36	1	8.3	3	12.5	4	11.1					

SWM (CANTAB) Between-Errors	Week 4	5	41.7	8	33.3	13	36.1	1.169	0.328 to 4.160	0.810	0.084	<0.001
	Week 12	0	0.0	6	25.0	6	16.7					
	Week 24	5	41.7	2	8.3	7	19.4					
	Week 36	2	16.7	8	33.3	10	27.8					
SWM (CANTAB) Strategy Fields	Week 4	2	16.7	8	33.3	10	27.8	0.336	0.091 to 1.234	0.100	0.134	<0.001
	Week 12	3	25.0	10	41.7	13	36.1					
	Week 24	4	33.3	1	4.2	5	13.9					
	Week 36	3	25.0	5	20.8	8	22.2					
TEA Telephone search whilst counting	Week 4	6	50.0	9	39.1	15	42.9	1.105	0.254 to 4.805	0.894	0.218	0.027
	Week 12	2	16.7	6	26.1	8	22.9					
	Week 24	0	0.0	5	21.7	5	14.3					
	Week 36	4	33.3	3	13.0	7	20.0					

*[1]Approximate likelihood-ratio test of proportionality of odds across response categories

*[2]Brant Test of Parallel Regression Assumption

Table 29 Best response time point from baseline (minus those Screening on TEA only)

		Arm										
		Placebo		Everolimus		Total		Odds Ratio	95% CI	p-value	LR test for proportionality of odds p-value*[1]	Brant test p-value *[2]
		n	%	n	%	n	%					
BMIPB List Learning Immediate Recall	Week 4	1	11.1	6	28.6	7	23.3	0.819	0.189 to 3.544	0.789	0.069	<0.001
	Week 12	5	55.6	4	19.0	9	30.0					
	Week 24	1	11.1	8	38.1	9	30.0					
	Week 36	2	22.2	3	14.3	5	16.7					
BMIPB List Learning Delayed Recall	Week 4	4	44.4	3	14.3	7	23.3	3.795	0.733 to 19.652	0.112	0.322	1.000
	Week 12	3	33.3	8	38.1	11	36.7					
	Week 24	0	0.0	5	23.8	5	16.7					
	Week 36	2	22.2	5	23.8	7	23.3					

BMIPB Complex Figure Test Immediate Recall	Week 4	1	11.1	4	19.0	5	16.7	0.593	0.136 to 2.578	0.486	0.184	0.218
	Week 12	2	22.2	6	28.6	8	26.7					
	Week 24	4	44.4	8	38.1	12	40.0					
	Week 36	2	22.2	3	14.3	5	16.7					
BMIPB Complex Figure Test Delayed Recall	Week 4	0	0.0	1	4.8	1	3.3	0.114	0.021 to 0.617	0.012	0.625	N/A
	Week 12	1	11.1	11	52.4	12	40.0					
	Week 24	4	44.4	6	28.6	10	33.3					
	Week 36	4	44.4	3	14.3	7	23.3					
SOC (CANTAB) Mean initial thinking time	Week 4	5	55.6	4	20.0	9	31.0	2.361	0.496 to 11.229	0.280	0.010	1.000
	Week 12	0	0.0	7	35.0	7	24.1					
	Week 24	2	22.2	4	20.0	6	20.7					
	Week 36	2	22.2	5	25.0	7	24.1					

SOC (CANTAB) Mean subsequent thinking time	Week 4	2	22.2	5	25.0	7	24.1	1.004	0.228 to 4.424	0.996	0.776	0.141
	Week 12	4	44.4	7	35.0	11	37.9					
	Week 24	0	0.0	3	15.0	3	10.3					
	Week 36	3	33.3	5	25.0	8	27.6					
SOC (CANTAB) Problems solved in minimum moves	Week 4	1	11.1	7	35.0	8	27.6	0.281	0.060 to 1.326	0.109	0.003	N/A
	Week 12	4	44.4	9	45.0	13	44.8					
	Week 24	4	44.4	3	15.0	7	24.1					
	Week 36	0	0.0	1	5.0	1	3.4					
SWM (CANTAB) Between-Errors	Week 4	4	44.4	7	33.3	11	36.7	1.196	0.279 to 5.124	0.809	0.339	0.049
	Week 12	0	0.0	5	23.8	5	16.7					
	Week 24	3	33.3	2	9.5	5	16.7					
	Week 36	2	22.2	7	33.3	9	30.0					

SWM (CANTAB) Strategy Fields	Week 4	1	11.1	7	33.3	8	26.7	0.344	0.079 to 1.489	0.153	0.040	<0.001
	Week 12	3	33.3	9	42.9	12	40.0					
	Week 24	4	44.4	1	4.8	5	16.7					
	Week 36	1	11.1	4	19.0	5	16.7					

*[1]Approximate likelihood-ratio test of proportionality of odds across response categories

*[2]Brant Test of Parallel Regression Assumption

Appendix 2 : Improvement by >1 SD from Baseline or screening – mixed effect regression analysis

Tables 30 provide mixed-effects logistic regression results for differences in arms over time for improvement by ≥1SD from baseline or screening (as relevant). Tables 30 respectively represent whole study population improvement from Baseline, while table 31 and 32 represents the whole study population over time from Screening visit, while Table 32 shows the study population over time from baseline excluding those who were eligible on TEA test

Mixed-effects logistic models are used as an approach to address intracluster correlation in binary regression studies analysis (e.g. ≥1SD improvement), due to repeated measurements of individuals in time (longitudinal studies).

'Model does not converge' indicates that the data available for the analysis in question were not sufficient for estimates to be calculated. Of those measures with adequate data, none were

TABLE 30 Improvement of ≥ 1SD on Measure from Baseline over time (whole study population)

Measure	Arm interacting with Time Point (Placebo and Week 4 are references)	Odds Ratio	95% CI	p-value	p-value for Global Interaction Effect of Arm with Time Point
	Week 12	0.244	0.001 to 65.196	0.620	0.437

BMIPB List Learning Immediate Recall	Week 24	8.266	0.034 to 1991.906	0.450	
BMIPB List Learning Delayed Recall	Week 12	0.369	0.020 to 6.907	0.505	0.716
	Week 24	0.889	0.041 to 19.159	0.940	
BMIPB Complex Figure Test Immediate Recall	Week 12	0.668	0.041 to 10.818	0.776	0.712
	Week 24	2.031	0.123 to 33.668	0.621	
BMIPB Complex Figure Test Delayed Recall	Week 12	Model	does	not	converge
	Week 24				
SOC (CANTAB) Mean initial thinking time	Week 12	Model	does	not	converge
	Week 24				
SOC (CANTAB) Mean subsequent thinking time	Week 12	Model	does	not	converge
	Week 24				
SOC (CANTAB) Problems solved in minimum moves	Week 12	5.695	0.022 to 1444.546	0.538	0.542
	Week 24	0.259	0.001 to 66.594	0.633	

SWM (CANTAB) Between-Errors	Week 12	Model	does	not	converge
	Week 24				
SWM (CANTAB) Strategy Fields	Week 12	Model	does	not	converge
	Week 24				
TEA Telephone search whilst counting	Week 12	20.135	0.527 to 768.623	0.106	0.242
	Week 24	1.417	0.077 to 26.088	0.815	

Table 31 Improvement of 1SD on Measure from Screening over time (whole study population)

Measure	Arm interacting with Time Point (Placebo and Week 4 are references)	**Odds Ratio**	**95% CI**	**p-value**	**p-value for Global Interaction Effect of Arm with Time Point**
BMIPB List Learning Immediate Recall	Week 12	Model	does	not	converge
	Week 24				
	Week 12	0.367	0.005 to 29.222	0.654	0.873

BMIPB List Learning Delayed Recall	Week 24	0.367	0.005 to 29.222	0.654	
BMIPB Complex Figure Test Immediate Recall	Week 12	2.642	0.143 to 48.854	0.514	0.807
	Week 24	1.791	0.110 to 29.041	0.682	
BMIPB Complex Figure Test Delayed Recall	Week 12	Model	does	not	converge
	Week 24				
SOC (CANTAB) Mean initial thinking time	Week 12	Model	does	not	converge
	Week 24				
SOC (CANTAB) Mean subsequent thinking time	Week 12	1.000	0.008 to 131.692	1.000	0.469
	Week 24	20.726	0.093 to 4627.380	0.272	
SOC (CANTAB) Problems solved in minimum moves	Week 12	0.176	0.009 to 3.431	0.252	0.507
	Week 24	0.312	0.015 to 6.606	0.454	
SWM (CANTAB) Between-Errors	Week 12	4.338	0.045 to 418.378	0.529	0.820
	Week 24	4.580×10^{-9}	incalculable	0.992	
SWM (CANTAB) Strategy Fields	Week 12	Model	does	not	converge
	Week 24				

TEA Telephone search whilst counting	Week 12	52.219	0.310 to 8808.169	0.131	0.244
	Week 24	28.020	0.317 to 2474.790	0.145	

Table 32 Improvement of 1SD on Measure from Baseline over time (minus those Screening as eligible on TEA only)

Measure	Arm interacting with Time Point (Placebo and Week 4 are references)	**Odds Ratio**	**95% CI**	**p-value**	**p-value for Global Interaction Effect of Arm with Time Point**
BMIPB List Learning Immediate Recall	Week 12	0.234	0.001 to 58.938	0.607	0.455
	Week 24	7.244	0.032 to 1658.884	0.475	
BMIPB List Learning Delayed Recall	Week 12	0.630	0.032 to 12.522	0.762	0.785
	Week 24	1.783	0.060 to 52.8843	0.738	
BMIPB Complex Figure Test Immediate Recall	Week 12	Model	does	not	converge
	Week 24				
BMIPB Complex Figure Test Delayed Recall	Week 12	Model	does	not	converge
	Week 24				
	Week 12	Model	does	not	converge

SOC (CANTAB) Mean initial thinking time	Week 24				
SOC (CANTAB) Mean subsequent thinking time	Week 12	Model	does	not	converge
	Week 24				
SOC (CANTAB) Problems solved in minimum moves	Week 12	5.419	0.022 to 1360.208	0.549	0.548
	Week 24	0.250	0.001 to 63.258	0.624	
SWM (CANTAB) Between-Errors	Week 12	Model	does	not	converge
	Week 24				
SWM (CANTAB) Strategy Fields	Week 12	Model	does	not	converge
	Week 24				

Appendix 3: Continuous Measures – mixed effect regression analysis

The following tables provide results of mixed-effects linear regression analysis (continuous measure) for differences in study arms over time. Tables respectively represent full study population (table 33), with baseline visit values; whole study population with screening visit values (Table 34) and study population with those eligible on TEA test excluded (Table 35) with baseline values interacting with reference time point (4 week).

Table 33 Continuous Measures over time from baseline (whole study population)

Measure	Arm interacting with Time Point (Placebo and Week 4 are references)	Coefficient	95% CI	p-value	p-value for Global Interaction Effect of Arm with Time Point
BMIPB List Learning Immediate Recall	Week 12	-5.708	-10.823 to -0.593	0.029	0.074
	Week 24	-1.375	-6.490 to 3.740	0.598	
BMIPB List Learning Delayed Recall	Week 12	0.125	-1.500 to 1.750	0.880	0.791
	Week 24	0.542	-1.083 to 2.166	0.513	

BMIPB Complex Figure Test Immediate Recall	Week 12	9.417	-3.082 to 21.916	0.140	0.333
	Week 24	5.417	-7.082 to 17.916	0.396	
BMIPB Complex Figure Test Delayed Recall	Week 12	10.917	0.188 to 21.645	0.046	0.036
	Week 24	-2.333	-13.062 to 8.395	0.670	
SOC (CANTAB) Mean initial thinking time	Week 12	0.122	-0.138 to 0.383	0.359	0.627
	Week 24	0.026	-0.234 to 0.287	0.844	
SOC (CANTAB) Mean subsequent thinking time	Week 12	0.704	-0.306 to 1.715	0.172	0.282
	Week 24	-0.013	-1.024 to 0.998	0.980	
SOC (CANTAB) Problems solved in minimum moves	Week 12	-0.636	-1.288 to 0.017	0.056	0.133
	Week 24	-0.496	-1.148 to 0.156	0.136	
SWM (CANTAB) Between-Errors	Week 12	0.159	-0.293 to 0.612	0.491	0.132
	Week 24	-0.299	-0.751 to 0.154	0.196	

SWM (CANTAB) Strategy Fields	Week 12	-0.056	-0.473 to 0.361	0.791	0.076
	Week 24	-0.443	-0.861 to -0.027	0.037	
TEA Telephone search whilst counting	Week 12	0.659	-2.298 to 3.616	0.662	0.904
	Week 24	0.186	-2.772 to 3.143	0.902	

Table 34 Continuous Measures over time from Screening (whole study population)

