Statistical Analysis Plan for

DOMINO-HD: MULTI-DOMAIN LIFESTYLE TARGETS FOR IMPROVING PROGNOSIS IN HUNTINGTON'S DISEASE

ISRCTN Number:	12103732	Version Number:	1.0 (18/07/2023)

Final Plan

Based on protocol version: 2.1, 21/05/2021

SAP Revision History				
Protocol version	Updated SAP version number	Section number changed	Description and reason for change	Date changed

ROLES AND RESPONSIBILITIES

Trial Statistician: Myrsini Gianatsi			
Role: SAP reviewer			
Date:		Signature:	
Senior Statistician: Dr Philip Pallmann			
Role: SAP reviewer			
Date:	18/07/2023	Signature:	See email dated 18/07/2023
Chief Investigator: Prof Monica Busse			
Role: SAP reviewer			
Date:		Signature:	
Other non-signatory contributor to the SAP: Andreas Markoulidakis			
Role: SAP author			
Other non-signatory contributors to the SAP: Dr Beth Ann Griffin, Prof Peter Holmans, Dr Cheney Drew			
Role: SAP reviewers			

TABLE OF CONTENTS

1. INTRODUCTION

This statistical analysis plan provides guidelines for the final analysis and presentation of results for the DOMINO-HD study. This plan, along with all other documents relating to the analysis of this study, will be stored in the Statistical Analysis Master File electronically. Location:

S:/Centre for Trials Research/Research/Mixed Studies/DOMINO/eTMF/8.0 Data Management & Statistics/8.5 Statistics/8.5.1 Statistical Analysis Plan

2. BACKGROUND

2.1 RATIONALE AND RESEARCH QUESTION

There are not many studies assessing the combined impact of lifestyle factors such as diet alongside direct longitudinal measures of physical activity and sleep on the progression and severity of HD, and none which have attempted to combine detailed genetic information with these modifiable factors.

DOMINO-HD is a 12-month observational study which aims to recruit participants from five clinical sites across Europe, with a combined target of 300-450 participants (at least 60, but no more than 90, participants per site). It utilises a digital platform developed during phase 1 of the study and a range of self-reported lifestyle questionnaires and wearable activity trackers (see the study protocol for more details) to gather phenotypical lifestyle data. As all participants are recruited from Enroll-HD (a global research platform), we will establish the feasibility of linking the DOMINO-HD lifestyle dataset to Enroll-HD clinical outcomes and where possible genome-wide association study (GWAS) data to explore their interplay with HD symptom severity. Ultimately, we aim to use these data to inform predictive outcome modeling and development of interventions aiming to slow the progression of HD or at the very least minimise functional impact of the disease.

2.1.1 Enroll-HD

Enroll-HD is a worldwide research platform which serves as the basis of a longitudinal observational study of HD families and provides an infrastructure to support clinical trials. Enroll-HD data are available to every researcher working with HD on reasonable request. All DOMINO-HD participants are also required to have consented to participation in Enroll-HD. Data from the annual Enroll-HD observational assessments will be merged with data obtained from the DOMINO-HD lifestyle assessments, including specific dietary questionaries and wearable activity tracker data to form a comprehensive lifestyle and clinical dataset in this population.

2.2 OBJECTIVES

The primary objective (and the focus of this statistical analysis plan) is to establish the feasibility of linking lifestyle factors and genetic risk factors to explore their interplay with HD symptom progression and severity.

3. STUDY MATERIALS

3.1 STUDY DESIGN

A consortium-led observational, prospective cohort study examining the effect of long-term (12 months) lifestyle factors (diet, sleep, physical activity), combined with genomic data, on the progression of HD symptoms. Participants will be recruited from Enroll-HD.

3.2 RANDOMISATION

Not applicable as DOMINO-HD is an observational study.

3.3 SAMPLE SIZE

Based on previous clinical trials with similar assessments and participant burden (see the protocol for more details), it is estimated that a minimum of 60 participants (and up to 90) per site will be required. This will result in a total of 300-450 participants across the five clinical sites. With a sample size of 300, conservative assumptions about the impact of the propensity score weighting, and a type I error rate of 0.05, we will have 80% power to detect effect sizes that are at least as large as 0.23 for each environmental measure when they are modelled as continuous predictors. This represents the minimum change in the HD composite outcome that would be detected as statistically significant with 80% power for a 1-unit change in the

environmental measure, suggesting the study has reasonable power for small to moderate effects.

3.4 FRAMEWORK

In people with HD, we will assess how lifestyle factors are related to the progression and severity of the disease.

3.5 INTERIM ANALYSES

There are no interim analyses planned.

3.5.1 PLANNED SAMPLE SIZE ADJUSTMENT

Not applicable.

3.5.2 STOPPING RULES

Not applicable.

3.6 TIMING OF FINAL ANALYSIS

Outcomes will be analysed once the final dataset becomes available.

3.7 TIMING OF OUTCOME ASSESSMENT

Outcome measures will be assessed at their corresponding time point (i.e., baseline at the time of recruitment, follow-up at the end of the 12-month period).

4. STATISTICAL PRINCIPLES

4.1 LEVELS OF CONFIDENCE AND P-VALUES

95% confidence intervals will be reported. All p-values will be compared to 5% level of significance.

4.1.1 ADJUSTMENT FOR MULTIPLICITY

None due to the exploratory nature of the study.

4.2 PROTOCOL DEVIATIONS

4.2.1 DEFINITION OF PROTOCOL DEVIATION

See protocol for full details. Examples of deviations include errors in applying inclusion/exclusion criteria and missed follow-up assessments due to error.

4.2.2 PRESENTATION OF PROTOCOL DEVIATIONS

These will be tabulated, with reasons (if given).

4.3 ANALYSIS POPULATION

All participants with available outcome data will be included in the primary analysis. As part of sensitivity analyses, a complete-case analysis (excluding participants with incomplete data) will be performed.

5. STUDY POPULATION

5.1 SCREENING DATA

Eligibility data: potential participants will be identified from registered Enroll-HD participants. If an individual is considered suitable for DOMINO-HD, but is not registered in Enroll-HD, it is acceptable to be recruited to both at the same time.

Informed consent: those meeting eligibility criteria and providing informed consent will be registered as study participants and assigned a unique study identification number prior to baseline data collection.

Safety data: the study is an observational study where no participant will receive any kind of intervention. We do not expect adverse events and there is no need for the participants to undergo any safety assessments.

