STATISTICAL ANALYSIS PLAN

FULL TITLE OF THE TRIAL: The impact of a holistic multifactorial management intervention in adults with early-onset type 2 diabetes (aged 16 to 45 years): The M3 randomised controlled trial.

SHORT TRIAL TITLE / ACRONYM: The M3 Study

SAP VERSION NUMBER AND DATE: 1.0 (12/08/2026)

IRAS NUMBER: 322040

SPONSOR: University of Leicester

SPONSOR REFERENCE NUMBER: 0827

TRIAL REGISTRATION: ISRCTN16890855

SAP AUTHOR: Dr Danielle Bodicoat

FUNDER(S): This study is funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (NIHR201165). The views expressed are those of the author(s) and not necessarily those of the NIHR or Department of Health and Social Care.

This trial is part of the following programme: Type 2 Diabetes (T2D) in 16 – 40 year olds: A **M**ultifactorial **M**anagement Intervention to Address **M**ultimorbidity in Early-Onset T2D in Adults (The M3 Research Programme).

FUNDED BY
NIHR | National Institute for Health and Care Research

DOCUMENT VERSION HISTORY

Version	Date	Section(s) Changed	Summary of Changes
1.0	12 August 2026	N/A	N/A - Initial version

SIGNATURE PAGE

The undersigned confirm that this Statistical Analysis Plan is scientifically sound, consistent with the IRAS protocol, and meets the requirements of ICH E9(R1), ICH E6(R3), and applicable regulatory guidelines.

I agree to ensure that the confidential information contained in this document will not be used for any other purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of the Sponsor.

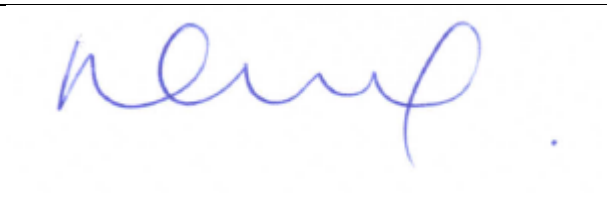
Chief Investigator/Principal Investigator for Lead Site:	
Signature:	Date:
Name: Professor Melanie Davies	
Lead Statistician:	
 Signature:	Date: 21/10/26
Name: Professor Laura Gray	

TABLE OF CONTENTS

1 LIST OF ABBREVIATIONS 6
2 PURPOSE OF THIS DOCUMENT 6
3 STUDY BACKGROUND 7
3.1 Background and Rationale 7
3.2 Research Question and Objectives 8
3.2.1 Research Question 8
3.2.2 Primary Objective 8
3.2.3 Secondary and Exploratory Objectives 8
3.3 Overall Design 9
3.4 Randomisation 11
3.5 Blinding 11
3.6 Treatment Arms 11
3.6.1 Intervention – The M3 Intervention 11
3.6.2 Control – Usual Care 11
3.7 Sample Size 11
3.8 Schedule of Assessments 11
3.8.1 Blood and Urine Testing 12
3.8.2 Glycaemic Level 12
3.8.3 Multimorbidity Management and Burden 12
3.8.4 Anthropometry and Body Composition 12
3.8.5 Blood Pressure (Systolic and Diastolic) 12
3.8.6 Psychosocial Wellbeing and Self-management support 12
3.8.7 Physical Activity, Sedentary Behaviour, and Sleep 14
3.8.8 Physical Function 14
3.8.9 Habitual Diet Questionnaire 15
3.8.10 Medical Management and Multimorbidity Status 15
3.8.11 Habitual Smoking and Alcohol Status 16
3.8.12 Self-Reported Sleepiness 16
3.8.13 Self-Reported Risk of Obstructive Sleep Apnoea (OSA) 17
3.8.14 Handgrip Strength 17
3.8.15 Disability and Life Participation 17
3.8.16 Health Economics Questionnaires 17
4 ENDPOINTS AND ESTIMANDS FRAMEWORK 17
4.1 Intercurrent Events and Strategies 17
4.2 Population 18
4.3 Treatment 19
4.4 Additional Estimand Details 19
4.4.1 Primary Endpoint 19
4.4.2 Secondary Endpoints 19
4.4.3 Exploratory Endpoints 20

