Statistical Analysis Plan

A single-blind, parallel-group, randomised trial to evaluate the clinical and cost-effectiveness of a tailored diabetes education and behavioural weight management programme versus diabetes education, in adults with overweight or obesity and a new diagnosis of type 2 diabetes.

Trial registration number

ISRCTN18399564

SAP revision history

Date	Version	Justification for SAP version
20/01/2022	1	First draft sent to ALA and SJG for initial review.
19/04/2022	2	Incorporated comments from ALA, sent to SJS
25/04/2022	3	Incorporated comments from SJS
13/05/2022	4	Incorporated comments from PB; added sensitivity analysis re: COVID changes
30/05/2022	5	Incorporated comments from SJG; included definitions of “completion” provided by WW
29/06/2022	6	Added Missing Data and Cleaned Up Comments and tracked changes
03/07/2022	7	Incorporated comments received from TSC
20/09/2022	8	Added exclusion criteria for objectively measured physical activity (i.e. min. weartime)
27/09/2022	9	Incorporated additional comments discussed during the Programme

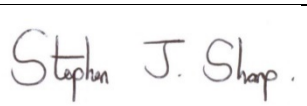
		Steering Committee meeting on 27/09/2022
05/10/2022		Incorporated final comments from Stephen

SAP responsibilities

Role in SAP development	Name, affiliation	Role in trial
SAP author	Julia Mueller, MRC Unit, University of Cambridge	Research Associate
SAP reviewer 1	Amy L. Ahern, MRC Unit, University of Cambridge	PI
SAP reviewer 2	Stephen J. Sharp, MRC Unit, University of Cambridge	Senior statistician
SAP reviewer 3	Simon Griffin, MRC Unit, University of Cambridge	PI

SAP signatures

Role	Name, affiliation	Date	Signature
Trial PI	Amy L. Ahern, MRC Unit, University of Cambridge	06/10/22	
Trial PI	Simon Griffin, MRC Unit, University of Cambridge	06/10/22	
SAP author	Julia Mueller, MRC Unit, University of Cambridge	06/10/22	

Overseeing statistician	Stephen J. Sharp, MRC Unit, University of Cambridge	06/10/22	

1 Introduction

1.1 Trial background and rationale

Treatment of diabetes and related complications (e.g. cardiovascular disease, amputation, kidney failure) uses 10% of the UK NHS budget.(1) This is predicted to rise to 17% in 2035 as the number of people with diabetes in the UK rises to 6.25 million, of which 5.6 million cases will be adults with type 2 diabetes (T2D).(1) Adults who are living with T2D are at increased risk of developing physical and mental health comorbidities, and have reduced quality of life and shorter life expectancy.(2,3) There are considerable social and economic costs to the individual living with diabetes as well as to wider society.(1,2,4)

While T2D is typically characterised as a progressive, irreversible condition, there is evidence that remission can be achieved by patients losing weight through bariatric surgery(5,6) or closely supervised very-low-calorie formula diets.(7),(8) However, many patients with T2D may be unsuitable for or unwilling to undergo these interventions, and given their high cost and reliance on specialists, they are unlikely to be widely adopted in the NHS in the near future. Partial or complete remission of type 2 diabetes has also been observed following smaller weight losses achieved through behavioural interventions.(9) Moreover, even without remission, behaviour change can produce improvements in health outcomes for people who have diabetes. We have shown that making healthy behaviour changes (e.g. increasing physical activity, reducing energy and fat intake) in the first year after diagnosis can reduce the likelihood of stroke or heart attack in the next 5 years.(10)

The Look AHEAD trial demonstrated that intensive specialist-led behavioural programmes could lead to weight loss and reductions in cardiovascular risk factors over 8 year follow up.(11) However, there are currently insufficient resources in the UK NHS to provide intensive, specialist-led behavioural programmes to the 3.2 million individuals who have T2D and the additional 200,000 who are diagnosed each year. Instead, current guidelines focus on structured diabetes education and dietary advice,(12) which is cheaper and scalable but has small, short term effects on weight and glycaemia, and relatively poor uptake.(13–15) A recent systematic review found that supportive behaviour-change programmes (with >11 hours of contact time) achieve greater reductions in weight and HbA_{1c} than structured education without additional support (<10hrs).(14) Integrating effective but scalable behaviour change programmes into care pathways for T2D could potentially improve glycaemic control and related risk factors and reduce complications. This would improve health and quality of life for people living with diabetes and reduce the burden of diabetes on health care resources.