5.2 ELIGIBILITY

Numbers eligible (as a proportion of those screened), consented (as a proportion of those eligible), and followed up (as a proportion of those consented) will be reported descriptively for the entire study and by site.

5.3 RECRUITMENT

Participants will be recruited from five current Enroll-HD sites across Europe. Up to 90 participants will be recruited from each site (minimum of 60).

5.4 WITHDRAWAL/FOLLOW UP

5.4.1 LEVEL OF WITHDRAWAL

Participants have the right to withdraw consent for participation in any aspect of the study at any time. If a participant initially consents but subsequently withdraws from the study, clear distinction must be made as to what aspect of the study the participant is withdrawing from.

Data collected prior to participant withdrawal will be used, unless the participant specifically asks that they are not.
These aspects could be:

1. Partial withdrawal from further data collection (e.g., some of sample collection, questionnaires, clinical assessments)
2. Complete withdrawal from further data collection
3. Withdrawal of permission to use data already collected

A participant will be deemed lost to follow-up if they do not attend the 12-month assessment, within a period of 8 weeks of the scheduled assessment (adhering to the Enroll-HD protocol). As all participants will be participants of Enroll-HD, we expect to obtain a proportion of 12-month assessment data for all participants unless they are lost to follow-up in the Enroll-HD study.

5.4.2 TIMING OF WITHDRAWAL

The timing and level of any withdrawals will be tabulated.

5.4.3 REASONS FOR WITHDRAWAL

Due to the participants' involvement in Enroll-HD, we do not anticipate a high rate of loss of follow-up from the study. Potential reasons for withdrawal include withdrawal of consent by the participant, or loss to follow-up as defined above. Reasons for withdrawal, if known, will be collected and tabulated.

5.4.4 PRESENTATION OF WITHDRAWAL/LOSS TO FOLLOW-UP

Numbers and percentages of withdrawal/loss to follow-up will be reported and tabulated for the entire study and by site.

5.5 BASELINE PARTICIPANT CHARACTERISTICS

5.5.1 LIST OF BASELINE DATA

The full list of variables in database can be found in the data dictionaries (S:\Centre for Trials Research\Research\Mixed Studies\DOMINO\eTMF\8.0 Data Management & Statistics\8.1 Data Management\8.1.3 Metadata\Data Dictionaries)

Physical activity data: Data from the Fitbit Charge 4 activity tracker will be collated for all participants across all clinical sites to provide a summary overview of physical activity across the cohort. Step counts from FitBit and Metabolic Equivalent of Task (MET)-minutes from the International Physical Activity Questionnaire (IPAQ) will be used as continuous variables, when controlling for bias, and will be dichotomised, when assessing their effect on the

progression of the disease. Scores from the Brunel Physical Activity Questionnaire and Lifetime Total Physical Activity Questionnaire (LTPAQ) will be used to report summary statistics based on these scores and be compared (e.g. with scatter plots and correlation coefficients) to FitBit step counts, and whether it is feasible to use these scores interchangeably. Percentage of missing values will be reported. The LTPAQ is an optional questionnaire, so reports of missing values should include which have missing components, and which are entirely missing (e.g. due to optional non-completion from the participant), if that information is available.

Diet data: Information from a modified Food Frequency Questionnaire (FFQ) will be collated to produce a summary overview of nutrition. Descriptive summary statistics from Malnutrition Universal Screening Tool (MUST) data will be reported, to understand the percentage of people who are malnourished at baseline. Additionally, summary descriptive statistics will be reported for data from a Swallowing Screening Tool (EAT10) about the number of people who have swallowing difficulties at baseline. Finally, summary statistics for BMI, calf and waist circumference will be reported at baseline. Percentage of missing values will be reported.

Sleep data: Sleep data obtained from activity trackers worn will be collated to generate a summary overview. These variables will be used as continuous covariates when controlling for bias, and will be dichotomised when assessing their effect on the progression of the disease. Data from the HD Sleep Questionnaire will be summarised and reported. These will be compared to data from activity trackers (e.g. using Pearson correlation or scatter plots, or other methods if appropriate), to assess consistency between self-reported and tracker data. Percentage of missing values will be reported. Linked Enroll-HD data from all sites will be summarised (e.g., demographic data, listed in Table 1). Percentage of missing values will be reported.Data from HD-ProTriad (severity of the disease) will be summarised and descriptive statistics will be reported. Percentage of missing values will be reported.

Scoring methods for the questionnaires mentioned above are presented in Appendix D.

Table 1: List of variables available from Enroll-HD

Variable Name	**Description**
Age	Age at baseline (at the first visit)
HD-Stage	- Late pre-manifest (LPM): No motor onset, age ≥ 18yrs, CAG ≥ 36, CAP > 80

	- Stage 1: 10 < TFC ≤ 13 - Stage 2: 6 < TFC ≤ 10 - Stage 3: TFC ≤ 6
Sex	Sex assigned at birth (male/female)
CAG	CAG repeat length
CAP	CAG-age product, defined as Age*(CAG-30)/6.49
race	Ethnicity
hddiagn	Age at clinical HD diagnosis
sxraterm	Rater's judgement of initial major symptom
ccdep	Has depression (includes treatment with antidepressants with or without a formal diagnosis of depression) ever been a part of the participant's medical history?
hdcat_I	Participant category (latest)
cmtrt__atc	Pharmacotherapy — ATC code
cmdostot	Pharmacotherapy – Total daily dose
cmenrf	Pharmacotherapy – Ongoing
cmtrt	Non-Pharmacologic Therapies – Therapy
cmfrq	Non-Pharmacologic Therapies – Number of times
cmdosfrq	Non-Pharmacologic Therapies – Frequency
cmenrf	Non-Pharmacologic Therapies – Ongoing
cmcar	Nutritional Supplements – Type
cmtrt__atc	Nutritional Supplements – ATC code(s)
cmenrf	Nutritional Supplements – Ongoing
mhbodysy	Comorbid Conditions – Body system code