Measure	Arm interacting with Time Point (Placebo and Week 4 are references)	**Coefficient**	**95% CI**	**p-value**	**p-value for Global Interaction Effect of Arm with Time Point**
BMIPB List Learning Immediate Recall	Week 12	-5.708	-10.823 to -0.593	0.029	0.074
	Week 24	-1.375	-6.490 to 3.740	0.598	
BMIPB List Learning Delayed Recall	Week 12	0.125	-1.500 to 1.750	0.880	0.791
	Week 24	0.542	-1.083 to 2.166	0.513	

BMIPB Complex Figure Test Immediate Recall	Week 12	9.417	-3.082 to 21.916	0.140	0.333
	Week 24	5.417	-7.082 to 17.916	0.396	
BMIPB Complex Figure Test Delayed Recall	Week 12	10.917	0.188 to 21.645	0.046	0.036
	Week 24	-2.333	-13.062 to 8.395	0.670	
SOC (CANTAB) Mean initial thinking time	Week 12	0.124	-0.132 to 0.379	0.342	0.594
	Week 24	0.020	-0.235 to 0.275	0.878	
SOC (CANTAB) Mean subsequent thinking time	Week 12	0.723	-0.184 to 1.629	0.118	0.157
	Week 24	-0.089	-0.996 to 0.818	0.848	
SOC (CANTAB) Problems solved in minimum moves	Week 12	-0.660	-1.302 to -0.018	0.044	0.102
	Week 24	-0.532	-1.174 to 0.110	0.104	
SWM (CANTAB) Between-Errors	Week 12	0.159	-0.293 to 0.612	0.491	0.132
	Week 24	-0.299	-0.751 to 0.154	0.196	
SWM (CANTAB) Strategy Fields	Week 12	-0.056	-0.473 to 0.361	0.791	0.076
	Week 24	-0.443	-0.861 to -0.027	0.037	
TEA Telephone search whilst counting	Week 12	1.304	-1.359 to 3.967	0.337	0.620
	Week 24	0.870	-1.793 to 3.533	0.522	

Table 35 Continuous Measures over time from baseline (minus those Screening as eligible on TEA only)

Measure	Arm interacting with Time Point (Placebo and Week 4 are references)	**Coefficient**	**95% CI**	**p-value**	**p-value for Global Interaction Effect of Arm with Time Point**
BMIPB List Learning Immediate Recall	Week 12	-6.317	-12.066 to -0.568	0.031	0.062
	Week 24	-0.698	-6.447 to 5.051	0.812	
BMIPB List Learning Delayed Recall	Week 12	0.619	-1.273 to 2.511	0.521	0.623
	Week 24	0.920	-0.972 to 2.813	0.340	
BMIPB Complex Figure Test Immediate Recall	Week 12	10.524	-3.978 to 25.026	0.155	0.364
	Week 24	5.476	-9.026 to 19.978	0.459	
BMIPB Complex Figure Test Delayed Recall	Week 12	11.429	-1.264 to 24.122	0.078	0.070
	Week 24	-2.619	-15.312 to 10.074	0.686	
SOC (CANTAB) Mean initial thinking time	Week 12	0.126	-0.186 to 0.438	0.428	0.723
	Week 24	0.082	-0.230 to 0.394	0.605	

SOC (CANTAB) Mean subsequent thinking time	Week 12	0.980	-0.259 to 2.215	0.121	0.191
	Week 24	-0.034	-1.271 to 1.203	0.957	
SOC (CANTAB) Problems solved in minimum moves	Week 12	-0.434	-1.192 to 0.323	0.261	0.353
	Week 24	-0.521	-1.278 to 0.237	0.178	
SWM (CANTAB) Between-Errors	Week 12	0.162	-0.349 to 0.672	0.535	0.153
	Week 24	-0.333	-0.843 to 0.177	0.201	
SWM (CANTAB) Strategy Fields	Week 12	-0.069	-0.546 to 0.408	0.777	0.063
	Week 24	-0.526	-1.003 to -0.049	0.031	

4-week, 12-week, 24-week and 36-weeks assessments across the whole study population

Appendix 4 : Imaging study

- **CUBRIC MRI data Preprocessing pathway**

The raw data of TRON participants was pre-processed using the standard Single-shell DTI processing pipeline for CUBRIC.

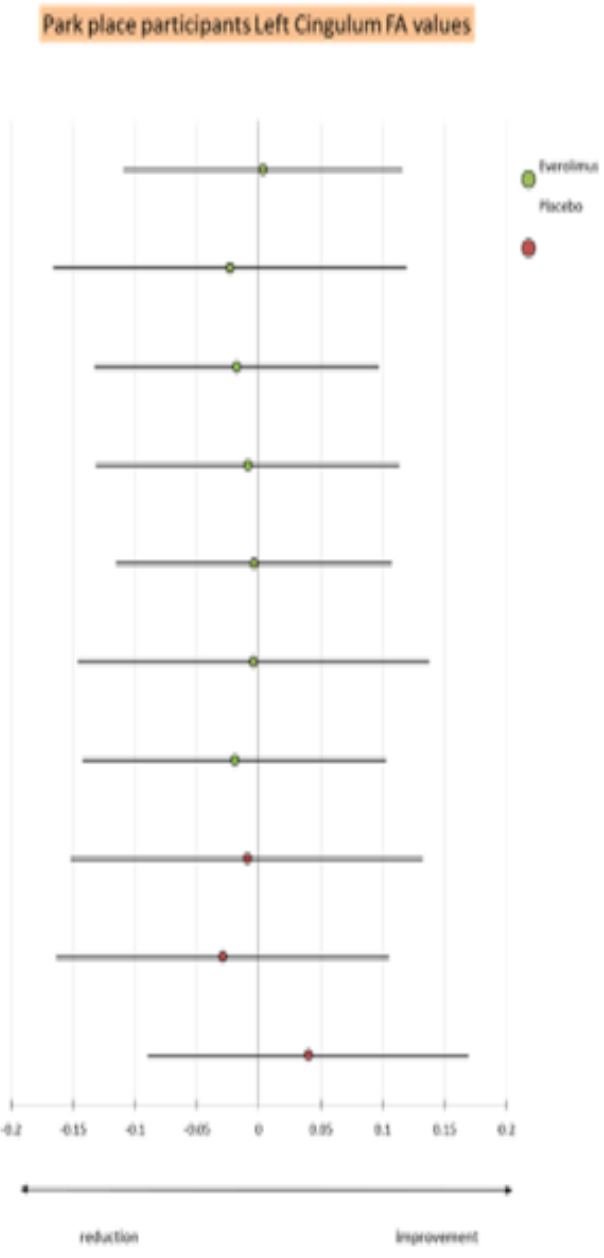
The outline for the analysis was:

1. Preprocessing pathway
    a. Converted DTI files to nifti (Neuroimaging Informatics Technology Initiative) files : 3 files – Main DTI files and .bval + .bvec files
    b. Created (structural) T1 nifti files for anatomical registration of the DTI files.
        i. Skull strip of T1 nifti file to create '.bet' file using Brain extraction tool(Smith 2002). This is crosschecked with FSLview for quality control.
        ii. Downsample T1 nifti file – created DS (down-sample) file by 1.5

2. Files needed for Tractseg using Explore DTI_4.8.6
    a. Preproc_DWIs.nii file using Plugins: *.bval/*.bvec file(s)
3. Bilateral Individual Tractography for the following tracts using TracSeg:
    a. Fornix
    b. Uncinate fasciculus
    c. Cingulum
    d. Superior longitudinal fasciculus 1, 2 & 3

4. Extract mean FA & MD from TractSeg - After pre-processing the raw data, the target tracts were analysed using TractSeg analysis standardised in CUBRIC.
5. Data inspection was done using FiberNavigator (Chamberland et al. 2014) in a few sample tracts.

- **Forest Plots (Fig 48 to 67) depicting the change in the FA and MD value for the TRON imaging study participants as per the CUBRIC1 or CUBRIC 2 cohort**



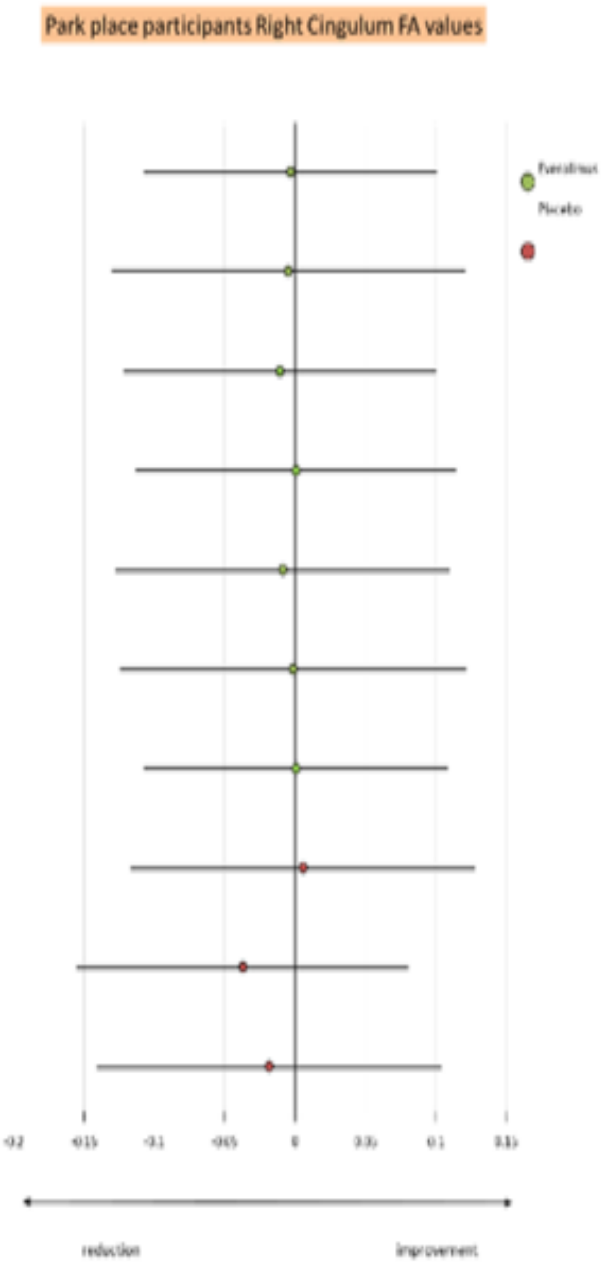


FIGURE 48 FA VALUE CHANGE IN CINGULUM (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