We have previously shown that commercial open-group behavioural weight management programmes, such as Weight Watchers, are a scalable and cost-effective way to help people lose weight and reduce risk of diabetes.(16–18) A randomised trial in the US showed that a combination of Weight Watchers classes and remote dietary counselling achieved greater weight losses and reductions in HbA_{1c} than standard care over 1 year in people with diabetes.(19) Around a quarter of participants randomised to this programme achieved good glycaemic control (HbA_{1c} below 53mmol/mol) at 12 months, compared with 14% of those receiving standard care. In the UK, a similar intervention has been developed for use in the NHS that combines referral to Weight Watchers with NICE-compliant diabetes education and dietary advice. However, this programme is unlikely to be widely commissioned without robust evidence of cost-effectiveness.

The aim of this study is to evaluate the effectiveness of a tailored diabetes education and behavioural weight management programme (DEW) compared to diabetes education (DE), for people with a new diagnosis of T2D (≤3years). This analysis plan refers to the assessment of effectiveness; a separate plan will be developed for the cost-effectiveness and health economic modelling analyses.

1.2 Trial objectives/hypotheses

Primary objective

To evaluate the effect of DEW vs DE on glycated haemoglobin (HbA_{1c}) at 12 months in adults with a recent (within the last 3 years) diagnosis of type 2 diabetes mellitus.

Secondary objectives

To evaluate the effect of DEW vs DE on:

- 6 and 12 month changes in body weight, body fat percentage, blood pressure, lipid profile, and modelled cardiovascular risk[1]
- the probability of achieving clinically significant weight loss, good glycaemic control or diabetes remission at 6 and 12 months
- 6 and 12 month changes in diet and physical activity
- 6 and 12 month changes in psychosocial factors associated with successful weight control
- 6 and 12 month changes in health-related quality of life and wellbeing

[1] Whether we are able to examine modelled cardiovascular risk will depend on data availability.

2 Methods

2.1 Trial design

This is a pragmatic, randomised, single-blind, parallel group, two-arm, superiority trial. Participants are randomised to either DEW or DE (standard care) using block randomisation with a 1:1 allocation stratified by gender and duration of diabetes.

The **DEW programme** is the Live Well With Diabetes programme. It combines remote diabetes education and dietetic counselling with a supportive group-based behaviour change programme. It involves:

three telephone calls with a registered dietitian (one triage session, two diabetes education and dietetic counselling sessions)

free of charge membership of Weight Watchers for 6 months (involves weekly open-group meetings, held in-person or online when required as part of COVID-19 regulations)

digital tools and online materials

The **DE programme** is the Diabetes and Education Self-Monitoring for Ongoing and Newly Diagnosed (DESMOND) programme.(13,15) This is a structured diabetes education programme for people with a new diagnosis of type 2 diabetes (held in-person or online when required as part of COVID-19 regulations).

For details on DEW and DE, see the GLoW protocol(20).

2.2 Randomisation

Participants were allocated to one of the two intervention arms in a 1:1 allocation using individual-level blocked randomisation stratified by sex (male, female) and duration of diabetes (<1 year, 1-3 years) with a block size of 6. The randomisation sequence was computer-generated by the trial statistician and programmed by the data manager. The sequence is unknown to all other personnel, including study coordinators, outcome assessors and investigators.

2.3 Sample size

The primary outcome is 12 month change from baseline in HbA_{1c}, adjusted for baseline. Based on data from a previous trial in adults with a recent diagnosis of T2D,(21) we assumed a 16mmol/mol SD, a 0.8 correlation between baseline and follow-up and 25% attrition. In a US trial of a similar intervention in people with T2D of any duration, a difference of 4mmol/mol was observed between intervention (-3mmol/mol) and control (+1mml/mol) at 12 months.(19) We

need 576 participants to detect a difference between randomised groups of 3mmol/mol HbA_{1c} with 90% power at a 5% significance level.

2.4 Framework

Superiority hypothesis testing framework, assessing if DEW is superior to DE.