mhterm_modify	Comorbid Conditions – Modified term (coded with ICD10)
mhenrf	Comorbid Conditions – Ongoing
height	Height (cm)
weight	Weight (kg)
bmi	BMI
alcab	Does the participant currently drink alcohol?
alcunits	Units of alcohol per week
tobab	Does the participant currently smoke?
packy	Pack-years
cafab	Current caffeine use?
cafpd	Do you drink more than 3 cups of coffee, tea and cola drinks combined per day?
mar	Drug use for non-medical reasons? – Abuse
marfrq	Drug use for non-medical reasons? – Frequency
isced	ISCED education level
jobclas	Employment (status)
motscore	Motor Score (TMS)
ocularh	Ocular pursuit – Horizontal
ocularv	Ocular pursuit – Vertical
sacinith	Saccade initiation – Horizontal
sacinitv	Saccade initiation – Vertical
sacvelh	Saccade velocity – Horizontal
sacvelv	Saccade velocity – Vertical
dysarth	Saccade velocity – Dysarthria

tongue	Saccade velocity – Tongue protrusion
fingtapr	Finger taps – Right
fingtapl	Finger taps – Left
prosupr	Pronate supinate (hands) – Right
prosupl	Pronate supinate (hands) – Left
luria	Pronate supinate (hands) – Luria
rigarmr	Rigidity (arms) – Right
rigarml	Rigidity (arms) – Left
brady	Rigidity (arms) – Bradykinesia body
dysttrnk	Maximal dystonia – Trunk
dystrue	Maximal dystonia – RUE (right upper extremity)
dystlue	Maximal dystonia – LUE (left upper extremity)
dystrle	Maximal dystonia – RLE (right lower extremity)
dystlle	Maximal dystonia – LLE (left lower extremity)
chorface	Maximal chorea – Face
chorbol	Maximal chorea – BOL (bucco-oral-lingual)
chortrnk	Maximal chorea – Trunk
chorrue	Maximal chorea – RUE (right upper extremity)
chorlue	Maximal chorea – LUE (left upper extremity)
chorrle	Maximal chorea – RLE (right lower extremity)
chorlle	Maximal chorea – LLE (left lower extremity)
gait	Gait
tandem	Tandem walking

retropls	Retropulsion pull test
diagconf	Diagnostic confidence level (DCL)
tfcscore	UHDRS TFC – Functional score
fascore	UHDRS function – functional assessment score
sdmt1	Core Cognitive Assessment – Total correct
swrt1	Core Cognitive Assessment – Total correct
anxscore	HADS-SIS – Anxiety subscore
hads_depscore	HADS-SIS – Depression subscore
irrscore	HADS-SIS – Irritability subscore
outscore	HADS-SIS – Outward irritability subscore
inwscore	HADS-SIS – Inward irritability subscore
tug	Physiotherapy – Timed “up and go” performed
tug1	Physiotherapy – Total time
scst	Physiotherapy – 30 sec chair stand test performed
scst1	Physiotherapy – Number of times the participant stands in 30 seconds
depscore	PBAs – Depression
irascore	PBAs – Irritability and aggression
psyscore	PBAs – Psychosis
aptscore	PBAs – Apathy
exfscore	PBAs – Executive function
pcs	SF-12 – Physical Component (PCS)
mcs	SF-12 – Mental Component (MCS)

5.5.2 DESCRIPTIVE STATISTICS

Categorical data will be presented using frequencies and percentages. Continuous data will be presented using number of patients, mean and SD or median and IQR to summarise the key demographics and patient characteristics. Baseline data will be tabulated for each site, and in total. Baseline data for those who were followed up at 12 months will also be tabulated and compared to those who were not followed up at 12 months. Summary statistics and percentage of missing values will be tabulated and reported for all covariates listed in Table 1 (baseline measures).

6. ANALYSIS

6.1 OUTCOME DEFINITIONS

6.1.1 PRIMARY OUTCOME(S)

The primary objective of this study is to establish the feasibility of linking lifestyle, clinical and genetic data and to explore their interplay with HD symptoms and severity.

We will explore the proportions of recruited participants with linked DOMINO-HD and Enroll-HD clinical data as well as linked GWAS data. We will also explore follow-up rates and lifestyle and wearable activity tracker data completeness and balance statistics in important variables. Our data will be used to provide progression criteria recommendations for future studies in this field.

Aspects of feasibility which will be assessed are:

- Completeness of key lifestyle questionnaires (at baseline: HD Pro-Triad, FFQ, LTPA, Sleep, IPAQ, Relationship) data [>80% returned with >80% completeness = feasible, >60% returned with >60% completeness = changes required, <60% returned or <60% completeness = not feasible]
- Completeness of activity tracker data [>70% feasible, 50-70% changes required, <50% not feasible]
- % participants linked with Enroll-HD clinical data [>80% feasible, 60-80% changes required, <60% not feasible]
- % participants linked with GWAS data [>60% feasible, 40-60% changes required, <40% not feasible]
- % participants with follow-up at 12m [>70% feasible, 50-70% changes required, <50% not feasible]
- Feasibility of balancing important covariates. Adequate balance of baseline covariates is achieved when the maximum Kolmogorov-Smirnov (KS) value is equal to or below 0.1. However, as the recommendation is for 60-80 observations per covariate per treatment

group (Markoulidakis, Holmans, et al., 2021), the feasibility of achieving balance should be determined alongside the available sample size.

- Agreement between questionnaire and activity tracker data

6.1.2 TIMING, UNITS AND DERIVATION OF PRIMARY

The primary outcomes will be assessed at 12-months follow-up.

6.1.3 LIST OF SECONDARY OUTCOMES

We will use a clinical composite score for the initial secondary analyses, namely the composite Unified Huntington's Disease Rating Scale (cUHDRS) defined by (Schobel et al., 2017) as follows:

$$cUHDRS = \frac{TFC - 10.4}{1.9} - \frac{TMS - 29.7}{14.9} + \frac{SDMT - 28.4}{11.3} + \frac{SWRT - 66.1}{20.1} + 10$$

where:

- Total Functional Capacity (TFC) is a standard assessment of overall function in HD and ranges from 13 (normal function) to 0 (complete loss of function).
- The Total Motor Score (TMS) has a maximum score of 124. Higher values indicate more severe impairment.
- SDMT is the Symbol Digit Modality Test. Higher scores reflect better cognitive functioning.
- SWRT is the Stroop Word Reading Test. Higher scores reflect better cognitive functioning.

The higher the value of cUHDRS, the better the condition of the patient (lower severity). As the disease progresses and the condition of the individual becomes more severe, the cUHDRS drops.

Individual components of cUDHRS (namely TFC, TMS, SDMT and SWRT) will also explored. This means that the entire analysis described below for the cUHDRS as an outcome will be reproduced for each of the four cUHDRS components (one at a time).

