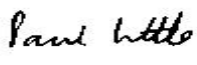
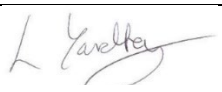
Active Brains

Statistical Analysis Plan

Trial name:	Active Brains
Trial registration number:	ISRCTN17349359
Protocol title and version number:	Active Brains Main Trial Protocol V9 23.09.24
SAP version number:	Version 1
SAP date:	31/10/2024

To be approved and reviewed by:

	Name	Signature	Date
Chief Investigator	Paul Little		18/11/24
	Lucy Yardley		4.11.2024
Trial Statistician (author)	Taeko Becque		31.10.2024
Senior Clinical Trial Statistician	Beth Stuart		15.11.2024

FUNDED BY
NIHR | National Institute for Health and Care Research

Table of Contents

1 Introduction

1.1 Purpose of Statistical Analysis Plan (SAP)

This statistical analysis plan (SAP) describes in detail the methods that will be used to analyse the data collected as part of the Active Brains trial. This will form the basis of the final trial publication. The final analysis will follow the SAP to ensure that the analyses are conducted in a scientifically valid manner and to avoid post hoc decisions which may affect the interpretation of the statistical analysis. Any deviations from the SAP will be detailed in the final report.

1.2 Trial background and rationale (short synopsis)

The prevalence of dementia is estimated to be between five and seven percent amongst those aged over 65 [1, 2] and the absolute number of cases is likely to increase due to people living longer [3]. It has been estimated that if interventions could delay both disease onset and progression by a modest 1 year, there would be nearly 9.2 million fewer worldwide cases of the disease in 2050, with nearly the entire decline attributable to decreases in persons needing a high level of care [3].

Cognitive impairment in the absence of dementia is more common, but prevalence estimates vary considerably depending on definitions [4]. Mild Cognitive Impairment (MCI) is the most commonly used definition, conventionally defined as a deterioration in at least one non-memory cognitive domain in addition to memory impairment, without severe functional impairment or loss of instrumental activities of daily living [4, 6]. An alternative way of characterising cognitive impairment is Age Associated Cognitive Decline (AACD). AACD is defined as one standard deviation (SD) below normal cognitive functioning in any cognitive domain, with some investigators having an additional criterion (self-report of a gradual decline in memory present for at least 6 months) [4].

It is now widely recognised that proactive management of modifiable risk factors can either delay onset or slow progression of dementia [5]. There is mounting evidence that healthy behaviour change (particularly physical activity) and cognitive exercises improve cognitive functioning and activities of daily living. Behavioural interventions are complex and resource intensive if delivered by purely face-to-face methods. In contrast, the internet is now used extensively and successfully by older people for self-management [6, 7]. Although many individuals may benefit from a free standing intervention (without additional support), the additional impact of behavioural facilitation may be important in helping initiate and maintain behaviour change [8, 9] - and the effectiveness and cost-effectiveness of a more intensive intervention may vary with the risk of developing dementia.

The primary aim of this study is therefore to assess the effectiveness and cost-effectiveness of a digital intervention (Active Brains) to help support older adults (60-85 years) with and without cognitive impairment in making healthy changes (physical activity, brain training, diet) to maintain cognitive function and prevent cognitive decline.

1.3 Objectives

Amongst older adults (60-85 years) both with and without cognitive impairment, the primary research objectives are at 1 and 5 years:

1. to estimate the clinical effectiveness and cost-effectiveness of the Active Brains digital intervention (compared with usual care) in maintaining cognitive functioning (assessed by verbal reasoning performance).

2. to estimate the clinical effectiveness and cost-effectiveness of the Active Brains digital intervention (compared with usual care) in maintaining cognitive functioning (assessed by verbal reasoning performance) when supported as necessary by a trained behavioural facilitator.

At the five-year follow-up point, a further primary objective is to estimate the effectiveness of the Active Brains digital intervention in delaying dementia diagnosis. This will be determined by data collected from patient notes review.

Secondary Research Objectives:

Secondary research objectives are to evaluate the effectiveness of the Active Brains digital intervention across a number of other key outcomes at 1 and 5 years:

1. Instrumental Activities of Daily Living

2. Additional cognitive outcomes (including spatial working memory, attention, and verbal short-term memory)

3. Health-related quality of life

4. Patient enablement

5. Depression

6. Changes in the target behaviours (physical activity (including sedentary time), healthy eating, engagement with brain training games)

7. Mortality data

1.4 Definition of endpoints

A complete list of measures is provided in table 1 at the end of the document. Participants in both the randomised trials and cohort study will be followed-up every 12 months for the five year duration of the study. At years one and five, participants will be asked to complete a full set of follow-up measures. In years two, three and four participants will only be requested to complete a subset comprising the most important outcome measures.

1.4.1 Definition of primary endpoint

The primary outcome is the Baddeley Verbal Reasoning score [10, 11] at 1 year.

The primary outcome at 5 years will be the diagnosis of dementia. The primary definition of dementia diagnosis will be based on diagnosis in the notes. Sensitivity analyses to this definition are provided in section 4.5.9.