2.5 Interim analyses and stopping guidance

This is a low-risk trial with no rules for early stopping, and no planned interim analyses.

2.6 Timing of final analysis

All statistical analyses will be undertaken after the database is closed for 12 month follow up data.

2.7 Timing of outcome assessments

Outcomes are assessed at 0, 6, and 12 months.

3 Statistical principles

3.1 Confidence intervals and p-values

We will conduct two-sided statistical tests and 95% confidence intervals will be calculated around estimates of effect.

Multiplicity: As we have specified a single primary endpoint (12 month change from baseline in HbA_{1c}) to test the effectiveness of the intervention (confirmatory analysis), adjustments for multiple endpoints are unnecessary (22). All other outcome measures are secondary and therefore subsidiary and exploratory (22,23). P-values will only be reported for the main effects and interaction analyses of the primary outcome; 95% confidence intervals will be reported for all outcomes/comparisons.

3.2 Adherence and protocol deviations

Study and intervention adherence will be explored in a separate process evaluation. In the present analysis, we will report study adherence in terms of the withdrawal/loss to follow-up outcomes detailed in section 4.4. We will report intervention adherence in terms of the number and proportion of participants in the DEW group who completed ≥1 dietitian call and ≥2 dietician calls, attended at least one WW meeting and attended at least 75% (i.e. >=18) WW meetings. Intervention completers will be defined as those who completed both dietician calls and at least 75% of meetings.

3.3 Analysis populations

Individuals will be included in the analysis in the group to which they were randomised, regardless of their adherence to the intervention. Participants with missing outcome data at follow-up will be excluded from the analysis.

4 Trial population

See section 5 of protocol. Participants will be 576 adults with overweight or obesity who have a new diagnosis of type 2 diabetes in the past 3 years.

4.1 Screening data

Participants complete a screening questionnaire to confirm their eligibility (see sections 4.2 and 4.3).

4.2 Eligibility criteria

Inclusion criteria

- BMI ≥25kg/m^2
- Age ≥18 years
- Diagnosis of type 2 diabetes within the previous 36 months (confirmatory blood test will not be required to enter trial)
- Capable of giving informed consent
- Have a good understanding of the English language (study materials are not tailored to support non-English language speakers)
- Willing to be randomised
- Willing to attend follow up visits at a local participating GP practice

Exclusion criteria

- Using insulin
- Previous/planned bariatric surgery
- Current/planned pregnancy
- Current diagnosis of eating disorder
- Already received a structured diabetes education programme
- GP considers unsuitable
- Participation in another structured behaviour change programme for diet and/or physical activity within the past 3 months

4.3 Recruitment

See Trial Flowchart and section 4 in protocol.

4.4 Withdrawal/loss to follow-up

Withdrawal/loss to follow-up will be reported separately for each study group using:

- Number of participants assessed for eligibility

- Number and reasons for exclusions
- Number of participants randomised
- Number of participants allocated to each group
- Number of participants lost to 6-month and 12-month follow up
- Number of participants withdrawing from the trial and reasons for withdrawal
- Number of participants analysed in each group (with number and reasons for any exclusions)

4.5 Baseline characteristics

The following baseline characteristics of the study sample will be summarised separately within each randomised group and for the total sample:

- Demographics (place of residency, race/ethnicity, occupation, gender, religion, education, socioeconomic status (Index of Multiple Deprivation, IMD, occupation, income bracket) social capital (relationship status), age, disability, caring responsibilities, car ownership, access to the internet)
- Height (cm)
- Weight (kg)
- HbA_{1c} (mmol/mol and %)
- Body-mass-index (BMI) and BMI group (25-<30, 30+) (kg/m^2)
- Body fat %
- Systolic and diastolic blood pressure (mm Hg)
- Total cholesterol, HDL cholesterol, and LDL cholesterol (mmol/L)
- Volume of total physical activity as measured using accelerometer (Mean Euclidian Norm Minus One [ENMO] in mg, see section 5.1)Plasma carotenoids
- Proportion of participants on glucose-lowering medication
- Self-report questionnaires (see section 5.1)

P-values for comparing the two study groups on baseline characteristics will not be reported as per the CONSORT statement (24).