6.1.4 ORDER OF TESTING

Not applicable.

6.1.5 TIMING, UNITS AND DERIVATION OF SECONDARIES

The secondary outcomes (cUHDRS, TMS, TFC, SWRT and SDMT) will be assessed 12 months after the initial (baseline) assessment.

6.2 ANALYSIS METHODS

The main modifiable lifestyle factors we are interested in, in terms of their impact on cUHDRS progression, are sleep, diet, and physical activity.

1. Sleep:

Total minutes of sleep are available per participant through the extracts from activity trackers. As there is no evidence in the literature about the minimal time of sleep required for individuals with HD, we will utilise the first tertile to dichotomise the participants into two categories (those with more sleep than two thirds of the sample vs. the rest). Dichotomisation based on the first quartile and the median, respectively, will be used as an alternative, to assess the sensitivity of our findings. Sleep efficiency – defined as the ratio of sleep duration (average total sleep time, TST) to duration in bed (average total time in bed, TBT) – will also be considered. TBT and TST will be extracted from FitBit. Alternatively, propensity score (PS) weighting for continuous treatment variables, as described in section 6.2.2, could be explored. The first tertile will be used as a cut-off point for the primary analysis, dichotomisation on the first quartile and the median, respectively, will also be explored, to understand the sensitivity of our findings.

The HD sleep questionnaire contains 45 questions that focus on different sleep-related issues such as duration, quality of sleep, abnormal nocturnal behaviour and quality of life. Variables emerging from the HD sleep questionnaire will be reported descriptively. A sleep disturbance score will be derived from a sub-set of the questions, namely 4-11, 14, 15, 18, 19, 14, 27, 28, 34 (Goodman et al., 2010).

Self-reported total sleep time (calculated from questions 1 and 2 of the HD sleep questionnaire) will be compared with average total sleep time from activity trackers. The values for sleep efficiency from activity trackers and sleep questionnaires (sleep disturbance) will also be assessed for correlation (association of high sleep disturbance with low sleep efficiency) and reported.

2. Diet:

We will derive a Mediterranean Diet Index (MDI) as defined by (Trichopoulou et al., 2003). This index's values range from 0 to 9, with higher values indicating higher adherence to an MD. The final score is the sum of nine elements, each of which is either 0 or 1. Beneficial components (namely 1. vegetables, 2. legumes, 3. fruits and nuts, 4. cereal, and 5. fish) contribute one point if their consumption is above the sex-specific sample median. For non-beneficial components, (6. dairy, 7. meat and poultry) one point is assigned to those whose consumption is below the sex-specific sample median. Finally, for alcohol consumption (8.), one point is assigned to men who consume between 10 and 50 g per day and to women who consume between 5 and 25 g a day. Finally, to score lipid intake, the unsaturated to saturated fatty acids (9.) ratio is calculated and one point assigned if the ratio exceeds the sex-specific sample median. A full description of the definition of the MDI is in Appendix A.

Additionally, the Healthy Eating Index (HEI) will be used as a diet score. Its values range from 0 to 100, with higher values representing better eating habits. The final score is a sum of 13 components, nine categories of adequate eating, and four moderators. A full description of the definition of the HEI is in Appendix B.

Both MDI and HEI will be used (one at a time) to assess the sensitivity of our findings. Additionally, the two scores will be compared as to their effect on the progression of HD – unless one is dropped (see section 6.2.3).

3. Physical activity:

Step counts over a 6-week period from the baseline assessment will be extracted from each participant's FitBit. 5000 steps per day (Tudor-Locke et al., 2013) will be used as a threshold to dichotomise participants into sedentary (< 5000 steps per day) and active (≥ 5000 steps per day).

A MET-minute is the amount of energy expended during a minute while at rest. It is a ratio of a person's working metabolic rate relative to their resting metabolic rate (one MET is the energy spent sitting at rest for one minute). As there is no evidence in the literature about a clear distinction between active and inactive lifestyle for individuals with HD, participants in the lowest tertile (≤ 33%) will be classified as inactive (Schiepers et al., 2018). MET-minutes will be extracted from FitBit. MET-minutes will be used as an alternative measure of physical activity (if available) to assess the sensitivity of our findings.

Construction of LifeHD score:

A LIBRA style index (Anstey et al., 2013; Kivipelto et al., 2006; Schiepers et al., 2018) will be created to investigate the association of a lifestyle-based score with the progression of HD

(as measured by cUHDRS and its four components). This will be called LifeHD. Modifiable lifestyle factors will be included, namely:

- Sleep;
- Diet;
- Physical activity;
- BMI.

Additionally, the following non-modifiable measures will be included:

- Age group (< 47yrs, 47 − 53yrs, > 53yrs);
- CAG repeat length;
- HD stage.

To achieve the scoring sums (LifeHD scores), the steps in Appendix C will be followed.

6.2.1 LIST OF METHODS AND PRESENTATION

A table with each feasibility criterion will be reported, alongside the percentage or nominal value for each of the sub-categories.
Means, medians, standard deviations, (empirical) 95% confidence intervals and ranges will be reported for continuous covariates, while for their respective binary transformations, we will report frequencies and percentages.
The impact of each lifestyle factor of interest (namely sleep, diet, physical activity) will be estimated as a causal treatment effect (e.g. ATE or ATT) (Markoulidakis, Taiyari, et al., 2021) using the doubly robust model. PS and balancing weights (assuming the availability of a large enough sample size, and that adequate balance is achieved – this is defined as the maximum KS statistic of all balancing covariates being below or equal to 0.1) will be used to control for confounding bias. The covariates which will be used to control for confounding bias are the lifestyle factors not under study in each case (e.g., when modelling the effect of sleep on the progression of cUHDRS, then diet and physical activity will be considered as confounders), as well as the baseline value of cUHDRS. If sample size allows, additional covariates will be added to control for confounding bias, including age and HD stage. If adequate balance is not achieved at baseline (defined as maximum KS statistic below or equal to 0.1), then mean differences of the outcome will be reported, and/or estimates using a regression model adjusting for the other two lifestyle factors, age, and HD stage. In the latter case, no balancing weights will be used. The full analysis will be re-attempted using the four components of cUHDRS as outcome variables, and similar results as for the original analysis will be reported.