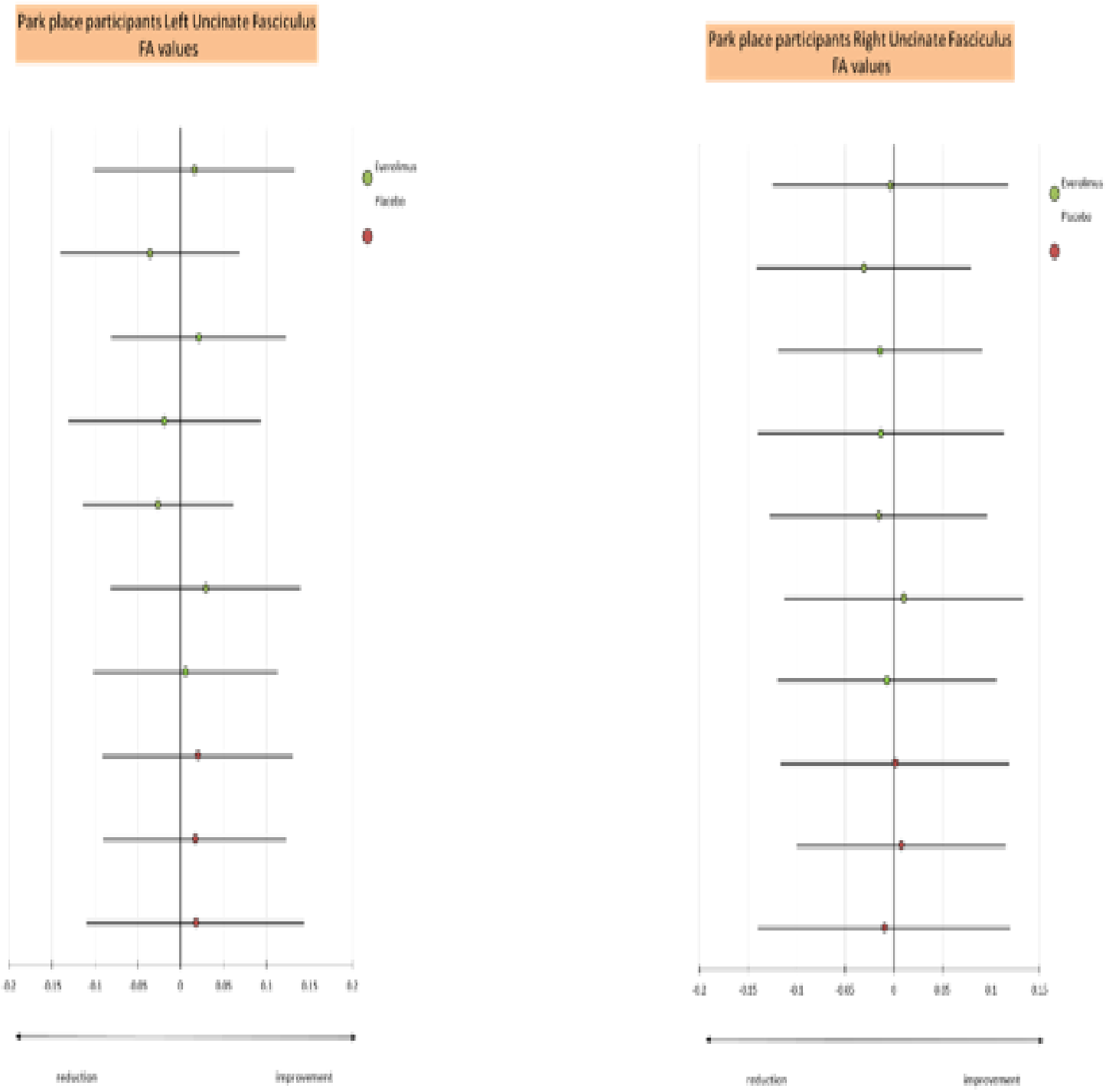


FIGURE 49 FA VALUE CHANGE IN UNCINATE FASCICULUS (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

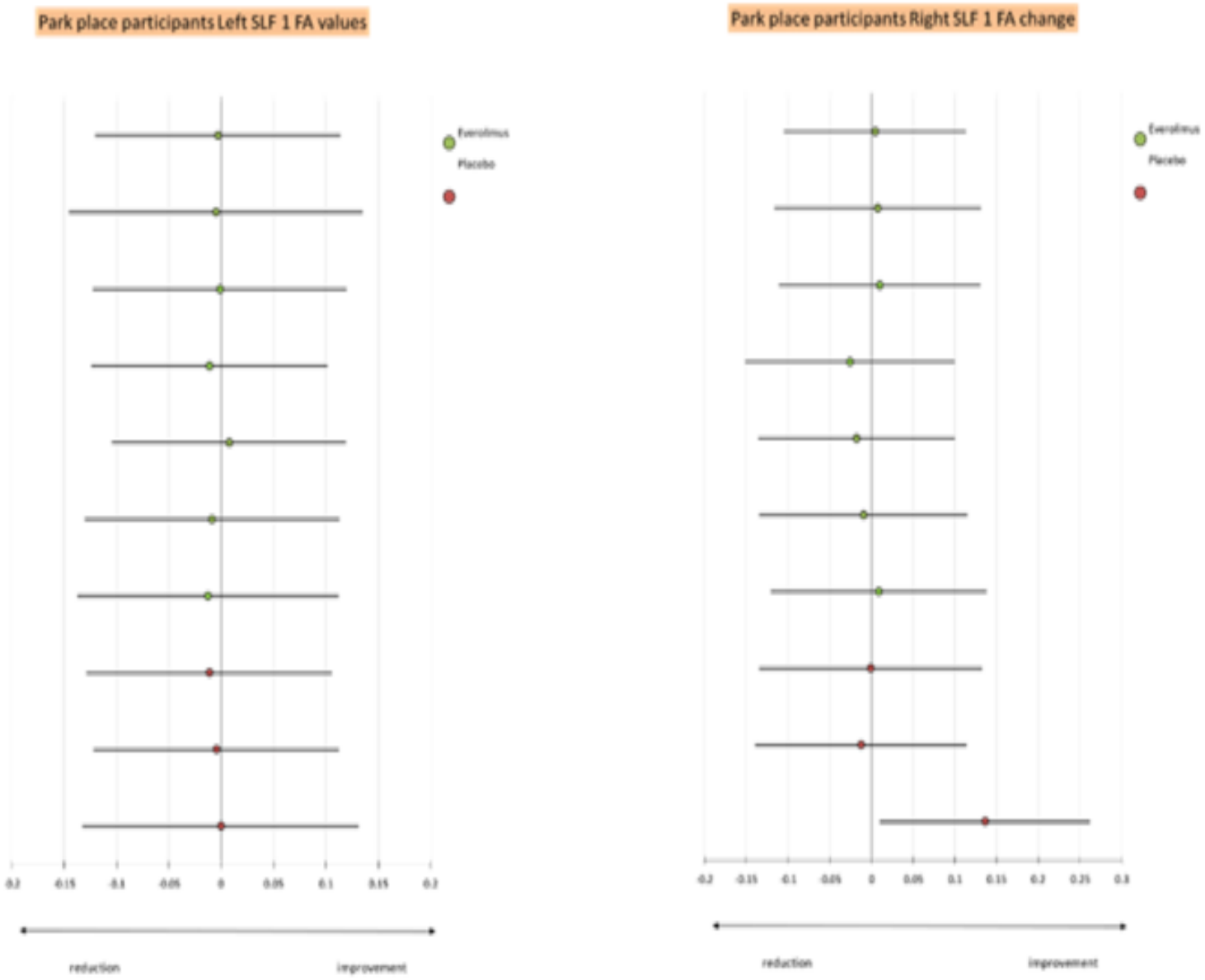


FIGURE 50 FA VALUE CHANGE IN SLF-1 (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

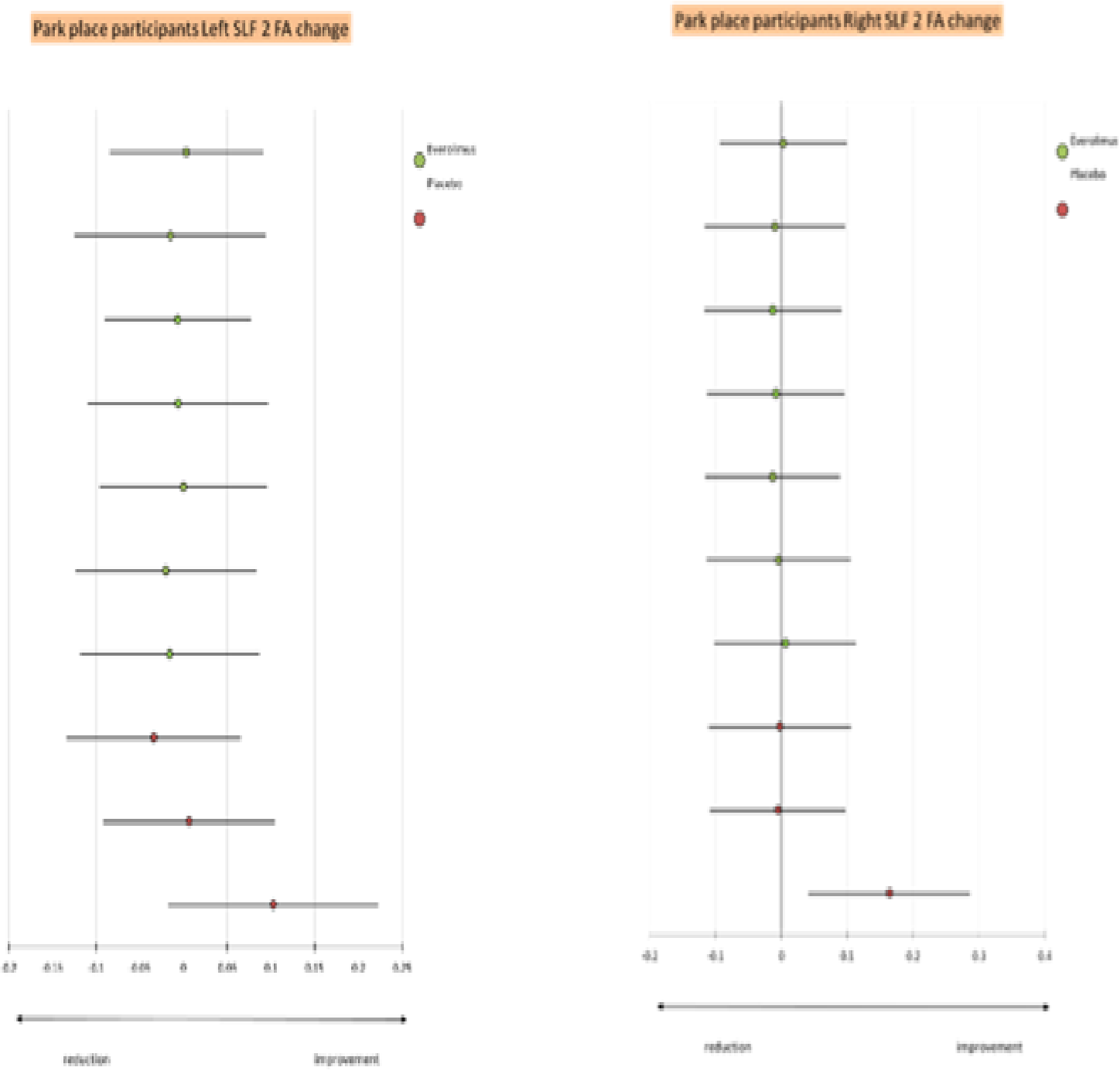


FIGURE 51 FA VALUE CHANGE IN SLF-2 (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

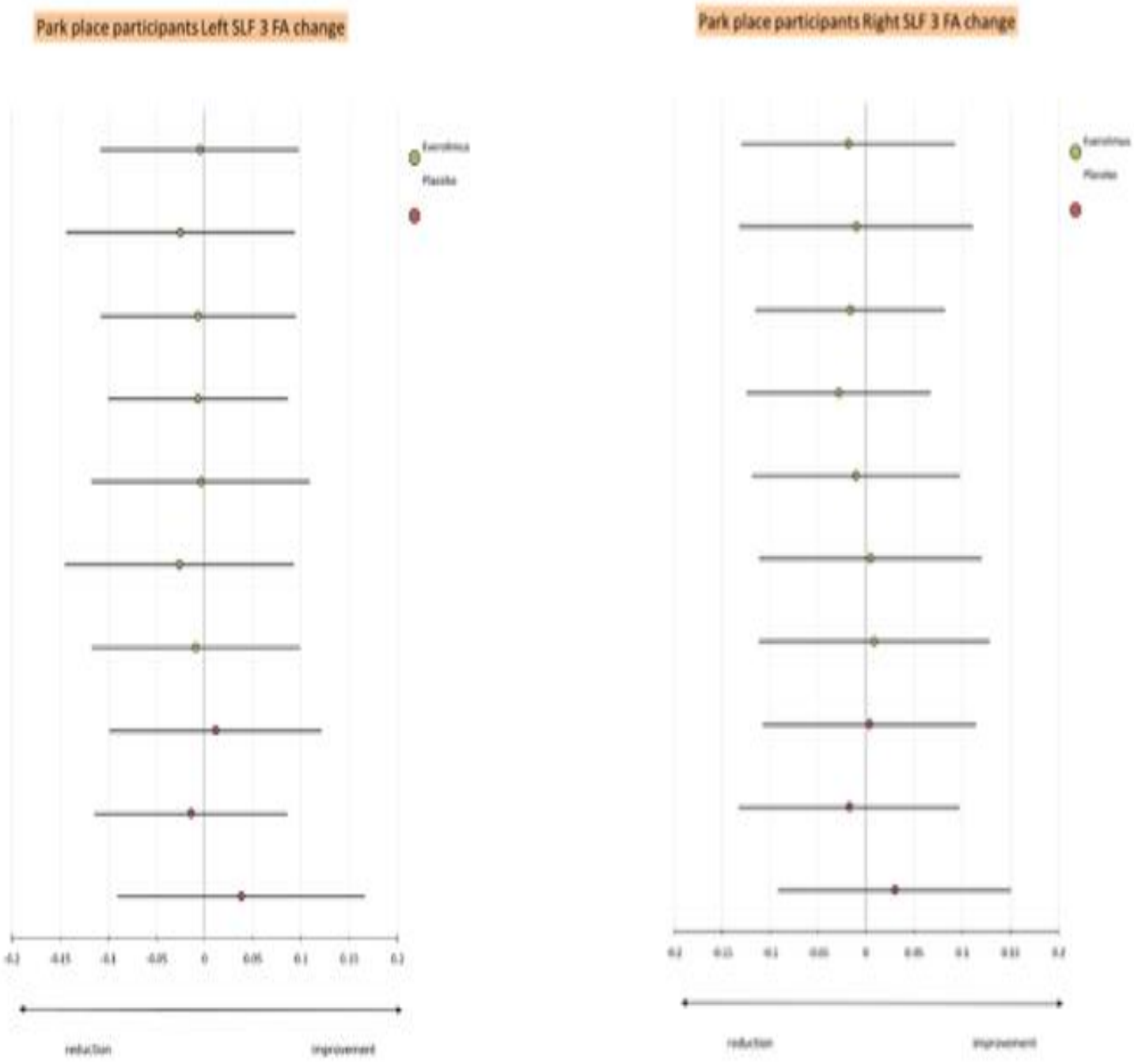
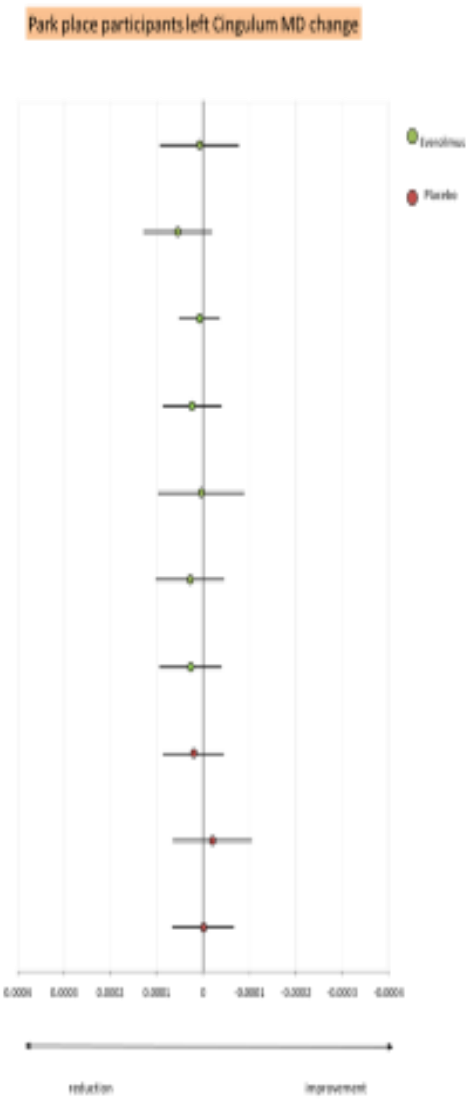