1.4.2 Definition of secondary endpoints

Secondary outcomes will include the following at both 1 year and 5 years: (Further details and derivations are provided in section 4)

- Spatial Working Memory (measured using Self ordered search test)
- Digital Vigilance (attention - measured using the 'digit' span task [12])
- Verbal Short-term memory (measured using the paired associates learning [13])
- Modified version of Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE SF) [14]
- Instrumental Activities of Daily living (IADL) [15]
- Quality of Life (EQ5D) [16]
- 12-Item Short Form Health Survey (SF12) [17]
- Patient Enablement Scale [18]
- Brief Geriatric Depression Scale (GDS) [19, 20]
- Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS) [21]
- Diagnosis of dementia
- Death
- International Physical Activity Questionnaire (IPAQ) plus strength and balance items [22]
- Dietary behaviour (food frequency questionnaire, with the addition of 2 items about nuts and oils) [23]
- Brain training behaviours

1.4.3 Process measures

- Problematic Experiences of Therapy Scale (PETS) [24]
- Self-efficacy for exercise scale [25]
- mMOS Social Support Survey – 8 item [26, 27]
- Social Support for Exercise scale [28]
- Locus of Causation in Exercise [29]
- Technology Acceptance Model Perceived Ease of Use scale [30-32]
- Pedometer use/purchase
- Cognitive performance measures, IQCODE, IADL, EQ5D and FFQ at 2, 3, and 4 years.

1.5 Analysis principles

All analyses will be reported according to CONSORT 2010 on planning, implementing and reporting statistical analyses [33] and ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials [34].

2 Design considerations

2.1 Description of trial design

There are two parallel trials with the same trial interventions:

1) one trial for patients with cognitive decline (lower cognitive score)
2) a second trial for patients without signs of cognitive decline (higher cognitive score)

Participants will be randomly allocated to one of three study groups (ratio 1:1:1)

1. Usual care
2. Access to the Active Brains website
3. Access to the Active Brains website with flexible human support from a central support facilitator

2.2 Trial power and sample size

Cognitive decline trial

For the primary outcome of cognitive decline at 5 years we assume an alpha of 0.05 and 80% power to detect a standardised effect of 0.15. This requires 698 per group for the MCI subgroup and the same again for the AACD subgroup. Assuming that 60% will be available to provide follow-up within 5 years, that rises to 698*3/0.6=3490 participants with MCI and the same with AACD per group or 6980 in total.

Although the reduction in cognitive decline as assessed using the Baddeley reasoning test is the primary outcome at one year, the limiting factor in the sample size calculation for those with cognitive impairment is the progression to dementia at 5 years (see outcome definition in section 4.5.9). We assume a 5% annual progression rate to dementia among those who are cognitively impaired. With alpha 0.05 and 80% power, we aim to detect a difference of 5% after 5 years (20% intervention and 25% control) in those diagnosed with dementia. This means we would need 1094 in each of the MCI and AACD groups or a total of 2188. Assuming that 60% will be available to provide follow up in 5 years requires 3647 in each of the 3 study groups or 10940 in total.

Non-cognitive decline trial

We assume few of those who have no cognitive impairment at baseline will develop dementia. For the primary outcome of cognitive decline we assume an alpha of 0.05 and 90% power to detect a standardised mean difference of 0.1 in the Baddeley. This requires 2103 per group. Assuming that 60% will be available to provide follow-up within 5 years, that rises to 3505 per group or 10,515 in total.

Therefore:

- In patients with cognitive decline we will recruit a minimum of 3647 patients to each of the three study groups (10,940 total); there will be a mixture of patients with MCI and AACD.
- In patients without cognitive decline we will recruit a minimum of 3505 patients to each of the three study groups (10,515 total).

This means our total sample size across both trials will be a minimum of 21,455 participants in the trials. Since the interventions are identical in both trials we will not only be able to analyse the two trials separately but also as a combined data set.

2.3 Randomisation details

Patients will be randomised online using the study software, which performs simple randomisation and concealment. Once randomised, participants will be informed of their group allocation (they will also be emailed this information). If participants are in one of the treatment arms they will be taken directly to Active Brains where they can begin the intervention. The participants, supporters and study team are unblinded, but the statistician will remain blinded.

2.4 Timing of planned analyses

All primary and secondary analyses will take place when all patients have completed one year follow up. Five-year follow-up is contingent on evidence of effectiveness in changing behaviour and improving cognition at 1 year.

2.4.1 Interim analyses and early stopping

No interim analyses are planned

2.4.2 Stopping rules

No stopping rules planned

2.5 Final analysis

All primary and secondary analyses will take place when all patients have completed one year follow up. If there is evidence of effectiveness in changing behaviour and improving cognition at 1 year, analysis will also be undertaken at 5 year follow up.

3 Statistical considerations

3.1 Definition of analysis populations

The primary analysis will analyse the two trials (cognitive decline and non cognitive decline) separately. We will then explore as a secondary analysis whether both trial datasets can be combined for modelling, particularly for any pre-planned subgroup analyses.

3.1.1 Intention-to-treat analysis population

The intention to treat (ITT) population includes all randomised patients in the arm to which they were randomised regardless of intervention adherence. All summaries and analysis will be on the ITT population unless otherwise specified.

3.2 Analysis software

Analysis will be carried out using Stata version 17 or higher.

3.3 Methods for handling data

3.3.1 Withdrawal from trial

All data up until the point of patient withdrawal from the trial will be used in analyses unless the patient withdrew consent and does not wish for the data already collected prior to withdrawal to be used for the trial.