6.2.2 COVARIATE ADJUSTMENT

Unweighted analyses (using regression models) will control for age, HD stage, as well as baseline values for the outcomes. Additionally, we will control for the two lifestyle (the continuous version of these) factors not included in each case.

The primary goal is to control for confounding bias, using PS and balancing weights (Markoulidakis, Taiyari, et al., 2021), for the two lifestyle factors other than the one under study in each case. Expecting limitations due to the restricted sample size (Markoulidakis, Holmans, et al., 2021), we will attempt to balance additionally for age, HD stage, and baseline values of the outcome.

The outcome analysis will also be performed using the LifeHD score constructed, as a categorical covariate, with and without the use of PS and balancing weights. The LifeHD score will be used both in addition to the individual lifestyle factors (e.g. using LifeHD score as 'treatment', and the three lifestyle factors as confounders), and solely – without taking into consideration the other lifestyle factors.

Outcome will be standardised, so that we report standardised effect sizes. The estimand of interest (e.g. ATE or ATT) will be determined based on whether adequate balance is achieved or not. If we are unable to achieve adequate balance, then unweighted regression and means difference will be reported.

Additionally, significance tests for difference of means of the covariates we wish to control for and the outcome will be reported. The pseudo-groups will be created using the binary indicators of diet, sleep, and physical activity.

Genetic information (via polygenic risk scores) will only be used in the analysis if it is feasible based on the available sample.

6.2.3 ASSUMPTION CHECKING

Standard model checking will be performed including fitted versus residual plots. Overlap will be checked using minimum and maximum values for each covariate as well as overlap density plots (Markoulidakis, Taiyari, et al., 2021) using the CoBWeb app. Sensitivity analysis plots (using OVtool) (Markoulidakis, Taiyari, et al., 2021; Pane et al., 2021) will be used to assess the assumption of no unobserved confounders left out of the analysis. This could be used even on simple regression analysis, without the use of PS and balancing weights. Alternatively, unmeasured confounding may be assessed using E-values (VanderWeele TJ & Ding P, 2017; Haneuse, et al., 2019).

6.2.4 ALTERNATIVE METHODS IF DISTRIBUTIONAL ASSUMPTIONS NOT MET

If the pseudo-treatment groups, created by the treatment indicators of sleep, diet, and physical activity, are not properly overlapped (Markoulidakis, Taiyari, et al., 2021) at baseline, trimming the sample will be considered. This would be case-specific for each treatment indicator separately. Feasibility of trimming would rely on the missingness rate and the total number of potential outliers compared to the total sample available (e.g. 5% of the total sample could be used as a decision rule as to whether to remove outliers or not).
The HEI and MDI scores will both be used as measures of diet, and binary indicators will be created to split the participants into treatment groups (healthy/unhealthy diet and high/low adherence to MDI, respectively). There is no preference as to which measure is favourable at the time of writing, but one of these could be dropped if either it creates treatment groups with very poor overlap of the baseline covariates (leading to a high rate of outliers), or in case the sizes of the two groups created by the dichotomisation are substantially different e.g. resulting in a 3:1 ratio or more extreme.

6.2.5 SENSITIVITY ANALYSES

OVtool will be used to assess the sensitivity of treatment effect estimation (Markoulidakis, Taiyari, et al., 2021; Pane et al., 2021) to unobserved confounders (for weighted analysis), and to other covariates that could affect the outcome (for unweighted analysis). Alternatively, E-values may be used to assess unmeasured confounding (VanderWeele TJ & Ding P, 2017; Haneuse, et al., 2019).

6.2.6 SUBGROUP ANALYSES

Numbers will not be high enough to conduct subgroup analyses.

6.3 MISSING DATA

Where missing data is likely to occur, it will most likely be due to participant drop-out or loss to follow-up. The amount of missing primary outcome data will be tabulated by pseudo-group and site.
Multiple imputation will be attempted, but the feasibility of this would rely on the sample size and percentage of missingness (>5% of the data points). If the missing rate is below 5%, but the missingness on a specific covariate exceeds 20%, then this covariate could be dropped or replaced, providing an alternative measure is available. If the overall missing rate is below 5%, complete case analysis will be performed.

6.4 ADDITIONAL ANALYSES

DOMINO-HD contains a trial within a study (TWAS), which is investigating methods to improve compliance with use of (and upload of data from) the FitBit trackers using SMS as prompts. This TWAS focuses on the comparison of different strategies of SMS notification sent to the participants, and explores whether any specific strategy resulted in better compliance rates (resulting in more data). The two strategies are characterised by routine and data-driven prompt notifications. Both groups receive an SMS per week, with the message in the first case always the same, while in the latter case it will be more personalised, informed by the success or failure to upload at least 50% of the previous week's data. In both cases, if the data are not updated properly for a certain amount of time, participants receive a phone call after four weeks (routine group), or an enhanced SMS after two weeks (data-driven group).
This is a randomised trial, and the outcome measure will be the total number of wear hours from week two to week nine of participating in the study (the total outcome time frame is eight weeks). Secondary outcomes are the total wear time across the first 3, 6, 9 and 12 months of the study. Statistical testing for difference of means between the two groups will be performed for all outcomes, to understand whether the two approaches lead to different levels of compliance. Two-sided t-tests will be performed to understand if one approach is inferior to the other. Finally, a simple regression model on the average wear hours per week will be fitted to evaluate the effect (if any) of the notification method (including covariates such as age, sex, HD stage, and a dummy variable indicating which group of notifications the participant is in). Descriptive statistics of wear time (in hours) will be reported by group, to examine any patterns of wear over time (decline in total adherence, etc.).
Summary statistics will be reported in a tabulated form, for all outcomes and groups (routine and data-driven), as well as summary statistics and 95% CIs for the mean differences.

6.5 HARMS

Safety reporting is covered in the main protocol. Adverse events will be tabulated.

6.6 STATISTICAL SOFTWARE

R version 4.3.0 or higher; R packages: twang, CBPS, entbal will be used to compute propensity scores and balancing weights; R package twang will be used to compute balance statistics (standardised mean difference and KS statistic); R package OVtool will be used to produce sensitivity analysis graphs; R package EValue will be used to calculate E-values; R Shiny app CoBWeb will be used for complete-case analysis (available at: https://andreasmarkoulidakis.shinyapps.io/cobweb/); R package mice will be used for the imputation of missing values; R package MatchThem could be used for missing data analysis; Python version 3.10.2 will be used for the initial extraction of FitBit data.