FIGURE 52 FA VALUE CHANGE IN SLF-3 (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANT



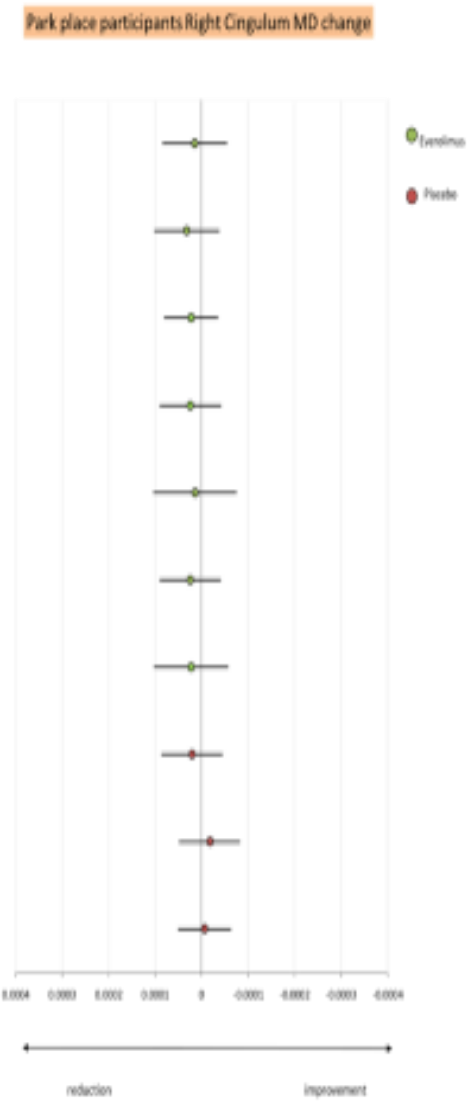


FIGURE 53 MD VALUE CHANGE IN CINGULUM (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

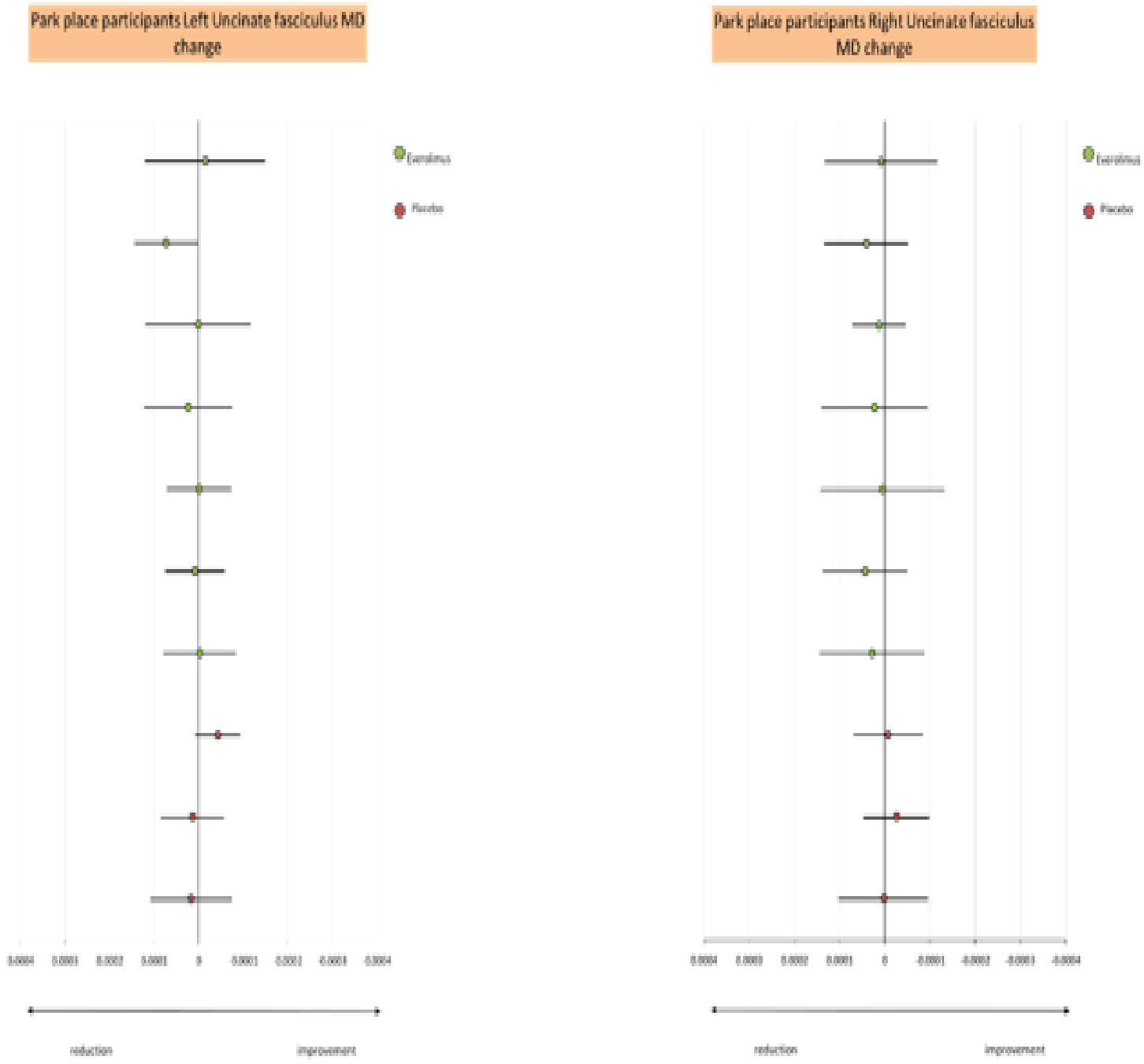


FIGURE 54 MD VALUE CHANGE IN UNCINATE FASCICULUS (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

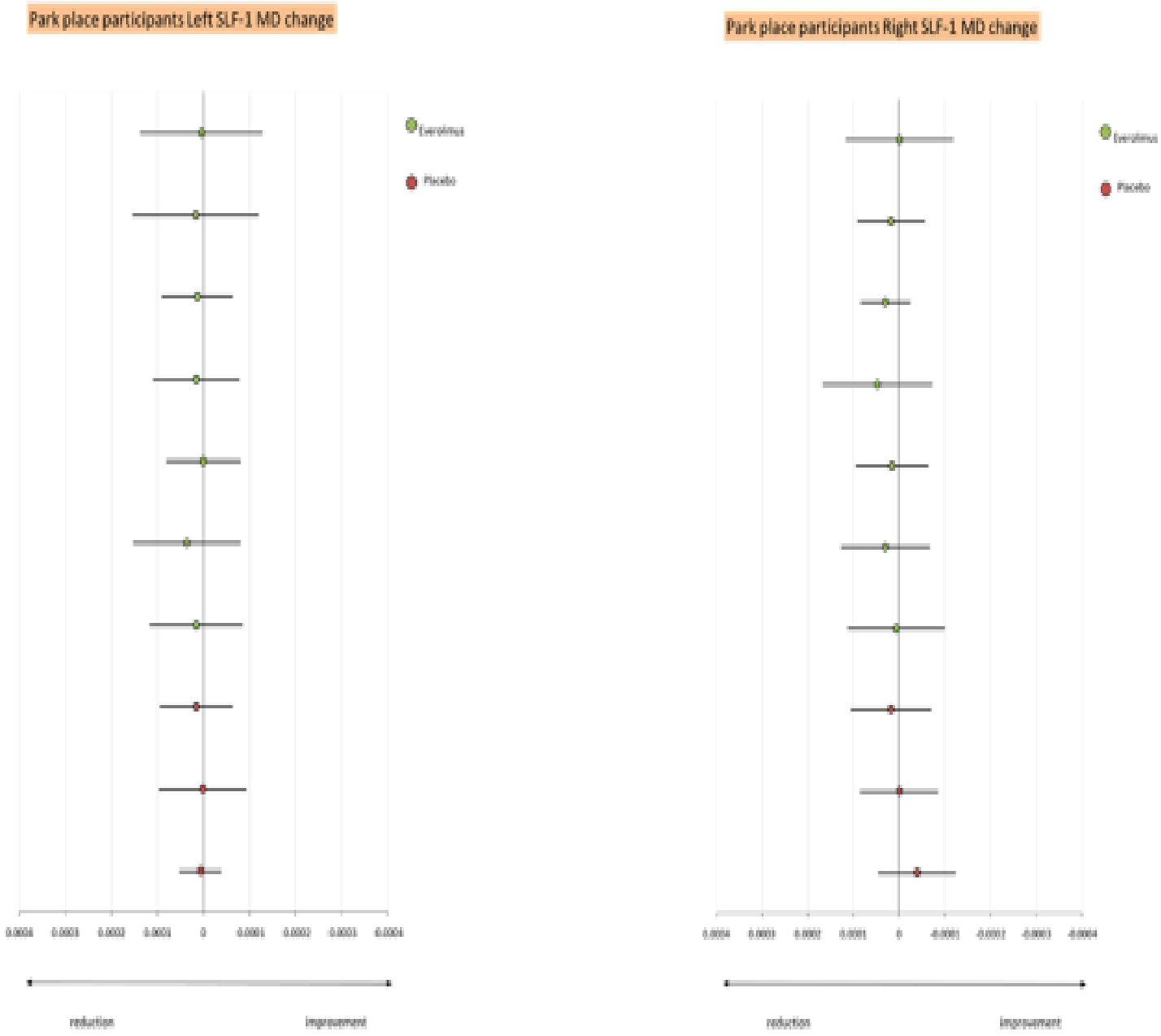
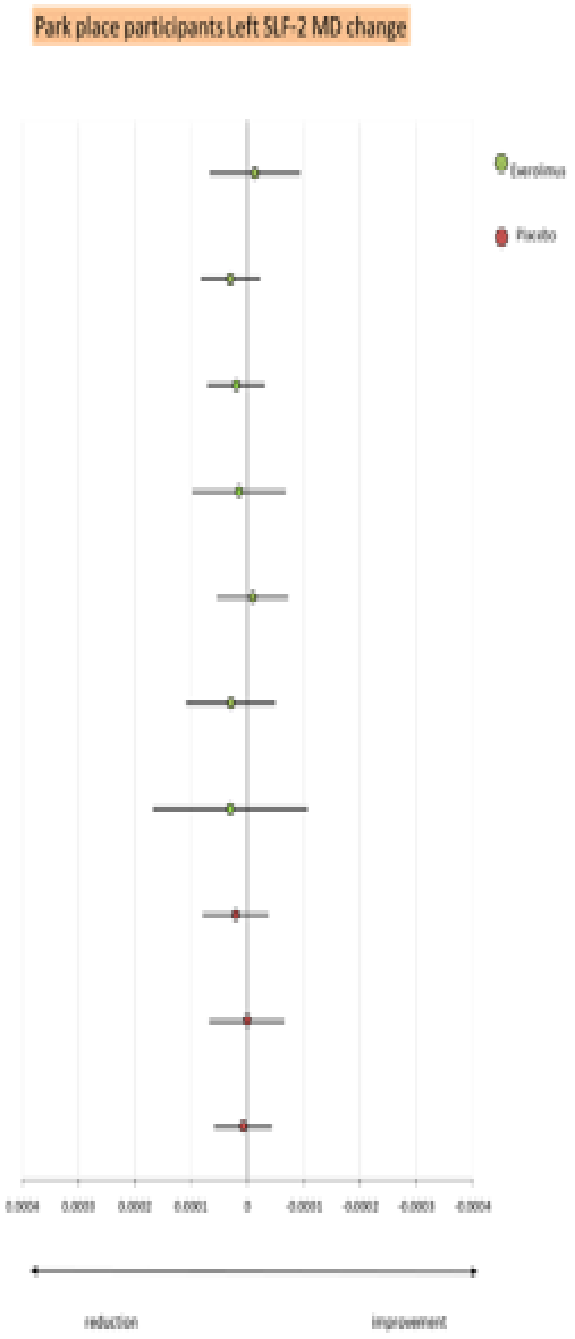