4.5 Endpoint Summary 25
5 QUALITY CONTROL 28
6 GENERAL ANALYSIS METHODS 28
6.1 Analysis Sets 28
6.2 Significance Level and Confidence Intervals 29
6.3 Multiplicity 29
6.3.1 Primary Multiplicity Control 29
6.3.2 Secondary and Exploratory Analyses 29
6.4 Handling of Missing Data 29
6.4.1 Expected Patterns and Mechanisms 29
6.4.2 Missing Baseline Values 29
6.4.3 Missing Covariate Values 29
6.4.4 Missing Endpoint Data 29
6.4.5 Safety Data 30
6.4.6 Date Imputation 30
6.5 Covariates in Analyses 30
6.6 Multicentre Considerations 30
6.7 Visit Windows 30
6.8 Statistical Software 31
7 DESCRIPTIVE STUDY ANALYSES 31
7.1 Trial Population / Disposition / Analysis Populations 31
7.2 Baseline Characteristics 31
7.3 General Medical History 31
7.4 Prior and Concomitant Medications 32
7.5 Measurements of Treatment Compliance 32
7.6 Protocol Deviations 33
7.7 Safety Data 33
8 EFFICACY ANALYSES 33
8.1 Primary Endpoint 33
8.1.1 Analysis Population 33
8.1.2 Model Specification 33
8.1.3 Hypothesis 34
8.1.4 Presentation 34
8.1.5 Sensitivity Analyses 34
8.1.6 Subgroup Analyses 35
8.2 Secondary Endpoints 35
8.3 Exploratory Endpoints 35
9 TABLES, LISTINGS, AND FIGURES (TLF) PLAN 36
10 SAP DEVIATIONS AND AMENDMENTS 36
10.1 Changes from the IRAS Protocol 36
10.2 Changes from the SAP 37
10.2.1 Planned vs. Unplanned Deviations 37

10.2.2 Amendment Process 37
11 APPENDIX A: Study Within a Trial (SWAT) 37
11.1 Rationale 37
11.2 SWAT Aim and Objectives 38
11.3 SWAT Methods 38
11.4 SWAT Inclusion and Exclusion Criteria 38
11.5 Sample Size Calculation 38
11.6 Methods of Allocation 38
11.7 Outcome Measures 39
11.8 SWAT analysis 39
12 PUBLICATION 39
13 REFERENCES 40

1 LIST OF ABBREVIATIONS

Abbreviation	Term
AE	Adverse Event
BMI	Body Mass Index
CGM	Continuous Glucose Monitoring
CI	Chief Investigator
CPRD	Clinical Practice Research Datalink
CRF	Case Report Form
CRP	C-Reactive Protein
DMSES	Diabetes Management Self-Efficacy Scale
DSAS-2	Type 2 Diabetes Stigma Assessment Scale
eGFR	Estimated Glomerular Filtration Rate
EQ-5D-5L	EuroQuol 5 Dimension 5 Levels (Quality of Life Questionnnaire)
HbA1c	Glycated Haemoglobin
GLP-1RA	Glucagon like peptide-1 receptor agonist
HDL	High-density Lipoprotein
IMD	Index of multiple deprivation
ITT	Intention to Treat
LDL	Low-density Lipoprotein
MAR	Missing At Random
MI	Multiple Imputation
MMRM	Mixed Model for Repeated Measures
MTBQ	Multimorbidity Treatment Burden Questionnaire
MVPA	Moderate-to-vigorous intensity physical activity
PAID	Problem Areas in Diabetes Scale
PHQ-9	Patient Health Questionnaire
PP	Per Protocol
PROM	Patient Reported Outcome Measure
OSA	Obstructive Sleep Apnea
RCT	Randomised Control Trial
SAP	Statistical Analysis Plan
SCS	Self-Compassion Scale
SF-36	36-Item Short Form Survey
SGLT2i	Sodium-glucose cotransporter-2 inhibitor
SPPB	Short Physical Performance Battery
SWAT	Study Within a Trial
T2D	Type 2 Diabetes
TMF	Trial Master File
UACR	Urine to Albumin Creatinine Ratio
UKDDQ	The UK diabetes and diet questionnaire
WHODAS	WHO Disability Assessment Schedule

2 PURPOSE OF THIS DOCUMENT

This Statistical Analysis Plan (SAP) defines the pre-specified statistical methodology for the analysis of the M3 study. It has been prepared in accordance with the CONSORT statement

for RCTs, ICH E9: Statistical Principles for Clinical Trials, ICH E9(R1): Addendum on Estimands and Sensitivity Analysis, and ICH E6(R3): Good Clinical Practice.

The SAP was finalised and version-controlled prior to database lock. Any amendments after approval of version 1.0 must be documented and justified. No formal stopping rules have been pre-defined and no interim efficacy analyses will be conducted.