3.3.2 Missing data

Multiple imputation with chained equations will be used for the primary outcome and will be implemented in the mice package in Stata. The imputation model will include all variables in the analysis model and auxiliary variables, which predict the outcome or the missingness of the outcome. Predictors of the outcome or its missingness will be ascertained using LASSO regression [35]. Further details are provided in section 4.4. However, we will also include sensitivity analyses using delta-based multiple imputation, to explore the impact of a better or worse outcome for the missing values than those imputed under MAR. We will also present a sensitivity analysis using complete cases only.

3.3.3 Outliers

Outliers will be defined as any data point more than 3 standard deviations from the mean. If outliers are found in regression modelling, then firstly the source data will be checked. If the source data shows that the data is correct, then sensitivity analysis excluding the outliers will be conducted to explore any difference in inferences.

3.3.4 Assumption checking and alternative methods

Assumptions for linear regressions will be checked using QQ plots for normality and residual vs fitted value plots for linearity and homoscedasticity. If linear modelling assumptions are not met, the data will be transformed using, e.g., Box Cox transformation and back transformed at values of interest.

3.3.5 Data transformations

No further data transformations will be used.

3.4 Definition of key derived variables

Derivations of scores are provided for each endpoint in section 4.

3.5 General principles for reporting and analysis

Analyses will in general be reported using a significance level of 5%, corresponding to 95% confidence intervals unless otherwise stated. There will be no adjustment for multiplicity. Each active intervention arm will be compared to usual care. Descriptive statistics will be reported to 1 decimal place as number and percentage for categorical variables, and mean and standard deviation for continuous variables, or median and interquartile range for variables with a skewed distribution. The groups will be labelled Usual Care, Active Brains and Active Brains plus support.

4 Planned analyses and reporting

4.1 Disposition of the study population

CONSORT flow diagram (following CONSORT guidelines) which should include:

- Screening data – total number screened, reasons for not entering trial)
- Summary of eligibility data to be presented - total number assessed for eligibility, breakdown of patients screened against the eligibility criteria
- Recruitment information - number consented, recruited/randomised, receiving allowed treatment, withdrawing/lost to follow up at each time point
- Analysis population - number included, reasons excluded

4.2 Protocol deviations

A protocol deviation is an unanticipated or unintentional divergence or departure from the expected conduct of a study inconsistent with the protocol, consent documents or other study procedures. Of particular importance are major deviations (violations) which may expose participants to increased risk; compromise the integrity of the entire study or affect participant eligibility. Protocol deviations, as reported on the eCRF page, will be listed with information on treatment group and the type of deviation. Full details of the protocol deviations will also be listed. Failure to engage with the randomised intervention will not be taken as a protocol deviation and neither will use of other websites by participants in the control group.

4.3 Baseline and demographic characteristics

Baseline outcome measures are listed in Table 1 and include:

- Patient socio-demographic measures: age, date of birth, gender, index of multiple deprivation indices; internet experience, education level, work status, ethnicity, height, weight
- Clinical/behavioural measures collected from patient report: Health conditions and medications, family history of dementia, smoking status, alcohol intake, BMI, potential dementia status.

4.4 Primary endpoint

The Baddeley verbal reasoning score at 1 year will be analysed using linear regression, adjusting for baseline Baddeley verbal reasoning score and age, gender, IMD score, number of comorbidities and polypharmacy (yes/no). Adjustment for prognostic covariates, specified a priori, is now commonly recommended in the primary analysis of trial data and can potentially increase power [36, 37]. Assumptions for linear regression will be checked using QQ plots for normality and residual vs fitted value plots for linearity and homoscedasticity. If linear modelling assumptions are not met, the data will be transformed using, e.g., a Box Cox transformation and back transformed at values of interest.

The primary analysis will use multiple imputation with chained equations. The imputation model will include all variables in the analysis model and any baseline covariates predictive of

the Baddeley verbal reasoning score or its missingness. Predictors of missingness will be ascertained using Lasso regression [35]. Potential auxiliary variables will include all baseline variables. A model for outcome that includes randomised group, baseline outcome and all potential auxiliary variables will be fitted using the LASSO with 10-fold cross-validation [38]. The regularization penalty (denoted by λ) will be chosen such that the cross-validation error was within one SE of the minimum [39]. The auxiliary variables that are included in the fitted model from the LASSO will be included in the imputation model for the outcome, i.e., the LASSO will used to select variables for the imputation model but not to estimate parameters of the imputation model. One hundred imputations will be performed and the imputed estimates will be combined using Rubin's rules. Imputation will be performed separately by randomised group and then combined [40].

All tests will be two-tailed with point estimates and 95% confidence intervals for the treatment effect presented. No formal adjustment for multiple significance testing will be applied. Analyses will be performed using Stata version 17 or above.

4.4.1 Key subgroup analysis

The key subgroup is the MCI subgroup, defined as 1.5 SD below the norm on the Baddeley reasoning test. Subgroup analysis will be performed by repeating the primary analysis and including a treatment by covariate interaction term for the MCI subgroup.

Further exploratory subgroups analyses are outlined in section 4.6.1.

Sensitivity analyses for the definition of cognitive impairment, for the multiple imputation and for the order of questions are outlined in section 4.6.2.