FIGURE 55 MD VALUE CHANGE IN SLF-1 (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS



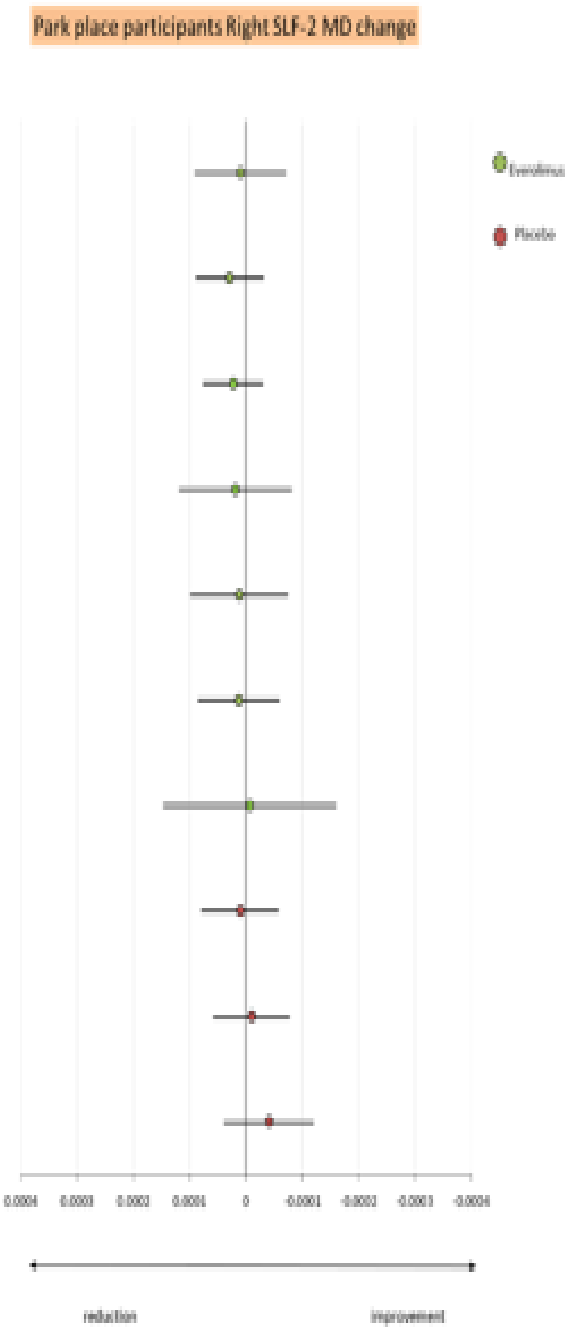


FIGURE 56 MD VALUE CHANGE IN SLF-2 (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

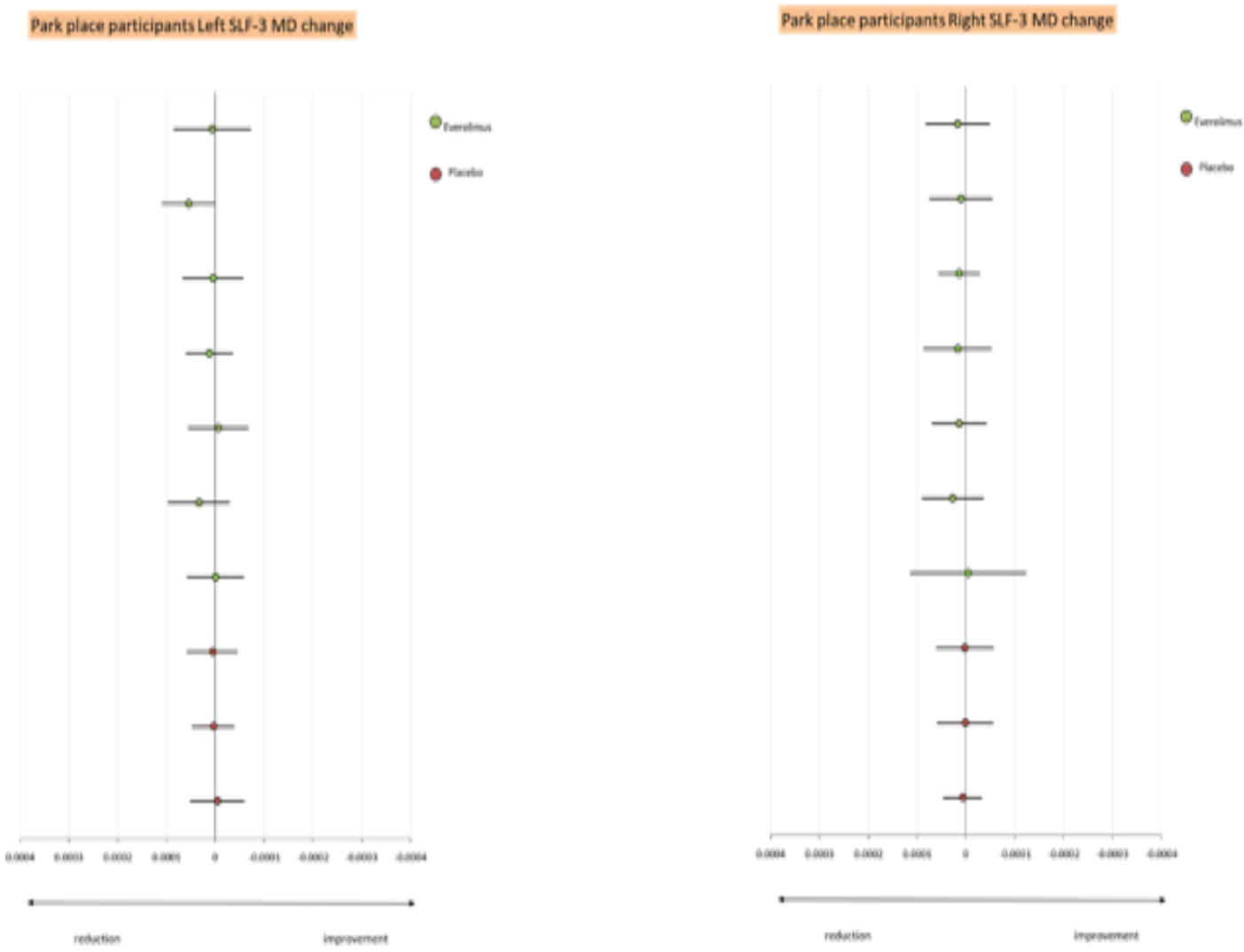
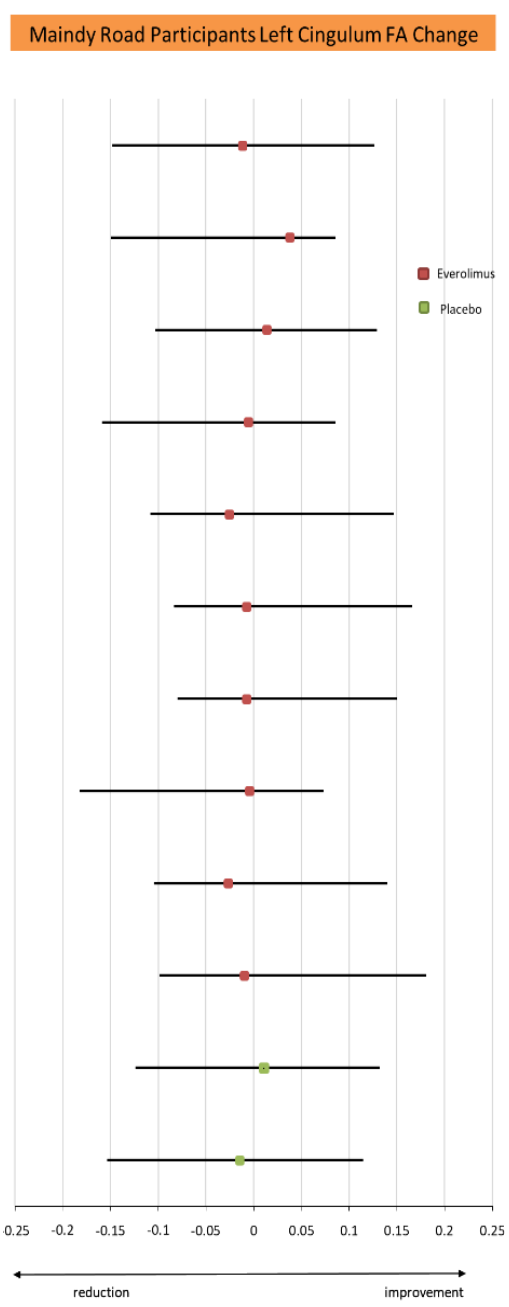


FIGURE 57 MD VALUE CHANGE IN SLF-3 (LEFT & RIGHT) FOR CUBRIC 1 PARTICIPANTS

Figure 58 to 62 depict the FA value change for CUBRIC 2 (Maindy road) participants for CG, UF, SLF-1, SLF-2 and SLF-3 respectively.