7. REFERENCES

7.1 NON-STANDARD STATISTICAL METHODS

7.2 DATA MANAGEMENT PLAN

S:/Centre for Trials Research/Research/Mixed Studies/DOMINO/eTMF/8.0 Data Management & Statistics\8.1 Data Management\8.1.1 Data Management Plan

7.3 STUDY MASTER FILE AND STATISTICAL MASTER FILE

S:/Centre for Trials Research/Research/Mixed Studies/DOMINO/eTMF

S:/Centre for Trials Research/Research/Mixed Studies/DOMINO/eTMF/8.0 Data Management & Statistics/8.5 Statistics

7.4 OTHER SOPS OR GUIDANCE DOCUMENTS

SAP DEVIATION LOG

Document number:		**Document version:**	
Reason for deviation:			

8. APPENDICES

Appendix A: Developing a Mediterranean Diet Index (MDI)

The MDI is a score (Trichopoulou A, Costacou D, 2003) taking integer values ranging from 0 to 9. It consists of nine components, each of which contributes either 0 (if the condition is not met) or 1 (if the condition is met). Table 2 reports the nine categories and when 1 point for each category is assigned. The MDI is the sum of the nine components. In (Trichopoulou A, Costacou D, 2003) adherence to a Mediterranean diet was classified as low for individuals with scores 0-3, moderate for individuals with scores 4-5, and high for those with score 6-9. Based on this classification, the lifestyle factor "diet" could be either dichotomised as high adherence to MD (MDI score 6-9) or non-high (MDI score 0-5), or used as a categorical covariate (with three levels), and PS and balancing weights computed for a treatment variable with three factors.

Table 2: The nine categories of foods included in the diet indices. *The ratio of polyunsaturated to saturated fatty acids is computed based on the standard reference values of polyunsaturated and saturated fatty acids included in the foods the individuals report to consume per week. **This corresponds to 2(1) glass(es) of wine or beer for men(women).

Category	Food category	Condition to obtain 1 point (for MDI)
1.	**Vegetables**	1 consuming ≥ sex-specific sample median (per week) 0 consuming < sex-specific sample median (per week)
2.	**Legumes**	1 consuming ≥ sex-specific sample median (per week) 0 consuming < sex-specific sample median (per week)
3.	**Fruits and nuts**	1 consuming ≥ sex-specific sample median (per week) 0 consuming < sex-specific sample median (per week)
4.	**Cereal**	1 consuming ≥ sex-specific sample median (per week) 0 consuming < sex-specific sample median (per week)
5.	**Fish**	1 consuming ≥ sex-specific sample median (per week) 0 consuming < sex-specific sample median (per week)
6.	**Diary**	0 consuming ≥ sex-specific sample median (per week)

		1 consuming < sex-specific sample median (per week)
7.	**Meat and poultry**	0 consuming ≥ sex-specific sample median (per week) 1 consuming < sex-specific sample median (per week)
8.	**Alcohol consumption**	0 otherwise 1 for consumption of 10-50g per day for men, and 5-25g per day for women**
9.	**Unsaturated to saturated fatty acids ratio***	0 otherwise 1 for ratio higher than the sex-specific sample median

The computation of each individual score for categories 1-8 (whether one or zero is assigned) uses the average number of servings per week consumed for each participant. For the computation of each individual's unsaturated to saturated fatty acids ratio, the total sum of unsaturated and the total sum of saturated fatty acids per week is used.

Appendix B: Developing a Healthy Eating Index (HEI)

The HEI is a score (Krebs-Smith et al., 2018) ranging from 0 to 100. It consists of 13 components, each of which contributes any value from 0 to 5 or 10 (depending on the component), with the maximum value (5 or 10) representing the best eating behaviour for each component. Table 3 reports the 13 categories, and the condition for the minimum/maximum points to be assigned for each category. The HEI is the sum of the nine components. To dichotomise the HEI, we will use 59 as a cut-off – higher values will be assigned 1, and 0 otherwise. Further, 79 could be used as a second cut-off point to trichotomise (if desired). If no or very few observations score over 59, then the median of the sample will be used as threshold for dichotomisation.

Table 3: The 13 categories of foods included in the HEI. *Ratio of poly- and monounsaturated fatty acids (PUFAs and MUFAs) to saturated fatty acids (SFAs).

Component	Maximum point	Standard for maximum score	Standard for minimum score of zero
Adequacy:			
Total fruits	5	≥0.8 cup equiv. per 1000kcal	No fruits
Whole fruits	5	≥0.4 cup equiv. per 1000kcal	No whole fruits
Total vegetables	5	≥1.1 cup equiv. per 1000kcal	No vegetables
Greens and beans	5	≥0.2 cup equiv. per 1000kcal	No dark green vegetables or legumes
Whole grains	10	≥1.5 oz equiv. per 1000kcal	No whole grains
Dairy	10	≥1.3 cup equiv. per 1000kcal	No dairy
Total protein foods	5	≥2.5 oz equiv. per 1000kcal	No protein foods
Seafood and plant proteins	5	≥0.8 oz equiv. per 1000kcal	No seafood or plant proteins
Fatty acids*	10	(PUFAs+MUFAs)/SFAs ≥2.5	(PUFAs+MUFAs)/SFAs ≤1.2
Moderation:			
Refined grains	10	≤1.8 oz equiv. per 1000kcal	≥4.3 oz equiv. per 1000kcal
Sodium	10	≤1.1 grams per 1000kcal	≥2 grams per 1000kcal

Added sugars	10	≤6.5% of energy	≥26% of energy
Saturated fats	10	≤8% of energy	≥16% of energy

Appendix C: LifeHD score

Table 4 reports the covariates which will be included in the LifeHD index. The following algorithm describes the development of the score.

Step 1. The risk factors will first be included separately in a regression model (together with non-lifestyle factors: age, sex, HD stage, time since diagnosis, and follow-up time since last visit). We will use a linear or logistic regression model, depending on the outcome (continuous or binary). This will also enable us to understand the impact of each category on the outcome variable and define the reference category for each factor. For diet, the sign and magnitude of each category will allow us to understand the most and least beneficial clusters – thus the beneficial category is group 1, and the other category is 0. The least beneficial cluster will be the reference category on the LifeHD index (group 0), such that any movement from this category to another (group 1 – or higher for more than two categories) will increase the index – thus, higher values of the index indicate "better" lifestyle.