5 Analysis

Primary and secondary outcomes are detailed below. For details on visits and measurements, please see the trial protocol (section 8).

5.1 Outcome definitions

Primary Outcome

- 12 month change from baseline in HbA_{1c}

Secondary Outcomes

- 6 month change from baseline in HbA_{1c}
- 6 and 12 month changes from baseline in body weight, body fat percentage, systolic and diastolic blood pressure, total cholesterol, HDL cholesterol, and LDL cholesterol.
- Probability of achieving good glycaemic control (HbA_{1c} <53mmol/mol) at 6 and 12 months
- Probability of achieving remission (HbA_{1c} <48mmol/mol and not currently prescribed glucose-lowering medication)[2] at 6 and 12 months
- Probability of losing ≥5% and ≥10% of initial body weight at 6 and 12 months
- Modelled cardiovascular risk (UKPDS) at 12 months

Behavioural and Psychosocial secondary outcomes

6 and 12 month changes from baseline in the outcomes specified in Table 1

Table 1. Secondary outcomes and measures.

Outcome	**Measure**	**Objective / self-reported**
Volume of total physical activity	ENMO (mg) (accelerometry)	Objective
Fruit and vegetable intake	Plasma carotenoids	Objective
Dietary restraint (self-reported)	Three Factor Eating Questionnaire – Restraint subscale(25)	Self-reported
Binge eating	The Binge Eating Scale (BES)(26,27)	Self-reported
Control over food cravings	Control of Eating Questionnaire(28)	Self-reported
Diabetes-related quality of life	Audit of Diabetes Dependent Quality of Life - ADDQoL(29)	Self-reported
Capability/Wellbeing	ICEpop CAPability measure for Adults (ICECAP-A)(30)	Self-reported
Health-related quality of life	EuroQol- 5 Dimension (EQ5D-5L)(31)	Self-reported
Dietary intake (Total energy intake in kilojoules)	Food Frequency Questionnaire	Self-reported

[2] This is a deviation from the protocol because we do not have sufficiently granular data to know whether medication was taken in the last 2 months.

Physical activity energy expenditure (PAEE)	Recent Physical Activity Questionnaire (RPAQ)(32,33)	Self-reported

5.2 Analysis methods

Descriptive

For all outcomes, the mean and SD at baseline and 6 and 12 month follow-up will be presented, together with the mean and SD of change from baseline, separately in each randomised group. For continuous variables, means and standard deviations (SDs) will be presented, unless the variable has a skewed distribution, in which case medians, 25th and 75th percentiles will be presented. For categorical variables, the number and percentage of individuals within each category will be presented. For each variable (continuous or categorical), the percentage of missing values will be calculated. For the categorical variables, percentages within sub-categories will be calculated using the number of non-missing values as the denominator.

Primary analysis

The intervention effect on HbA_{1c} at 12 months (and 95% CI) will be estimated from a random intercepts linear regression model, using measures of change from baseline in HbA_{1c} at 6 months and 12 months as outcomes. The model will include intervention group, visit, intervention by visit interaction, centre, the randomisation stratifiers (sex, diabetes duration) and baseline value of HbA_{1c} as fixed effects, and random intercepts to allow for the repeated measures on each individual:

$$(\text{Change in } HbA_{1c} \text{ from baseline to timepoint } t)_i = \beta_0 + u_i + \beta_1 \times group_i + \beta_2 \times (t=6)_i + \beta_{12} \times (t=6)_i * group_i + \beta_3 \times \text{baseline } HbA_{1ci} + \text{randomisation stratifiers} + \varepsilon_{it}$$

i=individual, t=6 or 12 months (follow-up timepoint).
$(t=6)_i$ *= 1 if timepoint=6 months, 0 otherwise.*
u_i *has a normal distribution with mean 0, variance* σ_u^2 *and represents the (level 2) between-individual error.*
ε_{it} *has a normal distribution with mean 0, variance* σ^2 *and represents the overall (level 1) residual error.*

Then β_1 *represents the baseline-adjusted difference in* HbA_{1c} *change at 12 months comparing intervention vs control.*

NOTE: $\beta_1 + \beta_{12}$ represents the baseline-adjusted difference in HbA_{1c} change at 6 months comparing intervention vs control.