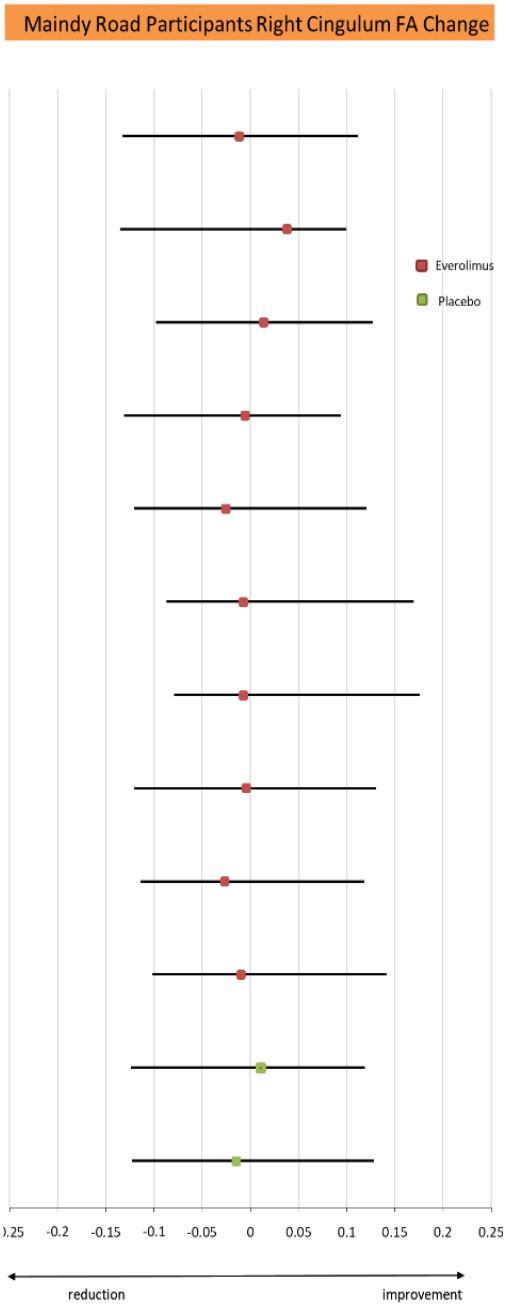


FIGURE 58 FA VALUE CHANGE IN CINGULUM (LEFT & RIGHT) FOR CUBRIC 2 PARTICIPANTS

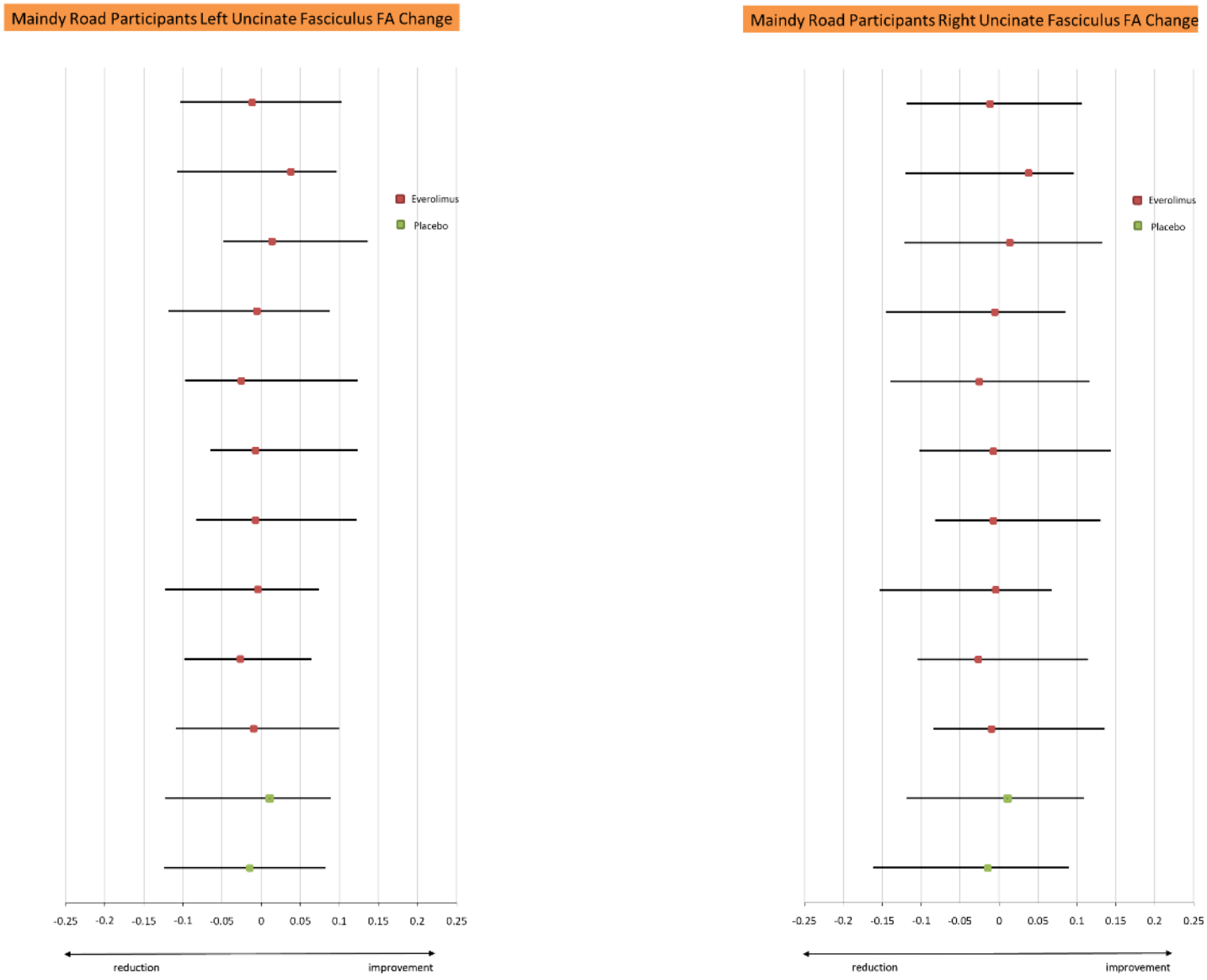
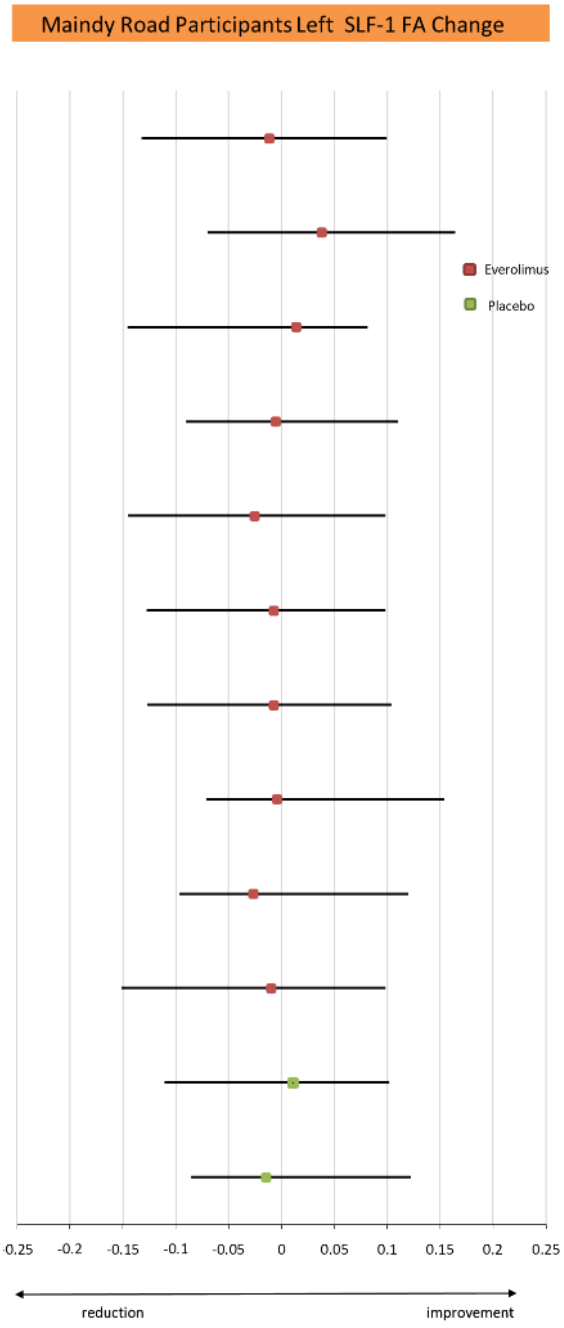


FIGURE 59 FA VALUE CHANGE IN UNCINATE FASCICULUS (LEFT & RIGHT) FOR CUBRIC 2 PARTICIPANTS



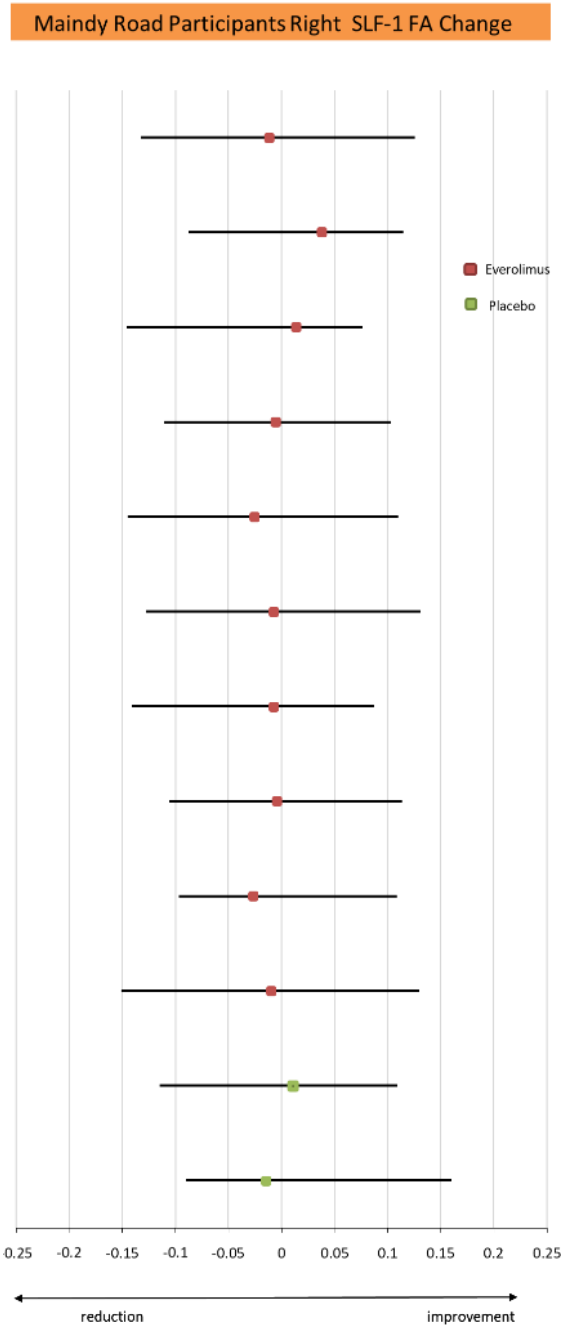
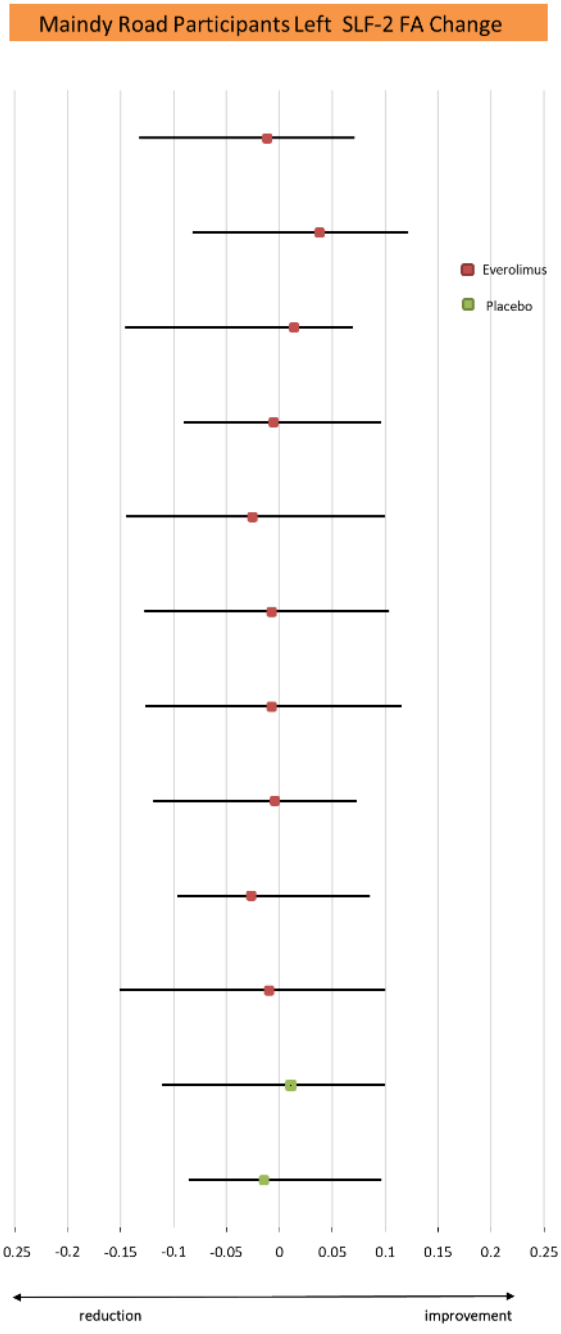