Step 2. Factors that were significant in the first step will be simultaneously placed into a single regression model – linear or logistic regression depending on the outcome. (For DOMINO-HD, as the outcome is cUHDRS – or its components – linear regression will be used.)

Step 3. From the final regression models of step 2, risk scores will be assigned for each factor with the respective β coefficients (steps 4 and 5).

Step 4. All β coefficients will be standardised so that the lowest one has a value of 1 — this is achieved by multiplying every β coefficient with 1/minimum(β).

Step 4b. (optional) Rounding each coefficient to the closest integer. In this case, all individual scores are positive integers. This will enable easier calculaiton of the scores, and higher values of the score would indicate better "lifestyle".

Step 5. The risk score for an individual will be obtained by summing the scores for the appropriate level of each of the risk factors.

Correlation and association of LifeHD with the outcome and the lifestyle factors under study (diet, sleep, physical activity) will be tabulated.

Table 4: Lifestyle factors included in LifeHD. *Modifiable; **Non-modifiable

Variable	**Calibration**
Diet*	MDI: 0-5; 6-9 HEI: 0-59; 60-100
Sleep*	SE > tertile/quantile SE ≤ tertile/quantile
Physical activity*	MET > tertile MET ≤ tertile
BMI*	BMI ≤ 25 normal 25 < BMI ≤ 30 overweight BMI > 30 obese
Age**	< 47 years old 47 – 53 years old ≥ 53 years old
CAG repeat length**	35 < CAG < 42 CAG ≥ 42

HD stage**	Participants could be either in stage: - Late pre-manifest (LPM): No motor onset, age ≥ 18yrs, CAG ≥ 36, CAP > 80 - Stage 1: 10 < TFC ≤ 13 - Stage 2: 6 < TFC ≤ 10 - Stage 3: TFC ≤ 6
Antidepressant medication*	Use; or No use

Table 5: Definition of variables/outcomes

Outcome/definition	Scale(s)	Timing of measurements	Score range	Interpretation
	Age	At baseline		The age of the participant at the time of the baseline assessment
	HD stage	At baseline	LPM, stage 1, 2, 3	The HD stage of the participant at the beginning of the study
	CAG	At baseline		The length of CAG repeat of the participant
	CAP	At baseline		The CAP score of the participant at the beginning of the study
	Sex	At baseline		The sex of the participant as specified at birth
Severity of the disease	cUHDRS	At baseline and 12 months after baseline		Measure of the severity of the disease at the baseline assessment

				and at the time of the follow-up assessment. Higher values indicate less severe impairment.
Severity of the disease – motor	TMS	At baseline and 12 months after baseline	0-124	Lower values indicate less severe impairment
Severity of the disease – functional	TFC	At baseline and 12 months after baseline	0-13	Higher values indicate less severe impairment
Severity of the disease – cognitive	SDMT	At baseline and 12 months after baseline	Raw number of correct answers	Higher values indicate less severe impairment
Severity of the disease – cognitive	SWRT	At baseline and 12 months after baseline	Raw number of correct answers	Higher values indicate less severe impairment
Sleep quality	Total minutes of sleep	At baseline	Time (in minutes)	Average (over one week) minutes of sleep per day
Sleep quality	Sleep efficiency	At baseline		The ratio of total minutes of sleep to time in bed

Adherence to Mediterranean diet	MDI score	At baseline	0-9	Higher values indicate better adherence to Mediterranean diet
Adherence to a healthy diet	HEI score	At baseline	0-100	Higher values indicate healthier diet
Physical activity	Step count	At baseline	Raw number of steps per day	Average (over one week) number of daily steps
Physical activity	MET-minutes	At baseline		A MET is the ratio of a person's working metabolic rate relative to their resting metabolic rate (one MET is the energy spent sitting at rest). Higher values indicate more active lifestyle.
	BMI	At baseline	Raw number	≤ 25 normal 25 < BMI ≤ 30 overweight > 30 obese
Lifestyle undex	LifeHD	At baseline		Higher values indicate a better overall lifestyle (better adherence to

				a healthy diet, sleep schedule, and physical activity)
Total FitBit wear time	Total wear hours	Throughout the study	Time (in minutes)	Adherence to wearing the FitBit. Higher values indicate better (more consistent) usage of the tracker.

Appendix D: Scoring of questionnaires

Questionnaire	Scope	Scoring
IPAQ	Assessing the types of intensity of physical activity and sitting time that people do as part of their daily lives, and using this time to estimate total physical activity in MET-minutes/week and time spent sitting.	To compute the MET-minutes using IPAQ the following formula is used: *8 × (minutes of vigorous activity per day) × (# of days)* + *4 × (minutes of moderate activity per day) × (# of days)* + *3.3 × (minutes of more than 1' walking per day) × (# of days)*
Brunel	The Brunel lifestyle physical activity questionnaire consists of 9 questions regarding the duration and severity of planned (6) and unplanned (3) physical activity.	Each question receives 1-5 points, depending on the frequency of the activity in question (e.g. 1 = Not at all, 5 = Highly). Planned physical activity is defined as, "… any activity that is scheduled into a daily routine, which may enhance your health, fitness, or well-being" (e.g. brisk walking,

		cycling). Unplanned physical activity is defined as any form of physical activity "excluding pre-planned physical activity" (e.g. heavy housework, playing with children). Factor scores for planned and unplanned dimensions of physical activity are calculated by adding scores from items 1–6 (planned) and 7–9 (unplanned), then dividing them by six and three, respectively. Factor scores range from 1 to 5, with higher scores indicating higher engagement in physical activity.
LTPAQ	Estimating the average physical activity of the participant during their lifetime.	<u>Categories of intensity</u> Category 1: Requires only sitting with minimal walking. Category 2: Requires a minimal amount of physical effort such as standing and slow walking. There is no increase in heart rate and there is no perspiration.