This analysis will be repeated also adjusting for glucose-lowering medication (number, dose, frequency) as one effect of these treatment programmes may be a reduction in the prescription of glucose-lowering medication. In addition to adjusting for glucose-lowering medication, we will report the proportion of participants that have been prescribed glucose-lowering medication at 6 and 12 months in each group.

We will conduct analyses with all observed data; random intercept models use all available data and assume missing data are missing at random. To explore this assumption, we will summarise the baseline characteristics of participants with and without missing outcome data.

Missing Data

Missing values of HbA_{1c} at baseline
Participants with a missing baseline value of HbA_{1c} will be included in the analysis using the missing indicator method (34), which is a valid method for pre-randomisation measures in trials, ensuring that, other than participants with missing outcome data (see below), no further participants are excluded, thereby maximising precision of the effect size estimates.

Missing values of HbA1c at 12 months
Participants with missing values of HbA_{1c} at 12 months will be excluded (i.e., a complete-case analysis which assumes outcome data are missing completely at random).

If there are > 10% of participants with missing values of HbA_{1c} at 12-month follow-up, a sensitivity analysis will be performed using multiple imputation by chained equations (MICE) – this assumes data are missing at random. The multiple imputation model will include values of HbA1c at baseline as well as other baseline characteristics that have univariate associations with missingness ($p<0.2$). We will run the MICE procedure with 10 cycles/iterations per dataset, to create 20 imputed datasets *(35,36)*. Analyses will then be run on the imputed datasets and pooled by Rubin's rules *(37)*.

Missing values of objective physical activity
To minimise potential bias in the assessment of total physical activity measured using the accelerometer, we will only include participants in the analysis who meet the following minimum wear time criterion:

- total wear>48hrs, ≥9hrs in each quadrant (where the 24 hour period is split into the following quadrants: midnight to 6am, 6am-12pm, 12pm-6pm, 6pm to midnight).

For participants who meet this criterion, we will use diurnal bias adjustment for any missing physical activity data.

Secondary analyses

The same approach will be used for continuous secondary outcomes. For secondary binary outcomes (glycaemic control; diabetes status; proportion losing ≥5% and ≥10% of initial body weight), a similar approach based on a random intercepts logistic regression model will be used.

Subgroup analyses

Potential interactions between the intervention effect and gender, index of multiple deprivation, educational qualification, and duration of diabetes (<1 year; 1-3 years) on the primary outcome will be examined by including the relevant multiplicative parameters in the model (e.g. 2-way interaction: intervention x gender). The relevant main effects will also be included in the model (e.g., intervention, gender). Educational qualification will be dichotomized into a variable grouping all education categories up to and including A-levels as 'below post-secondary' and categories above A-levels as 'post-secondary and above'. IMD will be added to the interaction terms as a continuous variable. If the data are compatible with an interaction effect, we will dichotomise IMD into most deprived (quintiles 1-5) and least deprived (quintiles 6-10) (38) and present intervention effect with 95% CIs within each subgroup.

Where data are compatible with an interaction effect ($p<0.05$), the intervention effect and 95% CIs within each subgroup will be presented (39).

Per-protocol analyses

To assess whether the findings are influenced by the degree of intervention completion, we will redo the primary outcome analysis (12 month HbA_{1c}) including only those in the DEW group who completed both dietitian calls and at least 75%[3] (i.e. 18/24) of WW meetings (or online equivalent) and those in the DE group who attended DESMOND or completed MyDESMOND.

Sensitivity analysis

We will conduct four sensitivity analyses to assess whether the estimated intervention effect on the primary outcome differs:

[3] in line with Tier 2 weight management service contracts, as informed by WW.

a) depending on whether the intervention was delivered during the COVID-19 pandemic or not,

b) depending on mode of intervention delivery (remote vs. in-person vs. hybrid),

c) depending on measurement method for the primary outcome, HbA_{1c} and

d) when taking the follow-up duration into account.

a) COVID-19: We will explore the influence of the COVID-19 pandemic by conducting a sub-group analysis to assess whether intervention effectiveness differed depending on whether participants' interventions took place primarily after or before the onset of the pandemic.