4.4.2 Adherence to intervention

To explore the effect of adherence to the randomised intervention, we will estimate the treatment effect among the subpopulation of participants would adhere with the intervention under either treatment assignment. The estimand is defined in Table 2, and adherence with the intervention is defined as:

1) Completing the first session of Active Brains as far as the pages that advise on what to do (Active Lives homepage)

AND

2) Adhering to at least some of the recommended behaviours, defined as answering 'less than a month' or more of at least one of the Getting Active/Strength and Balance/Breaks from Sitting/Eat for Health recommendations.

OR

3) Accessed Brain Training more than once

Table 2. Intercurrent events and estimands

Population	All trial participants eligible at baseline
Treatment	Active Brains + Support vs Active Brains vs usual care
Endpoint	Baddeley verbal reasoning score at 1 year
Summary measure	Mean difference for intervention vs usual care
Intercurrent event	**Strategy for handling intercurrent event**
Did not complete core sections of Active Brains intervention	Principal stratum - subpopulation of patients who would complete core sections of the intervention under either treatment assignment

This will be estimated using a Complier Average Causal Effect (CACE) analysis. The CACE analysis will be carried out using instrumental variables regression, with randomised group as the instrumental variable, and adherence defined above as the endogenous variable.

4.5 Secondary endpoints

The following secondary outcomes will be analysed at 1 and 5 years:

4.5.1 Key secondary endpoint

Change in the proportions of participants to a different category:

- from no cognitive impairment to cognitive impairment (AACD or MCI),
- from AACD to MCI or severe cognitive impairment,
- from MCI to severe cognitive impairment.

The Baddeley verbal reasoning task will be used to indicate AACD (defined as scores more than 1 SD below the normative score), MCI (defined as scores more than 1.5 SD below the normative score), and severe cognitive impairment (defined as scores more than 2 SD below the normative score). The normative score is determined by PROTECT database, a large (n>15,000) cohort of older adults' data on battery of cognitive assessment tasks (https://www.protectstudy.org.uk/)

4.5.2 Cognitive performance tests

Cognitive performance tests include:

- Spatial Working Memory (measured using Self ordered search test)
- Digital Vigilance (attention - measured using the 'digit' span task)
- Verbal Short-term memory (measured using the paired associates learning)

These are continuous scores and will be analysed in the same way as the primary outcome, adjusting for the baseline test score and the same covariates as in the primary analysis.

4.5.3 Informant Questionnaire on Cognitive Decline in the Elderly

Modified version of Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE SF) consists of 16 questions with a Likert response ranging from 1=much improved to 5=much worse. An overall change score is calculated by averaging the scores on each item. This score will be analysed using linear regression adjusting for baseline IQCODE score and the same covariates as in the primary analysis.

4.5.4 Patient Enablement Index (PEI)

Patient Enablement Index (PEI) consists of six questions and produces a mean score ranging from 1 to 7 with lower scores indicating higher enablement. The PEI score will be analysed using linear regression adjusting for the same covariates as in the primary analysis.

4.5.5 Instrumental Activities of Daily Living (IADL)

Instrumental Activities of Daily Living (IADL) consists of 7 questions, each with two parts, about managing day to day activities. The first part of each question rates how well these activities can be managed 0=On my own; 1=With some help; 2=With full help; 3 = Could not do it (so it wasn't done at all or someone did it for me); 8=Not relevant to me (always done by someone else or not needed). The second part of each question rates how difficult it is (or would be) to do the activity with 0=No difficulty; 1=Some difficulty; 2=Great difficulty. The total score is the sum of all the subscales, creating a score from 0 to 35. The total IADL score will be analysed using linear regression, adjusting for the same covariates as in the primary analysis.

4.5.6 International Physical Activity Questionnaire (IPAQ)

International Physical Activity Questionnaire (IPAQ) plus strength and balance items

MET minutes achieved in each category (walking, moderate activity and vigorous activity) and total MET minutes of physical activity a week are calculated as follows:

- Walking MET-minutes/week = 3.3 * walking minutes * walking days
- Moderate MET-minutes/week = 4.0 * moderate-intensity activity minutes * moderate days
- Vigorous MET-minutes/week = 8.0 * vigorous-intensity activity minutes * vigorous-intensity days
- Total physical activity MET-minutes/week = sum of Walking + Moderate + Vigorous MET-minutes/week scores.

Note: Bouts of activity lasting less than 10 minutes duration are not counted. It is also recommended that activity bouts of greater than 3 hours are truncated, i.e., in each category a maximum of 21 hours of activity are permitted a week (3 hours X 7 days).

Total MET minutes will be analysed using linear regression adjusting for baseline MET minutes and the same covariates as in the primary analysis.

4.5.7 Quality of Life (EQ5D)

Quality of Life (EQ5D) EQ-5D-5L index values will be derived using the appropriate value set. These utility scores will be analysed using linear regression adjusting for baseline EQ5D score and the same covariates as in the primary analysis.

4.5.8 Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS)

Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS) is a 7-item scale on a Likert scale 1=none of the time to 5=all of the time. It is scored by first summing the scores for each of the seven items. The total raw scores are then transformed into metric scores using the SWEMWBS conversion table which can be found here:

https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto/swemwbs_raw_score_to_metric_score_conversion_table.pdf

Interpretation: Scores range from 7 to 35 and higher scores indicate higher positive mental wellbeing. To see how individual's scores compare with national survey data (from adults) which can be found here: https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto/wemwbs_population_norms_in_health_survey_for_england_data_2011.pdf

The metric scores will be analysed using linear regression adjusting for baseline SWEMWBS score and the same covariates as in the primary analysis.