The SAP is only intended to cover the main statistical analyses for the study, as well as the Study Within A Trial (SWAT). The SAP does not cover the following parts of the M3 trial:

- Feasibility assessment (an interim analysis conducted prior to the SAP being drafted).
- Process evaluation (a mixed-methods evaluation described in the study IRAS protocol).
- Health economic analyses (a separate analysis plan will be developed for these analyses by SCHARR).
- CPRD analyses (exploratory analyses described in the study IRAS protocol).

The analyses described in this SAP will be conducted after hard database lock.

3 STUDY BACKGROUND

3.1 Background and Rationale

There is an increasing prevalence of type 2 diabetes (T2D) in early adulthood (16-40 years, termed "early-onset adult T2D").[1,2] Adults with early-onset T2D appear to have a pre-disposition to accelerated disease progression[3-5] and display an extreme risk phenotype compared with those diagnosed later in life.[1,3,6-15] These are exacerbated by difficulties instigating and maintaining optimal self-care practices.[1,5,14,16]

These factors, coupled with the longer exposure to the harmful effects of hyperglycaemia, lead to a higher risk of adverse physical and mental health outcomes, poorer quality of life, and early mortality.[1,6-10,13-15,17-42] Consequently, early-onset T2D has huge implications for healthcare use,[37,38] complicated by poorer adherence to treatment,[43-45] as well as indirect societal costs.[38,46] Furthermore, people from minority ethnic backgrounds, with lower socioeconomic status, and/or with learning disability syndromes are disproportionately represented in early-onset T2D.[38,47-49]

While there is an urgent need for tailored, effective care for people with early-onset T2D,[47,50] there is a dearth of targeted interventions tailored to their needs. This may explain current care delivery being met with low concordance and higher levels of dissatisfaction.[37,51-53]

Evidence (predominantly in older adults) suggests that intensive multifactorial interventions, combing pharmacological and non-pharmacological management strategies can exert long-term benefits in adults with T2D, including reductions in HbA1c, cardio-renal complications,

and mortality.[54-59] Collectively, multicomponent interventions may prove effective in improving health outcomes for people with early-onset T2D, and may be further advanced when combined with digital components involving support, education, and wearable technology. It remains unknown whether combining all these components in a multi-component invention alongside a holistic, minimally disruptive approach,[60,61] and tailoring the components to the individual needs of people with early-onset T2D, could provide additional benefit, whilst remaining cost-effective.

3.2 Research Question and Objectives

3.2.1 Research Question

Is an innovative multifactorial management intervention that is tailored specifically to the needs of adults with early-onset T2D ("the M3 intervention") effective and cost-effective at improving diabetes management and holistic health (including addressing multimorbidity) in this high-risk group?

3.2.2 Primary Objective

To test the effectiveness of the M3 intervention, compared to standard care, on diabetes management (assessed via HbA1c) at 52 weeks in adults with early-onset T2D.

3.2.3 Secondary and Exploratory Objectives

1) To test the effectiveness of the M3 intervention, compared to standard care, on diabetes management (assessed via HbA1c) at 26 and 104 weeks in adults with early-onset T2D.

2) To test the effectiveness of the M3 intervention in adults with early-onset T2D, compared to standard care, on a range of secondary and exploratory holistic health domains at 26, 52 and 104 weeks.

3) To test the cost-effectiveness of the M3 intervention, compared to standard care, in adults with early-onset T2D.

4) To evaluate key constructs of implementation (fidelity, dose delivered and reach) and mechanisms of impact (recruitment, exposure and satisfaction) of the M3 intervention.

5) To evaluate the impact of a pragmatic remote data collection approach versus in-person data collection on data quality and participant retention.

Objective 3 will be addressed by the health economic analyses and Objective 4 will be addressed by the process evaluation; therefore, these two objectives will not be addressed in this SAP. Objective 5 will be addressed by the SWAT (see Appendix A; Section 11).

3.3 Overall Design

The M3 study is a 104-week, multi-centre, open-label, parallel-group randomised controlled trial (RCT), with embedded health economic analysis, internal feasibility assessment, mixed-methods process evaluation, and SWAT. This SAP only addresses the analyses for the main RCT and the SWAT.

Individuals aged 16 to 45 years (inclusive) with T2D diagnosed before age 40 are eligible (see study protocol for full list of eligibility criteria). Eligible participants are randomised 1:1 to one of two groups to receive the intervention (the M3 intervention) or control (ongoing standard care) for 104 weeks.