DEW: Since duration of access to WW sessions in DEW was 6 months, participants who received their referral 3 months or less prior to the onset of the pandemic (23rd March, i.e. when the first lockdown was announced by the prime minister) will be considered to have spent ≥50% of their intervention during the pandemic (i.e. "primarily post onset of COVID-19 pandemic").

DE: For DE, intervention duration was shorter since, pre-pandemic, it involved only one full-day session (or two half-day sessions). Therefore, participants who received their referral 1 month or less prior to the onset of the pandemic (23rd March) will be considered "primarily post onset of COVID-19 pandemic".

b) Intervention delivery mode: Due to the COVID-19 pandemic, interventions in both study groups were only available in remote delivery mode from March 2020. We will explore how mode of delivery affected intervention effectiveness in sub-group analyses for three groups: in-person, remote, mixed. Definitions for modes of delivery are provided in Table 2.

Table 2. Definitions for mode of delivery for the two study groups, DEW and DE.

	DEW	**DE**
In-person	All WW sessions completed in-person (regardless of whether online materials were also accessed)	DESMOND session(s) attended in-person (regardless of whether MyDESMOND was also accessed)
Remote	All WW sessions completed remotely OR only online materials	Only MyDESMOND accessed; no in-person

	accessed (no in-person WW sessions)	DESMOND session attended
Mixed	Some WW sessions attended online and some in person	N/A
Did not access Intervention	No WW sessions attended	Did not attend DESMOND or access MyDESMOND

c) Measurement method for HbA_{1c}: HbA_{1c} measurements may be derived from three different sources:

- Measurement at the GP practice by a research nurse using a blood sample
- Home-testing finger-prick blood sample kit
- HbA_{1c} value obtained from medical notes review

We will redo the primary analysis separately for each group of participants depending on how their HbA_{1c} value was assessed.

For the subgroups detailed above (a-c), we will redo the primary outcome analysis in each of the subgroups, reporting effect estimates and 95% Cis for each group.

d) Duration of follow-up: We will redo the primary analysis adjusting for the follow-up duration for HbA_{1c}.

Inputs for health economic modelling

To inform long-term health economic modelling, we will provide the regression coefficients and the covariance matrix for the primary and secondary analyses. Additionally, we will redo the primary outcome analysis including interaction terms for weight loss, systolic blood pressure, total cholesterol and HDL cholesterol.

Checking assumptions

Prior to analyses, we will assess the assumptions underlying linear and logistic regression.

Linear regression

Normality of residuals (within-person error and between-person error):
Normality will be assessed by visually inspecting the frequency distribution of the standardised residuals. Variables with a skewed distribution will be log-transformed, and the distribution of the residuals then re-checked for normality. Departures from the normality assumption should not affect the validity of the method as the sample size is large (40).

Heteroscedasticity: Heteroscedasticity will be explored by visually examining the regression plot plotting the standardized residuals of the outcome against the standardized predicted values of the model.

Logistic regression

We will examine goodness of fit of logistic regression models by examining the area under the curve (AUC) of the Receiver Operating Characteristic (ROC) curve.

5.3 Safety data

Any Adverse Events (AE) and Serious Adverse Events (SAE) will be reported by randomised group (number and percentage of individuals with at least 1 AE, total number of AEs, number and percentage of individuals with at least 1 SAE, total number of SAEs).

AEs are recorded from the start of the intervention only if they result in cessation of the intervention and the event is considered related to the intervention, or they are related to measurement procedures. All AEs are filed in the study's Trial Master File (TMF) and added to MRC Epidemiology Unit's central log of adverse events.

SAEs are recorded from the start of the intervention. If an SAE is deemed **unrelated** or **expected** by the PIs, it will be documented on the SAE form but no further reporting is required. If an SAE is deemed as **related** and **unexpected**, the SAE is subject to expedited reporting. All documentation will be filed in the TMF and added to the log of adverse events.

5.4 Statistical software

R version 4.1.2 and R Studio version 1.0.153 (or newer versions if these are released before analyses start).

6 References

1. Hex N, Bartlett C, Wright D, Taylor M, Varley D. Estimating the current and future costs of Type1 and Type2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. Diabet Med. 2012;29(7):855–62.

2. HQIP, NHS, DUK. National Diabetes Audit , 2015-16 Report 2a : Complications and Mortality. 2017;(March).

3. Ali S, Stone MA, Peters JL, Davies MJ, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: A systematic review and meta-analysis. Vol. 23, Diabetic Medicine. 2006. p. 1165–73.