FIGURE 60 FA VALUE CHANGE IN SLF-1 (LEFT & RIGHT) FOR CUBRIC 2 PARTICIPANTS



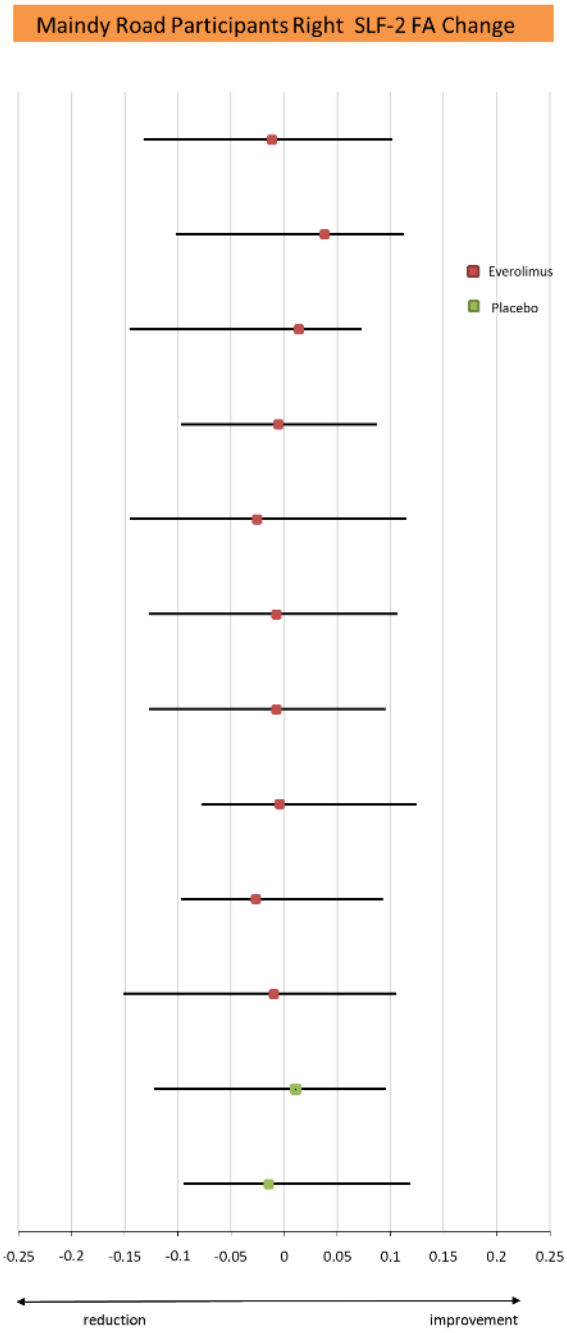
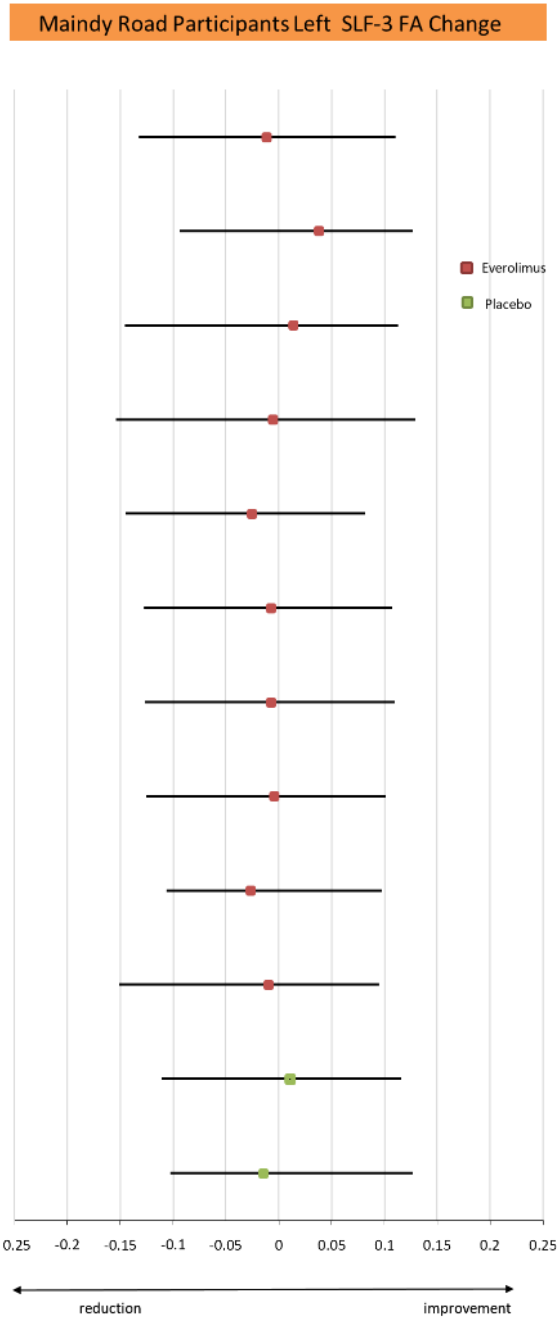


FIGURE 61 FA VALUE CHANGE IN SLF-2 (LEFT & RIGHT) FOR CUBRIC 2 PARTICIPANTS



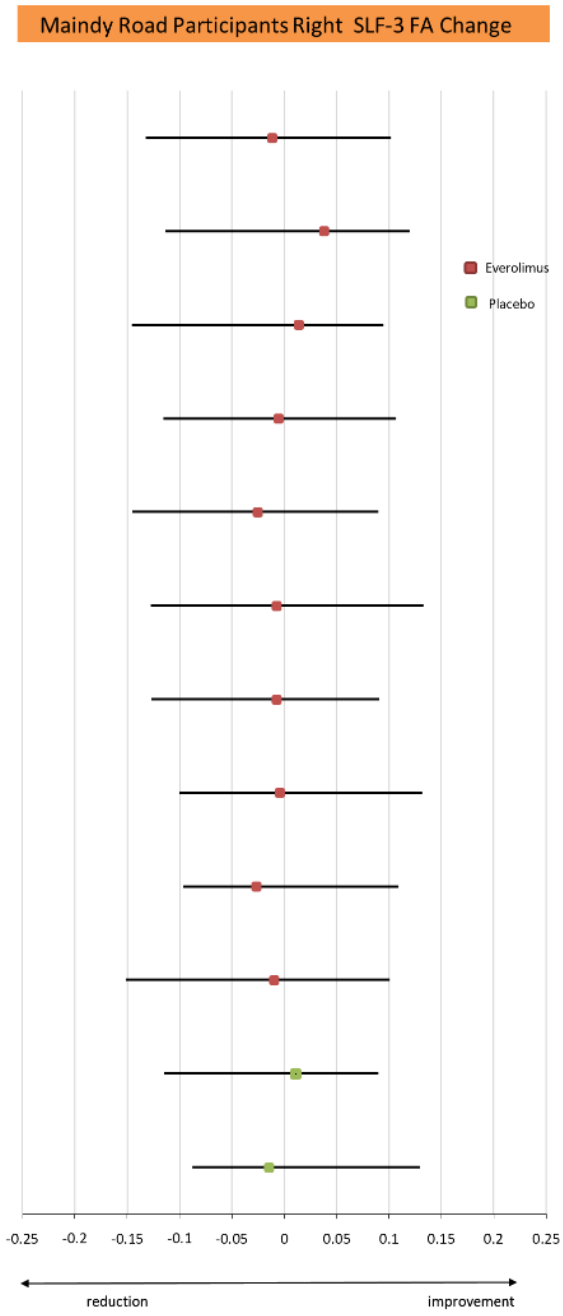
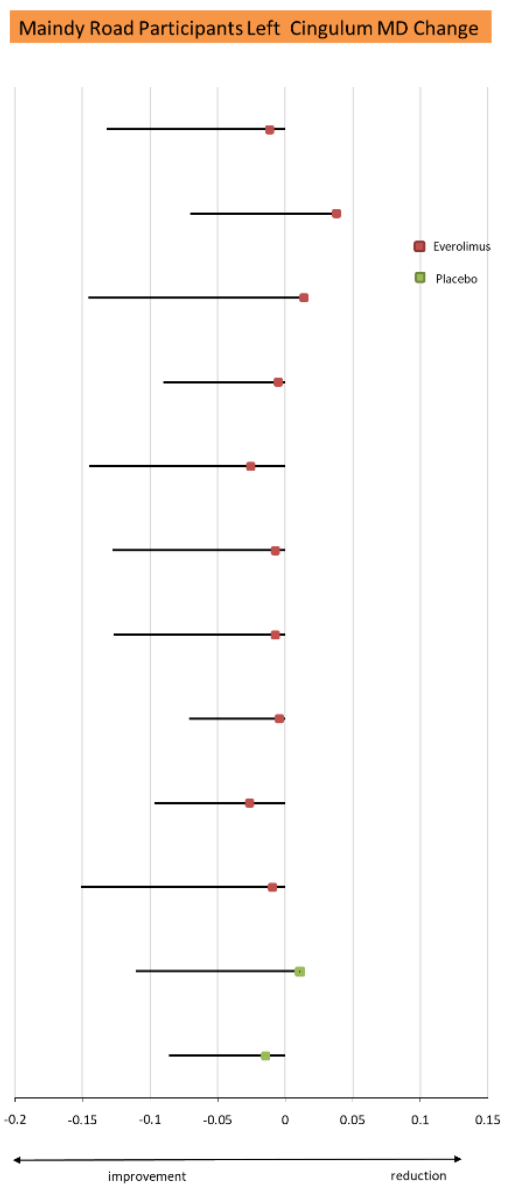


FIGURE 62 FA VALUE CHANGE IN SLF-3 (LEFT & RIGHT) FOR CUBRIC 2 PARTICIPANTS

Figure 63 to 67 depict the MD value change for CUBRIC 2 (Maindy road) participants for CG, UF, SLF-1, SLF-2 and SLF-3 respectively.



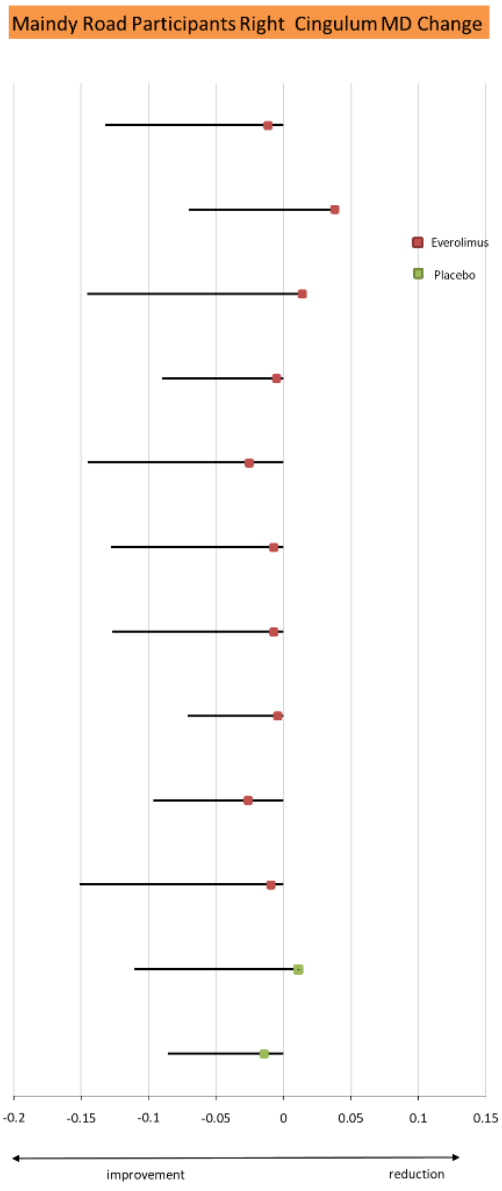
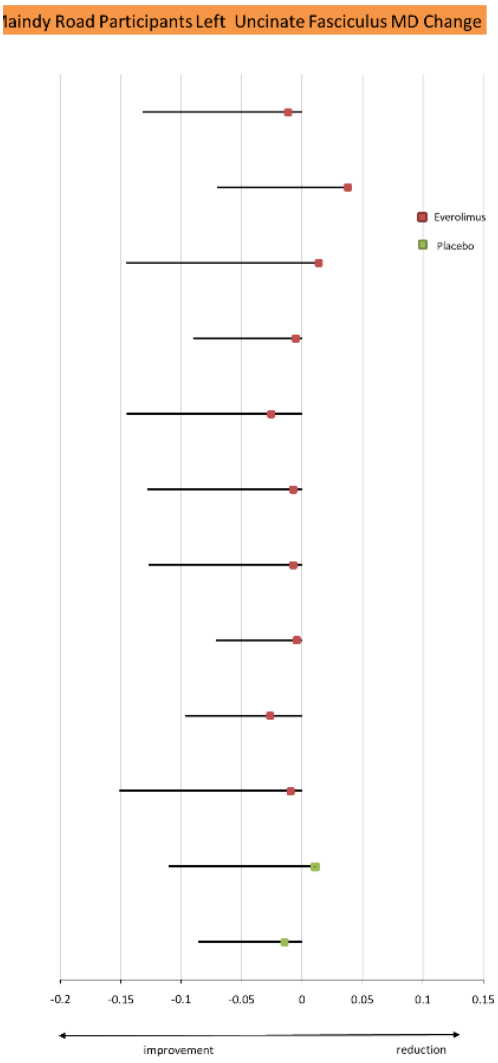