4.5.9 12-Item Short Form Health Survey (SF12)

The 12-Item Short Form Health Survey (SF12) is a self-reported quality of life measure assessing the impact of health on an individual's everyday life. Two summary scores are reported from the SF-12 – a mental component score (MCS-12) and a physical component score (PCS-12). The scores may be reported as Z-scores (difference compared to the population average, measured in standard deviations). The population average PCS-12 and MCS-12 are both 50 points. The population standard deviation is 10 points, so each 10 increment of 10 points above or below 50, corresponds to one standard deviation away from the average. The MCS-12 and PCS-12 scores will be analysed using linear regression, adjusting for the baseline scores and the same covariates as in the primary analysis.

4.5.10 Diagnosis of dementia

Diagnosis of dementia will be obtained from notes review. This will be analysed using logistic regression, adjusting for the same covariates as in the primary analysis.

The primary definition of dementia will be based on:

1. If there is a notes diagnosis of dementia participants will be classed as having dementia.

As sensitivity analyses, we will use the following methods to ascertain the dementia outcome:

2. As recording in notes is suboptimal, consider self-report item "Has a doctor or other medical worker ever said that you have dementia?" *(Y/N)* - this data currently collected in baseline and year 1 and year 5 follow-up online/paper measures.

3. Evidence of significant decline in Baddeley (baseline to follow-up) **AND** evidence of functional deterioration (**EITHER**: a) significant change in self-report IQCODE (baseline to follow-up) **OR** b) two self-report questions on difficulty in finance and planning* (NB the Caerphilly study was based on informant not self-report). Difficulty in finance and planning items collected in year 5 online/paper measures.

* The items are:

Thinking about how you were ten years ago, how are you at:
a) handling your personal finances (banking, pension, etc)?
b) thinking about future events and planning ahead?
(Much improved/ A bit improved/ Not much change/ A bit worse/ Much worse)

4. Self-report of difficulty in finance and planning alone (since these questions had an AUROC of 0.9 – with the caveat about self-report).

5. Multiple imputation of the missing dementia outcome will be carried out following the same approach as in the primary analysis section 4.4. Sensitivity analyses to the imputation will include assuming all missing have/do not have dementia.

4.5.11 Death

Death will be obtained from notes review, or reported by the practice, supporter or relative. This will be analysed using logistic regression, adjusting for the same covariates as in the primary analysis.

4.5.12 Dietary behaviour

Dietary behaviour (food frequency questionnaire, with the addition of 2 items about nuts and oils). A prudent pattern score is calculated for each participant as follows:

i) fat spreads and milks are categorised as full fat or reduced fat versions (reduced fat spreads <69g fat/100g, milks <3.5g fat/100g),
ii) weekly frequencies of consumption are calculated as: never=0, <1/month=0.2, 1-3/month=0.5, 1/ week=1, 2-4/week=3, 5-6/week=5.5, 1/day=7, 2-3/day=17.5, 4-5/day=31.5, ≥6/day=42,
iii) food variables are standardised by subtracting the means and dividing by the SDs for the Hertfordshire Cohort Study (HCS) population,
iv) the coefficient for each food (Table 2 of [38]) was multiplied by the standardised food variable,
v) these values are summed resulting in one score for each subject.

The score indicates the participant's compliance with the prudent pattern and is interpreted as a marker of their diet quality (see [41] for further interpretation).

This score will be analysed using linear regression adjusting for baseline prudent pattern score and the same variables as in the primary analysis. The additional questions on consumption of nuts and oils will be summarised descriptively by randomised arm.

4.5.13 Brain training behaviours

Brain training behaviours will be summarised descriptively by randomised arm.

4.5.14 Problematic Experiences of Therapy Scale (PETS)

Problematic Experiences of Therapy Scale (PETS) The scale comprises 12 items divided into four subscales: ‘‘symptoms too severe or aggravated by therapy’’ (items 1–3), ‘‘uncertainty about how to carry out the treatment’’ (items 4–5), ‘‘doubts about treatment efficacy’’ (items 6–8), and ‘‘practical problems’’ such as lack of time or opportunity, forgetting (items 9–12). All items are scored on a scale ranging from 1 (disagree strongly) to 5 (agree strongly). The scores will be recoded into binary categories as follows: participants who respond ‘‘strongly disagree’’ to all items in a subscale are recoded as ‘‘no barriers’’, and all other scores are recoded as ‘‘some barriers or doubts’’. For each subscale, the categories will be summarised descriptively by intervention arm.

4.5.15 Brief Geriatric Depression Scale (GDS-4)

Brief Geriatric Depression Scale (GDS-4) scores one point for each answer indicating depression (see codebook). Interpretation: 0 = Not Depressed; 1 = Uncertain; 2 to 4 = Depressed. This will be analysed using logistic regression with categories depression vs not depressed/uncertain, adjusting for baseline depression and the same variables as in the primary analysis.