		Category 3: Requires carrying light loads and continuous walking. These activities would increase the heart rate slightly and may cause some light perspiration. Category 4: Requires carrying heavy loads, brisk walking, and climbing. These jobs/activities would increase the heart rate substantially and cause heavy sweating. The outcome variables are estimated as the number of hours spent in each type of activity for different time periods in a respondent's lifetime and at different intensity levels. Total physical activity is estimated as the sum of occupational, household, and exercise/sports activities in hours per week. It is also possible to convert these data into energy expended by multiplying the hours spent by the estimated metabolic cost of that activity. The resulting data would be denoted as MET-hours/week. The MET corresponding to each activity are listed in the LTPQA Comprehensive Users' Guide.
Modified FFQ	The FFQ requires respondents to report the frequency of consumption of a predefined list of	A food file compiled from the reported food consumption and matched to the associated nutrition information that has been

	food items over a period of time. Some FFQs may include additional questions on portion sizes and preparation methods. To improve accuracy for reporting portion sizes, food atlas photographs, food replicas and food models may be utilised in non-weighed assessment methods.	established to use alongside the reported consumption. This information is then used to compute the HEI or MDI.
MUST	MUST is a five-step screening tool to identify adults who are malnourished, at risk of malnutrition (undernutrition), or obese. It also includes management guidelines which can be used to develop a care plan.	The MUST is scored in the database as 0, 1 or 2.
EAT10	The Eating Assessment Tool (EAT-10) is a screening for self-perceived oropharyngeal dysphagia (OD) in community-dwelling elders. A summated EAT-10 total score ranges from 0 to 40, with a score ≥ 3 indicative of OD.	The EAT-10 total score out of 40 is calculated in the database.

Sleep questionnaire	Identify sleep disturbances in patients with HD	For each question, either 0, 1 or 2 points are allocated for each reply (depending on the question). The total score is the sum of the points aggregated.
HD-PRO-Triad	The HD-PRO-Triad is an HD-specific, patient-reported outcome (PRO) instrument of the HD symptom triad (cognitive decline, emotional/behavioural dyscontrol, and motor dysfunction).	Each item in HD-PRO-Triad is scored on a scale of 1 to 5, with higher scores indicating worse functioning or health-related quality of life on each domain. The total score for each domain (cognition, emotional and behavioural dyscontrol, motor function) is computed as a mean based on the sum of scores of item responses divided by the number of items answered. The possible maximum total score for each domain is therefore 5 if the patient answered 5 to all items. The HD-PRO-Triad total score is computed as the sum of the three domain total scores, with a possible maximum of 15.
Apathy Evaluation Scale (Clinician)		The questionnaire consists of 18 items, scored from 1 to 4 each. The total score is the sum of the sub-scores, with higher score representing greater apathy.

DOMINO-HD:	STATISTICAL ANALYSIS PLAN

References

Anstey, K. J., Cherbuin, N., & Herath, P. M. (2013). Development of a new method for assessing global risk of Alzheimer's disease for use in population health approaches to prevention. *Prevention Science*, *14*(4), 411–421. https://doi.org/10.1007/s11121-012-0313-2

Goodman, A. O. G., Morton, A. J., & Barker, R. A. (2010). Identifying sleep disturbances in Huntington's disease using a simple disease-focused questionnaire. *PLoS Currents*, *2*. RRN1189. https://doi.org/10.1371/currents.RRN1189

Haneuse S *et al*., Using the E-Value to assess the potential effect of unmeasured confounding in observational studies. JAMA (2019);321(6):602–603. doi: 10.1001/jama.2018.21554. PMID: 30676631.

Kivipelto, M., Ngandu, T., Laatikainen, T., Winblad, B., Soininen, H., & Tuomilehto, J. (2006). Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *The Lancet Neurology*, *5*(9), 735–741. doi: 10.1016/S1474-4422(06)70537-3. PMID: 16914401.

Krebs-Smith, S. M., Pannucci, T. E., Subar, A. F., Kirkpatrick, S. I., Lerman, J. L., Tooze, J. A., Wilson, M. M., & Reedy, J. (2018). Update of the healthy eating index: HEI-2015. *Journal of the Academy of Nutrition and Dietetics*, *118*(9), 1591–1602. doi: 10.1016/j.jand.2018.05.021. Erratum in: J Acad Nutr Diet. 2019 Aug 20;: PMID: 30146071; PMCID: PMC6719291.

Markoulidakis, A., Holmans, P., Pallmann, P., Busse, M., & Griffin, B. A. (2021). *How balance and sample size impact bias in the estimation of causal treatment effects: A simulation study*. [Online]. Cornell University. Available at: https://arxiv.org/abs/2107.09009

Markoulidakis, A., Taiyari, K., Holmans, P., Pallmann, P., Busse, M., Godley, M. D., & Griffin, B. A. (2023). *A tutorial comparing different covariate balancing methods with an application evaluating the causal effects of substance use treatment programs for adolescents.* Health services & outcomes research methodology, 23(2), 115–148. https://doi.org/10.1007/s10742-022-00280-0

Schiepers, O. J. G., Köhler, S., Deckers, K., Irving, K., O'Donnell, C. A., van den Akker, M., Verhey, F. R. J., Vos, S. J. B., de Vugt, M. E., & van Boxtel, M. P. J. (2018). Lifestyle for Brain Health (LIBRA): a new model for dementia prevention. International Journal of Geriatric Psychiatry, 33(1), 167–175. doi: 10.1002/gps.4700. Epub 2017 Feb 28. PMID: 28247500.

Schobel, S. A., Palermo, G., Auinger, P., Long, J. D., Ma, S., Khwaja, O. S., Trundell, D.,

Cudkowicz, M., Hersch, S., Sampaio, C., & others. (2017). Motor, cognitive, and functional declines contribute to a single progressive factor in early HD. *Neurology*, *89*(24), 2495–2502. doi: 10.1212/WNL.0000000000004743. Epub 2017 Nov 15. PMID: 29142089; PMCID: PMC5729794.

Trichopoulou A, Costacou D, B. C. & T. D. (2003). Adherence to a Mediterranean diet and survival in a Greek population. *New England Journal of Medicine*, *348*(26), 2599–2608. doi: 10.1056/NEJMoa025039. PMID: 12826634.

Tudor-Locke, C., Craig, C. L., Thyfault, J. P., & Spence, J. C. (2013). A step-defined sedentary lifestyle index:< 5000 steps/day. *Applied Physiology, Nutrition, and Metabolism*, *38*(2), 100–114. doi: 10.1139/apnm-2012-0235. Epub 2012 Nov 8. PMID: 23438219.

VanderWeele TJ and Ding P, Sensitivity analysis in observational research: introducing the E-value, Annals of Internal Medicine (2017);167(4):268-274. doi: 10.7326/M16-2607. Epub 2017 Jul 11. PMID: 28693043.