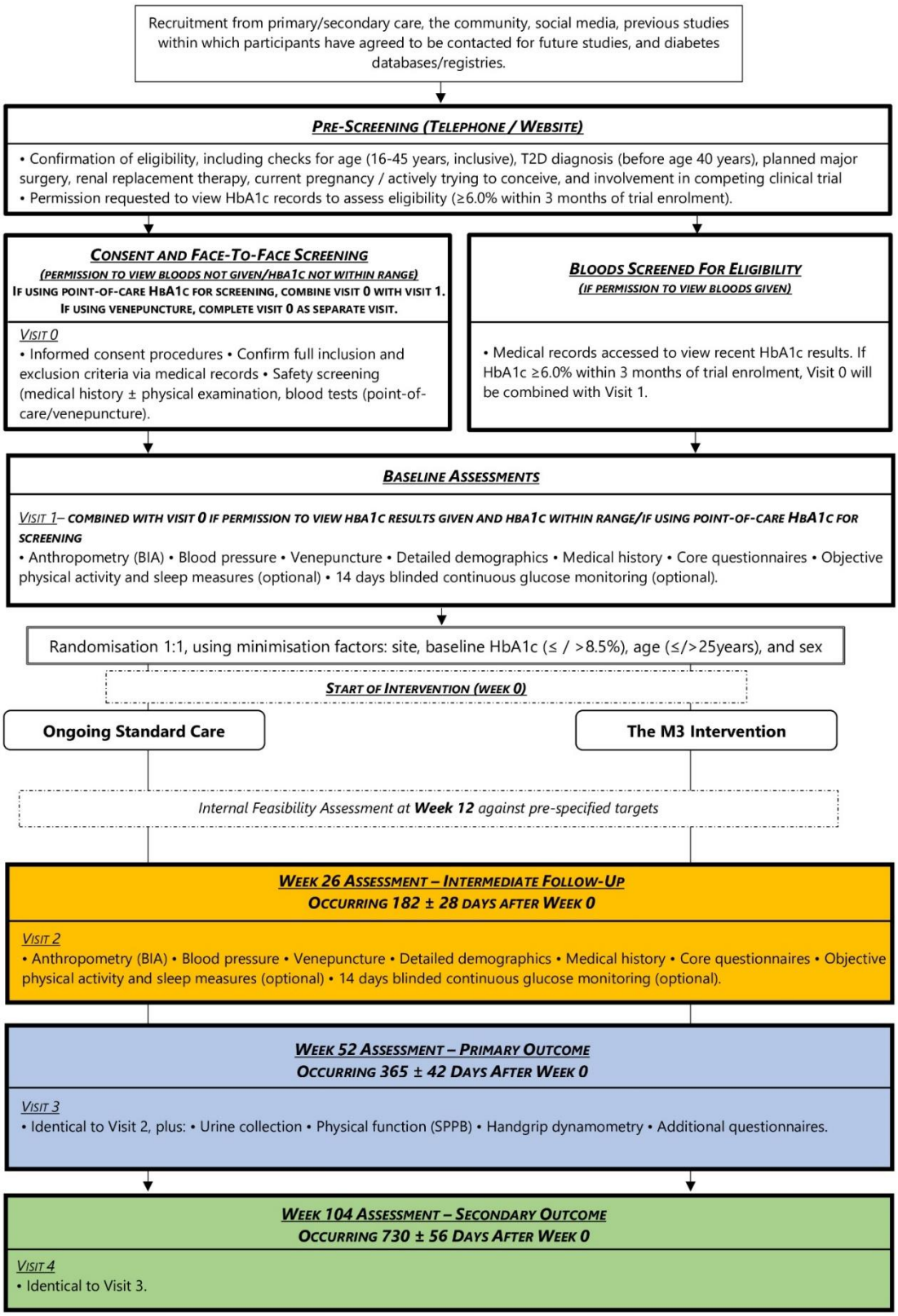
The planned sample size for the main trial is 246 participants (123 in control and 123 in intervention). No specific sample size is set for the SWAT.

Participants are recruited at three research centres (Leicester, Derby, and Imperial). The visit schedule is shown in Figure 3-1 and is as follows:

- Pre-screening via telephone/website.
- Visit 0 (Up to Day -42) – Consent and face-to-face screening (Can be combined with Visit 1).
- In-person trial assessments:
    - Visit 1 – Week 0 (Day 0) – Baseline assessments.
    - Visit 2 – Week 26 (Day 182±28) – Intermediate follow-up assessments.
    - Visit 3 – Week 52 (Day 365±42) – Primary follow-up assessments (including primary endpoint).
    - Visit 4 – Week 104 (Day 730±56) – Extended follow-up assessments.
- Additionally, there are brief study contacts to review consent and collect AEs at Week 13 (Day 91±14), Week 39 (Day 274±14), Week 65 (Day 456±14), Week 78 (Day 547±14), and Week 91 (Day 638±14).
- An unscheduled visit can be used when a participant misses their appointment, is outside of their visit window due to unforeseen circumstances, or repeat assessments are required.