4. Kanavos P, Aardweg S Van Den, Schurer W. Diabetes Expenditure, Burden of Disease and Management in 5 EU Countries. LSE Heal. 2012;(January):1–113.

5. Ribaric G, Buchwald JN, McGlennon TW. Diabetes and weight in comparative studies of bariatric surgery vs conventional medical therapy: A systematic review and meta-analysis. Obes Surg. 2014;24(3):437–55.

6. Sjöström L, Peltonen M, Jacobson P, Ahlin S, Andersson-Assarsson J, Anveden Å, et al. Association of bariatric surgery with long-term remission of type 2 diabetes and with microvascular and macrovascular complications. JAMA. 2014;311(22):2297–304.

7. Steven S, Hollingsworth K, Al-Mrabeh A, Avery L, Aribisala B, Caslake M, et al. Very-Low-Calorie Diet and 6 Months of Weight Stability in Type 2 Diabetes: Pathophysiologic Changes in Responders and Nonresponders. Diabetes Care. 2016;March 21.

8. Lean MEJ, Leslie WS, Barnes AC, Brosnahan N, Thom G, McCombie L, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): An open-label, cluster-randomised trial. The Lancet. 2017;

9. Gregg EW, Chen H, Wagenknecht LE, Clark JM, Delahanty LM, Bantle J, et al. Association of an intensive lifestyle intervention with remission of type 2 diabetes. JAMA. 2012;308(23):2489–96.

10. Long H G, Cooper AJM, Wareham J N, Griffin J S, Simmons K R. Healthy Behavior Change and Cardiovascular Outcomes in Newly Diagnosed Type 2 Diabetic Patients: A Cohort Analysis of the ADDITION-Cambridge Study. Diabetes Care. 2014;37(6):1712–20.

11. Wadden TA. Eight-year weight losses with an intensive lifestyle intervention: The look AHEAD study. Obesity. 2014;22:5–13.

12. NICE NG28. Type 2 diabetes in adults: management. Natl Inst Heal Care Excell. 2015;(December 2015):1–86.

13. Davies MJ, Heller S, Skinner TC, Campbell MJ, Carey ME, Cradock S, et al. Effectiveness of the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cluster randomised controlled trial. BMJ. 2008;336(7642):491–5.

14. Pillay J, Armstrong M, Butalia S, Donovan L, Sigal R, Vandermeer B, et al. Behavioral Programs for Type 2 Diabetes Mellitus: A Systematic Review and Network Meta-analysis. Ann Intern Med [Internet]. 2015 Dec 1 [cited 2021 Jul 9];163(11):848–60. Available from: https://pubmed.ncbi.nlm.nih.gov/26414227/

15. Gillett M, Dallosso HM, Dixon S, Brennan A, Carey ME, Campbell MJ, et al. Delivering the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cost effectiveness analysis. BMJ. 2010;341:c4093.

16. Hartmann-Boyce J, Johns DJ, Jebb SA, Summerbell C, Aveyard P. Behavioural weight management programmes for adults assessed by trials conducted in everyday contexts: systematic review and meta-analysis. Obes Rev. 2014 Nov;15(11):920–32.

17. Ahern AL, Wheeler GM, Aveyard P, Boyland EJ, Halford JCG, Mander AP, et al. Extended and standard duration weight-loss programme referrals for adults in primary care (WRAP): a randomised controlled trial. Lancet. 2017 Jun 3;389(10085):2214–25.

18. Meads DM, Hulme CT, Hall P, Hill AJ. The cost-effectiveness of primary care referral to a UK commercial weight loss programme. Clin Obes. 2014;4(6):324–32.

19. O'Neil PM, Miller-Kovach K, Tuerk PW, Becker LE, Wadden TA, Fujioka K, et al. Randomized controlled trial of a nationally available weight control program tailored for adults with type 2 diabetes. Obesity. 2016;24(11):2269–77.