4.5.16 Self-efficacy for exercise scale (SESS)

Self-efficacy for exercise scale consists of 9 items scored from 0 (not confident) to 10 (very confident). A total score is calculated by summing the responses to each question, giving a total score from 0-90. A higher score indicates higher self-efficacy for exercise. This score will be analysed using linear regression adjusting for baseline SESS and the same variables as in the primary analysis.

4.5.17 Social Support Survey (mMOS)

mMOS Social Support Survey – 8 item includes 8 items on instrumental and emotional subscales.The mMOS-SS is scored as the average score of subscale items transformed to a zero to 100 scale, with higher scores indicating more support. See http://www.rand.org/health/surveys_tools/mos/mos_socialsupport_scoring.html
This score will be analysed using linear regression adjusting for baseline mMOS score and the same variables as in the primary analysis.

4.5.18 Social Support for Exercise scale

Social Support for Exercise scale is a 13-item scale with responses on a 5-point Likert scale how often (1, *none* to 5, *very often*). An average social support score is calculated for friends, family, and combined friends and family, where higher scores indicate greater social support for physical activity and exercise. The score will be summarised descriptively by intervention arm.

4.5.19 Locus of Causation in Exercise (LCE)

Locus of Causation in Exercise consists of three items, with responses ranging from 1 (strongly disagree) to 7 (strongly agree). To score the LCE, reverse scores on items two and three and then calculate the mean for the three items. High scores indicate greater self-determination or a more internal perceived locus of causality and low scores less self-determination. The score will be analysed using linear regression adjusting for baseline LCE and the same variables as in the primary analysis.

4.5.20 Technology Acceptance Model Perceived Ease of Use scale

Technology Acceptance Model Perceived Ease of Use scale consists of six items on a 1 (extremely likely) to 7 (extremely unlikely) Likert scale. The total score is obtained by taking the average score and transforming to a 0 to 100 scale as follows:
PEU = (AVERAGE(TAM07, TAM08, TAM09, TAM10, TAM11, TAM12) – 1)(100/6)
The score will be summarised descriptively by intervention arm.

4.5.21 Pedometer use

Pedometer use/purchase will be summarised descriptively by randomised arm.

4.6 Additional analyses

4.6.1 Subgroup analyses

For the primary outcome at 1 year, exploratory subgroup analyses will be performed by repeating the primary analysis and including a treatment by covariate interaction term for the following baseline subgroups:

- Age: above/below 75 years
- Gender: male/female
- Deprivation: IMD more deprived (deciles 1-5) vs less deprived (deciles 6-10)
- Comorbidities: < 2 vs ≥ 2
- Polypharmacy: number of medications ≥ 5 vs ≤ 4
- Health diet: high/low prudent diet score (FFQ)
- Physical activity: high/low IPAQ score
- Depression: depression vs no depression, based on GDS at baseline
- Retirement: retired/working

The same subgroups will be analysed in the same way using the primary outcome at 5 years.

4.6.2 Sensitivity analyses

1. Sensitivity analyses to the primary analysis multiple imputation will include:
   - Delta-based multiple imputation, assuming missing verbal reasoning scores are on average 0.15 points (on a standardised scale) better or worse than the MAR imputed values of verbal reasoning score.
   - complete cases analysis

2. Sensitivity analyses for the definition of cognitive impairment will include:
   - MCI defined as 1.5 SD below the norm in another non-memory cognitive domain plus memory impairment.
   - AACD defined as 1 SD below the norm for the Baddeley reasoning test and IADL.
   - Decrements in the other computerised tests (spatial working memory, digital vigilance and verbal short term memory) combined with IQCODE in indicating a diagnosis of dementia.

3. A sensitivity analysis to check for any differences resulting from changing the order of the Baddeley verbal reasoning task from the third task to the first task will be also carried out.

4.6.3 Further analyses

The health economic analysis plan and the process evaluation plan will be outlined in separate documents. The cognitive and behavioural outcomes at 2, 3, 4 years will be analysed in the process evaluation.

4.7 Safety reporting

For this study Serious Adverse Event (SAE) is defined as any untoward medical occurrence or effect that at any dose:

- Results in death
- Is life-threatening – refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe
- Requires hospitalisation, or prolongation of existing inpatients' hospitalisation
- Results in persistent or significant disability or incapacity
- Other important medical events - based upon appropriate medical judgment; they may jeopardise the participant and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

SAEs will be summarised descriptively by severity and relatedness. No SAEs are considered expected for the purposes of this trial. Occurrence of SAEs among all randomised participants will be analysed using logistic regression.

5 Tables, listings and figures templates

Table 1 – List of Measures

Measure		Baseline				Year 1 and Year 5 follow-up				Year 2, year 3 and Year 4 follow-up			
		Usual	Active Brains	AB + Support	Cohort group	Usual	Active Brains	AB + Support	Cohort group	Usual	Active Brains	AB + Support	Cohort group
Patient Socio Demographic measures: (age, date of birth, gender, postcode (to estimate deprivation indices); internet experience, education level, work status, ethnicity, height, weight)		X	X	X	X								
Clinical/behavioural measures collected from patient report: Health conditions and medications, family history of dementia [b], smoking status, alcohol intake, BMI, potential dementia status.		X	X	X	X	X	X	X	X				
Self-reported respiratory infections						X	X	X	X				
COVID-19 and long-COVID status		X[c]	X[c]	X[c]	X[c]	X	X	X	X	X	X	X	X
Out of pocket expenditure						X	X	X	X				
Clinical measures from notes review (comorbidities, systolic blood pressure, cholesterol, frailty score)						X*	X*	X*					
Cognitive Performance Tasks	**Primary Outcome:** Reasoning (measured using the Baddeley reasoning test)	X	X	X	X	X	X	X	X	X	X	X	X
	Spatial Working Memory (measured using the Self ordered search test)	X	X	X	X	X	X	X	X	X	X	X	X