FIGURE 63 CINGULUM MD CHANGE FOR CUBRIC 2 PARTICIPANTS



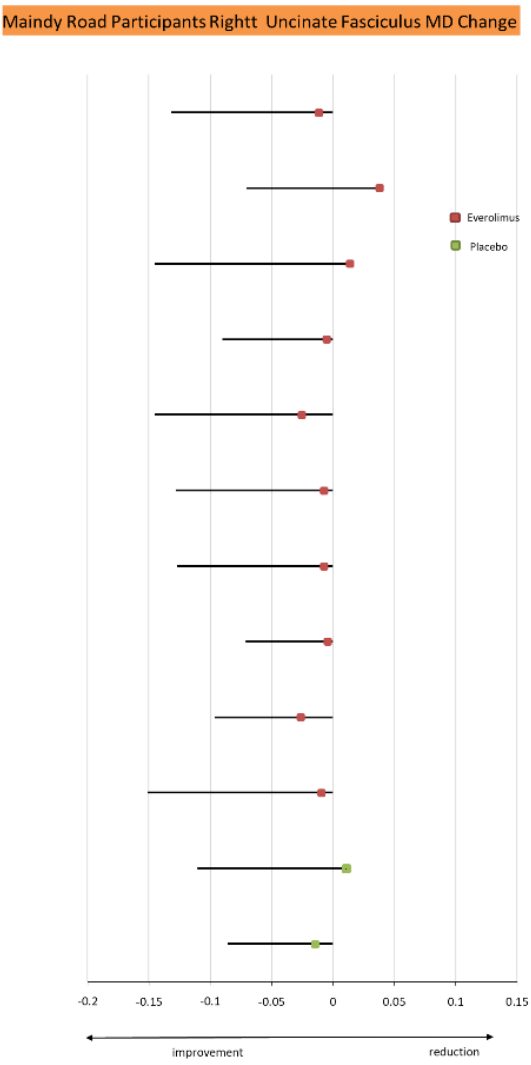
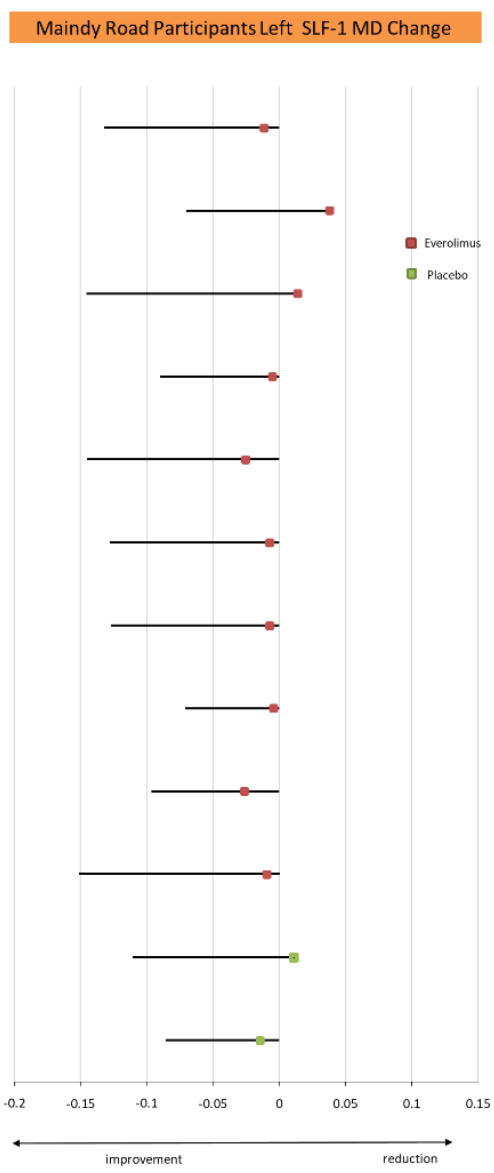


FIGURE 64 MD CHANGE FOR UF IN CUBRIC 2 PARTICIPANTS



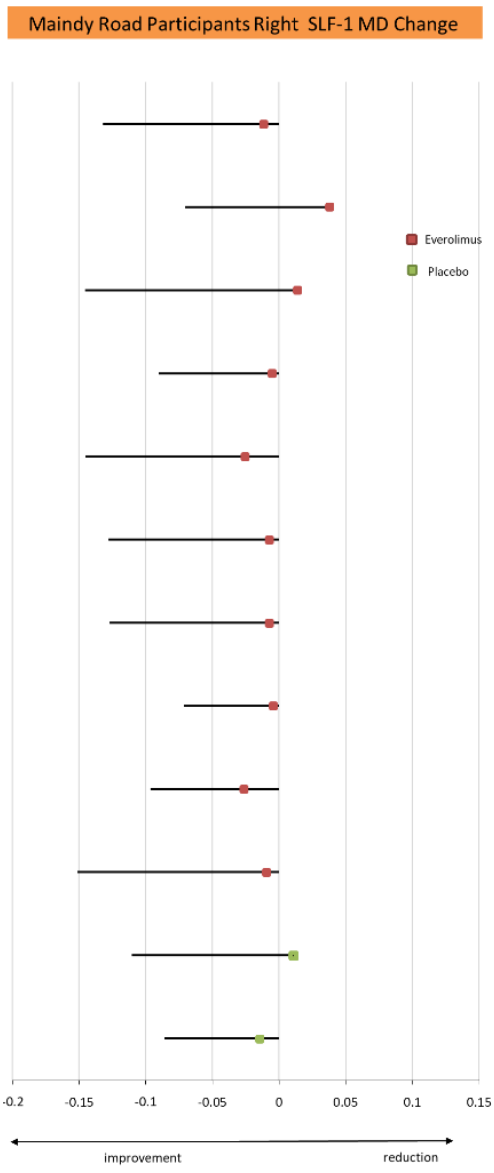
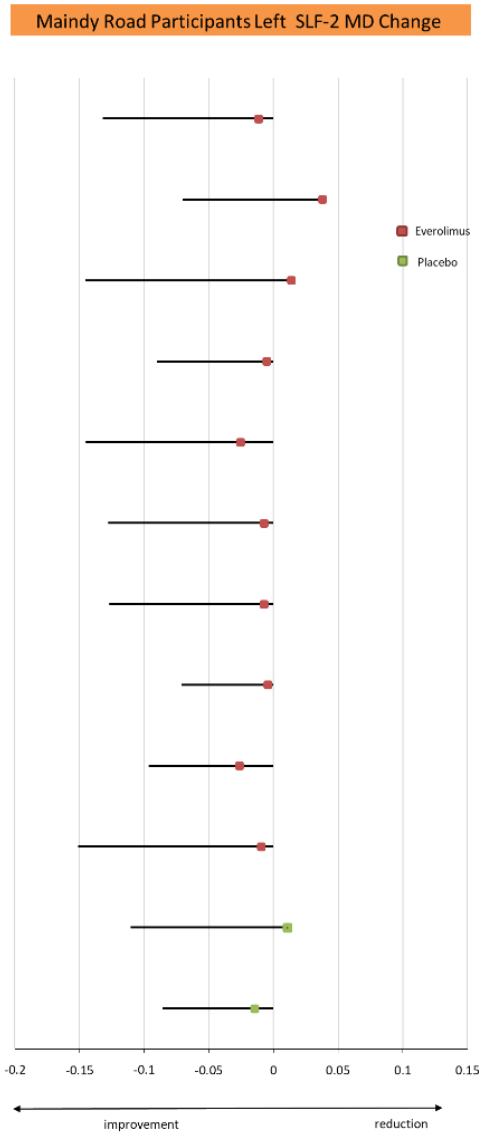


FIGURE 65 MD CHANGE IN SLF-1 TRACT FOR CUBRIC 2 PARTICIPANTS



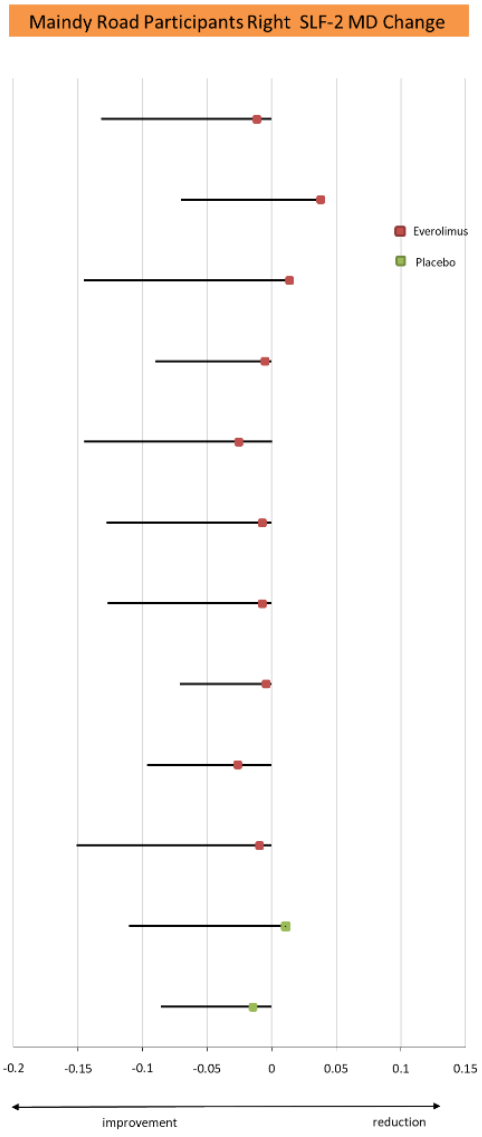
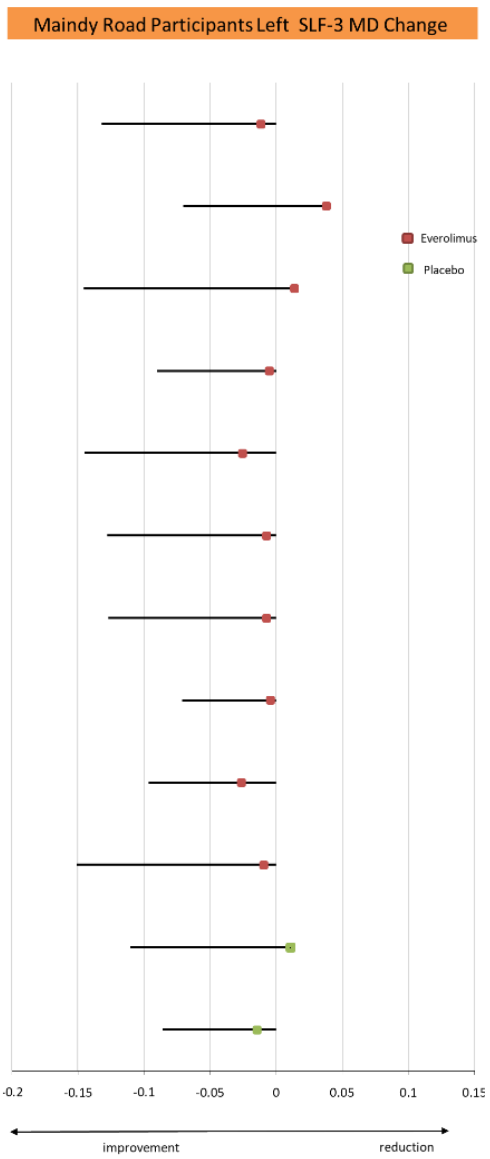


FIGURE 66 MD CHANGE SLF-2 TRACT FOR CUBRIC 2 PARTICIPANTS



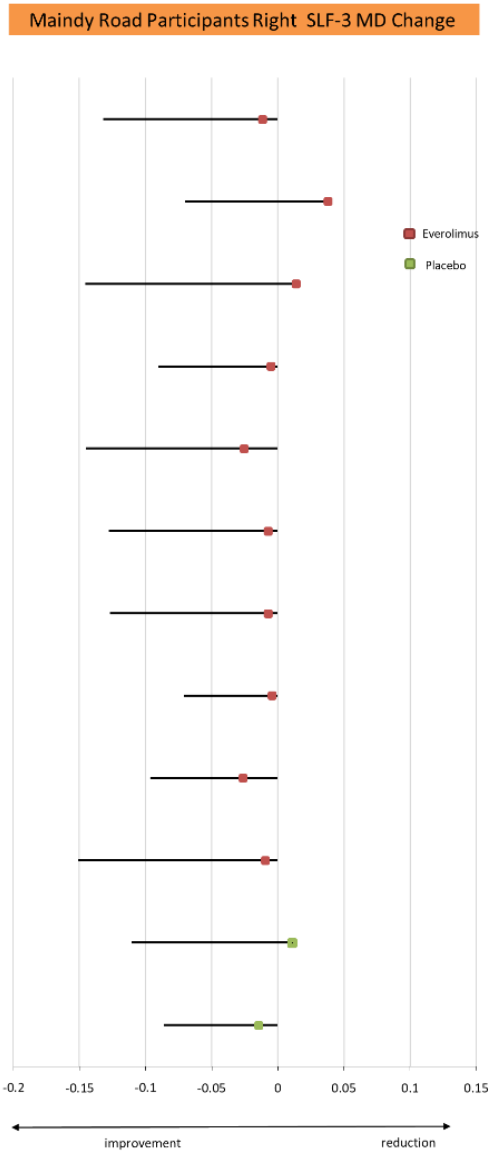


FIGURE 67 MD CHANGE FOR SLF-3 TRACT IN CUBRIC -2 PARTICIPANTS