20. Ahern AL, Woolston J, Wells E, Sharp SJ, Islam N, Lawlor ER, et al. Clinical and cost-effectiveness of a diabetes education and behavioural weight management programme versus a diabetes education programme in adults with a recent diagnosis of type 2 diabetes: study protocol for the Glucose Lowering through Weight management (GLoW) randomised controlled trial. BMJ Open [Internet]. 2020 Apr 28 [cited 2022 Apr 20];10(4). Available from: https://pubmed.ncbi.nlm.nih.gov/32350016/

21. Griffin SJ, Borch-Johnsen K, Davies MJ, Khunti K, Rutten GE, Sandbæk A, et al. Effect of early intensive multifactorial therapy on 5-year cardiovascular outcomes in individuals with type 2 diabetes detected by screening (ADDITION-Europe): a cluster-randomised trial. Lancet. 2011 Jul;378(9786):156–67.

22. Bender R, Lange S. Adjusting for multiple testing - When and how? J Clin Epidemiol [Internet]. 2001 [cited 2021 Jan 20];54(4):343–9. Available from: https://pubmed.ncbi.nlm.nih.gov/11297884/

23. Li G, Taljaard M, Van den Heuvel ER, Levine MA, Cook DJ, Wells GA, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. - PubMed - NCBI. Int J Epidemiol [Internet]. 2016 [cited 2020 May 15];46(2):527–39. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28025257

24. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. BMJ. 2010 Mar 27;340(7748):698–702.

25. Westenhoefer J, Stunkard AJ, Pudel V. Validation of the flexible and rigid

control dimensions of dietary restraint. Int J Eat Disord. 1999;26:53–64.

26. Gormally J, Black S, Daston S, Rardin D. The assessment of binge eating severity among obese persons. Addict Behav. 1982;7(1):47–55.

27. Grupski AE, Hood MM, Hall BJ, Azarbad L, Fitzpatrick SL, Corsica JA. Examining the binge eating scale in screening for binge eating disorder in bariatric surgery candidates. Obes Surg. 2013;23(1):1–6.

28. Dalton M, Finlayson G, Hill A, Blundell J. Preliminary validation and principal components analysis of the Control of Eating Questionnaire (CoEQ) for the experience of food craving. Eur J Clin Nutr. 2015;69(12):1313–7.

29. Bradley C, Todd C, Gorton T, Symonds E, Martin A, Plowright R. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: The ADDQoL. Qual Life Res. 1999;8(1–2):79–91.

30. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. Qual Life Res [Internet]. 2012 Feb [cited 2021 May 12];21(1):167–76. Available from: https://pubmed.ncbi.nlm.nih.gov/21598064/

31. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res [Internet]. 2011 Dec [cited 2022 May 30];20(10):1727–36. Available from: https://pubmed.ncbi.nlm.nih.gov/21479777/

32. Golubic R, May AM, Benjaminsen Borch K, Overvad K, Charles MA, Diaz MJT, et al. Validity of electronically administered Recent Physical Activity Questionnaire (RPAQ) in ten European countries. PLoS One. 2014;9(3).

33. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. Int J Epidemiol. 2002;31(1):168–74.

34. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med [Internet]. 2005 Apr 15 [cited 2021 Feb 15];24(7):993–1007. Available from: http://doi.wiley.com/10.1002/sim.1981

35. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: What is it and how does it work? Int J Methods Psychiatr Res. 2011 Mar;20(1):40–9.

36. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007 Sep 5;8(3):206–13.

37. Rubin DB. Multiple Imputation for Nonresponse in Surveys [Internet]. Rubin DB, editor. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 1987 [cited

2020 May 18]. (Wiley Series in Probability and Statistics). Available from: http://doi.wiley.com/10.1002/9780470316696

38. Lee JA, Meacock R, Kontopantelis E, Gittins M, Matheson J. Deprivation and primary care funding in Greater Manchester after devolution: a cross-sectional analysis. Br J Gen Pract [Internet]. 2019 Nov 1 [cited 2022 Apr 19];69(688):e794–800. Available from: https://bjgp.org/content/69/688/e794

39. Wang R, Ware JH. Detecting Moderator Effects Using Subgroup Analyses. Prev Sci. 2013;14(2):111–20.

40. Lumley T, Diehr P, Emerson S, Chen L. The Importance of the Normality Assumption in Large Public Health Data Sets. Annu Rev Public Health. 2002 May;23(1):151–69.