	Digital Vigilance (attention - measured using the 'digit' span task)	X	X	X	X	X	X	X	X	X	X	X	X
	Verbal Short-term memory (measured using the paired associates learning)	X	X	X	X	X	X	X	X	X	X	X	X
Modified version of IQCODE SF (to allow self-report)		X	X	X	X	X	X	X	X	X	X	X	X
Instrumental Activities of Daily living		X	X	X	X	X	X	X	X	X	X	X	X
The patient Enablement Scale						X	X	X	X				
International Physical Activity Questionnaire (IPAQ) plus strength and balance items		X	X	X	X	X	X	X	X	X	X	X	X
Quality of Life (EQ5D)		X	X	X	X	X	X	X	X	X	X	X	X
SWEMWBS		X	X	X	X	X	X	X	X				
SF12		X	X	X	X	X	X	X	X				
Diagnosis of dementia						X*	X*	X*					
Mortality						X*	X*	X*					
Health economic analysis of cost effectiveness will require resource usage (meds, consultations, hospitalisation, A&E attendance, outpatient visits)						X*	X*	X*					
Personal resource usage						X	X	X					
Dietary behaviour (food frequency questionnaire, with the addition of 2 items about nuts and oils)		X	X	X	X	X	X	X	X	X	X	X	X
Brain training behaviours						X	X	X	X	X	X	X	X
Patient rated adherence to changes in behaviour and list of barriers they felt contributed to their adherence (measured with the PETS questionnaire)							X	X	X				
Brief Geriatric Depression Scale		X	X	X	X	X[1]	X[1]	X[1]	X[1]				
Self-efficacy for exercise scale		X	X	X	X	X	X	X	X				
mMOS Social Support Survey – 8 item		X	X	X	X	X	X	X	X				

Social Support for Exercise scale						X	X	X				
Locus of Causation in Exercise	X	X	X	X	X	X	X	X				
Technology Acceptance Model Perceived Ease of Use scale						X	X	X				
Pedometer use/purchase					X	X	X	X				
Objective patient data												
Usage of the Active Brains website throughout (including what components viewed and any data entered online such as at goal reviews).		X	X	X		X	X	X		X	X	X
Objective supporter data												
Supporters' Usage of Active Brains website (throughout study)			X				X				X	
Emails sent to participants throughout study			X				X				X	
Qualitative data												
Interviews with patients about their experiences of the study and/or intervention					X[a]	X[a]	X[a]	X[a]				
Interviews with Central Support Facilitators about their experiences of the study and intervention							X					
Proxy-Measures (not cohort group)												
	Baseline				Years 1 and 5				Years 2-4			
Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (short form) [53]							X					
Quality of Life (EQ5D5L-proxy version 2) to be completed by the nominated contact person (if necessary)							X					
IADL proxy version to be completed by the nominated contact person (if necessary)							X					

*=notes a measure collected at notes review. [1] = collected at year one only (not year five also)
[a] = process interviews to be conducted 2-12 months after commencement of study; [b] = baseline only;
[c] = COVID-19 questions only (not long-COVID) at baseline

6 References

1. Prince, M., et al., *The global prevalence of dementia: A systematic review and metaanalysis.* Alzheimer's & Dementia, 2013. **9**(1): p. 63-75.e2.
2. Matthews, F.E., et al., *A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II.* Lancet, 2013. **382**(9902): p. 1405-1412.
3. Brookmeyer, R., et al., *Forecasting the global burden of Alzheimer's disease.* Alzheimers Dement, 2007. **3**(3): p. 186-91.
4. Stephan, B., et al., *Early cognitive change in the general population: how do different definitions work?* Journal of the American Geriatrics Society, 2007. **55**(10): p. 1534-1540.
5. Organization, W.H., *Risk reduction of cognitive decline and dementia: WHO guidelines*, in *Risk reduction of cognitive decline and dementia: WHO guidelines*. 2019.
6. Aalbers, T., M.A.E. Baars, and M.G.M.O. Rikkert, *Characteristics of effective Internet-mediated interventions to change lifestyle in people aged 50 and older: A systematic review.* Ageing Research Reviews, 2011. **10**(4): p. 487-497.
7. Stellefson, M., et al., *Web 2.0 Chronic Disease Self-Management for Older Adults: A Systematic Review.* J Med Internet Res, 2013. **15**(2): p. e35.
8. Dennison, L., et al., *Does brief telephone support improve engagement with a web-based weight management intervention? Randomized controlled trial.* Journal of medical Internet research, 2014. **16**(3): p. e95.
9. Riper, H., et al., *Effectiveness of guided and unguided low-intensity internet interventions for adult alcohol misuse: a meta-analysis.* PLoS One, 2014. **9**(6): p. e99912.
10. Baddeley, A.D., *A 3 min reasoning test based on grammatical transformation.* Psychonomic Science, 1968. **10**(10): p. 341-342.
11. Wild, K., et al., *Status of computerized cognitive testing in aging: a systematic review.* Alzheimer's & Dementia, 2008. **4**(6): p. 428-437.
12. Conklin, H.M., et al., *Verbal working memory impairment in schizophrenia patients and their first-degree relatives: evidence from the digit span task.* American Journal of Psychiatry, 2000. **157**(2): p. 275-277.
13. Owen, A.M., et al., *Visuospatial memory deficits at different stages of Parkinson's disease.* Neuropsychologia, 1993. **31**(7): p. 627-644.
14. Jorm, A.F., *A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): development and cross-validation.* Psychol Med, 1994. **24**(1): p. 145-53.
15. Ashford, S., S. Brown, and L. Turner-Stokes, *Systematic review of patient-reported outcome measures for functional performance in the lower limb.* Journal of rehabilitation medicine, 2015. **47**(1): p. 9-17.
16. Pickard, A.S., et al., *Health utilities using the EQ-5D in studies of cancer.* Pharmacoeconomics, 2007. **25**(5): p. 365-384.
17. Ware Jr, J.E. and C.D. Sherbourne, *The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection.* Medical care, 1992: p. 473-483.
18. Little, P., et al., *Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain.* Bmj, 2008. **337**: p. a884.
19. Almeida, O.P. and S.A. Almeida, *Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV.* International journal of geriatric psychiatry, 1999. **14**(10): p. 858-865.
20. Lesher, E.L. and J.S. Berryhill, *Validation of the geriatric depression scale-short form among inpatients.* Journal of clinical psychology, 1994. **50**(2): p. 256-260.
21. Haver, A., et al., *Measuring mental well-being: A validation of the Short Warwick–Edinburgh Mental Well-Being Scale in Norwegian and Swedish.* Scandinavian Journal of Public Health, 2015. **43**(7): p. 721-727.

22. Hurtig-Wennlof, A., M. Hagstromer, and L.A. Olsson, *The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility.* Public Health Nutr, 2010. **13**(11): p. 1847-54.
23. Little, P., et al., *The validity of dietary assessment in general practice.* Journal of Epidemiology & Community Health, 1999. **53**(3): p. 165-172.
24. Kirby, S., M. Donovan-Hall, and L. Yardley, *Measuring barriers to adherence: validation of the problematic experiences of therapy scale.* Disability and rehabilitation, 2014. **36**(22): p. 1924-1929.
25. Resnick, B. and L.S. Jenkins, *Testing the Reliability and Validity of the Self-Efficacy for Exercise Scale.* Nursing Research, 2000. **49**(3): p. 154-159.
26. Sherbourne, C.D. and A.L. Stewart, *The MOS social support survey.* Social science & medicine, 1991. **32**(6): p. 705-714.
27. Moser, A., et al., *The eight-item modified Medical Outcomes Study Social Support Survey: psychometric evaluation showed excellent performance.* Journal of clinical epidemiology, 2012. **65**(10): p. 1107-1116.
28. Sallis, J.F., et al., *The development of scales to measure social support for diet and exercise behaviors.* Preventive medicine, 1987. **16**(6): p. 825-836.
29. Markland, D. and L. Hardy, *On the factorial and construct validity of the Intrinsic Motivation Inventory: Conceptual and operational concerns.* Research Quarterly for exercise and Sport, 1997. **68**(1): p. 20-32.
30. Venkatesh, V. and F.D. Davis, *A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies.* Management Science, 2000. **46**(2): p. 186-204.
31. Davis, F.D., R.P. Bagozzi, and P.R. Warshaw, *User acceptance of computer technology: a comparison of two theoretical models.* Management science, 1989. **35**(8): p. 982-1003.
32. Davis, F.D., *Perceived usefulness, perceived ease of use, and user acceptance of information technology.* MIS quarterly, 1989: p. 319-340.
33. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332. PMID: 20332509
34. *https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.p*df
35. Mainzer, R. M., Nguyen, C. D., Carlin, J. B., Moreno-Betancur, M., White, I. R., & Lee, K. J. (2024). A comparison of strategies for selecting auxiliary variables for multiple imputation. *Biometrical Journal*, 66, 2200291. https://doi.org/10.1002/bimj.202200291
36. Morris, T.P., Walker, A.S., Williamson, E.J. *et al.* Planning a method for covariate adjustment in individually randomised trials: a practical guide. *Trials* **23**, 328 (2022). https://doi.org/10.1186/s13063-022-06097-z
37. Kahan, B.C., Jairath, V., Doré, C.J. *et al.* The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials* **15**, 139 (2014). https://doi.org/10.1186/1745-6215-15-139
38. Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22.
39. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021) An introduction to statistical learning (2nd ed.). Springer.
40. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? *Statistical Methods in Medical Research*. 2018;27(9):2610-2626. doi:10.1177/0962280216683570
41. Robinson, S., Crozier, S., Borland, S. *et al.* Impact of educational attainment on the quality of young women's diets. *Eur J Clin Nutr* **58**, 1174–1180 (2004). https://doi.org/10.1038/sj.ejcn.1601946

SAP revision history

Version number	Revision history	Author	Date