Trial procedures and study outcomes related to the trial (performed for all participants) are distinct from measures undertaken as part of the M3 intervention (only assessed for the intervention group during intervention visits). To reduce participant and researcher burden, where overlap between the trial assessments and intervention measures occurs, duplicate measures are not performed. Where intervention measures are used as trial assessments, these are entered into the study database as an ongoing process and do not require a separate data source for analysis.

Figure 3-1. Trial Flow Chart.



3.4 Randomisation

Randomisation occurs after baseline assessments. Eligible participants are randomly assigned in a 1:1 ratio to one of two arms: 1) **Intervention group:** The M3 intervention, OR 2) **Control group:** Usual care under their usual diabetes service providers.

The randomisation uses minimisation. The minimisation factors are:

- **Trial site:** Leicester, Derby, or Imperial.
- **Baseline HbA1c value:** ≤8.5% (≤69mmol/mol) or >8.5% (>69mmol/mol).
- **Age:** ≤ 25 years or >25 years.
- **Sex:** Male or female.

3.5 Blinding

The nature of the intervention prevents blinding of participants and intervention providers. Outcome assessors will be blinded to the primary outcome of HbA1c.

3.6 Treatment Arms

3.6.1 Intervention – The M3 Intervention

The M3 intervention is a multi-component, complex intervention for people with early-onset T2D. The combination of components and targets is tailored for each individual participant. A fuller description of the M3 intervention is provided in the study IRAS protocol. In brief, the key features of the M3 intervention are: delivery of care; technology; targeting risk factor domains; promotion of newer therapies; and earlier, more intensive intervention.

3.6.2 Control – Usual Care

Ongoing standard care is delivered by participants' usual NHS service providers in primary and secondary care and is not altered or standardised as part of the trial.

3.7 Sample Size

The planned sample size for the main trial is 246 participants (123 in control and 123 in intervention). Based on previous data,[31,62,63] we require data from 86 participants per group to detect a 0.5% absolute difference in HbA1c at 12 months, with 90% power, assuming a SD of 1.0% and two-sided 5% alpha. Assuming 30% dropout rate, we will recruit 246 participants. A 0.5% difference in HbA1c is chosen as this reflects a clinically meaningful difference, associated with 9% reduction in CVD risk.[64]

3.8 Schedule of Assessments

See IRAS protocol for further details on assessments and the schedule of procedures.

3.8.1 Blood and Urine Testing

At all visits, blood samples are taken to measure circulating biomarkers (including HbA1c). If venepuncture is not possible at 52 weeks, HbA1c may be obtained via electronic medical records. Where HbA1c is obtained via medical records, these data are entered into the study database as an ongoing process and do not require a separate data source for analysis.

At the Week 52 and Week 104 visits, a urine sample is also collected. Fasting assessments of glucose, insulin, c-peptide, and triglycerides are optional.

3.8.2 Glycaemic Level

At all visits, blinded continuous glucose monitoring (CGM) is used to assess:

- **Time in range:** % of time between 3.9–10.0 mmol/L.[65]
- **Time above range:** % of time >10.1 mmol/L.[65]
- **Time below range:** % of time <3.8 mmol/L.[65]

Participants wear a CGM for 14 days. Only valid days of CGM wear (defined as valid data for ≥70% of the day, i.e. ≥16.8 hours) will be analysed.[66,67] There is no limit on the number of valid days required for CGM data to be analysed.

3.8.3 Multimorbidity Management and Burden

At all visits, the 10-item Multimorbidity Treatment Burden Questionnaire (MTBQ) is used to assess the treatment burden.[68] Each item is scored 0 (Not difficult/Does not apply) to 4 (Extremely difficult), with a global score calculated by finding the average score (sum of scores for all answered questions / number of questions answered) then multiplying by 25 to produce a 0-100 range. Participants with more than 50% missing responses should be excluded from the score calculation. Higher scores indicate worse burden, with scores categorised as: 0 (No burden), <10 (Low), 10-22 (Medium), and >22 (High).

3.8.4 Anthropometry and Body Composition

At all visits, body weight (to the nearest 0.1 kg), BMI, body fat percentage, lean body mass, and fat mass are measured.

3.8.5 Blood Pressure (Systolic and Diastolic)

At all visits, arterial blood pressure is measured. Three measurements are obtained and the average of the last two measurements will be used. If only two measurements are available then the average of those will be used. If only one measurement is available then that will be used.

3.8.6 Psychosocial Wellbeing and Self-management support

At all visits, the following questionnaires are completed.

3.8.6.1 Self-compassion

The Self-Compassion Scale (SCS), a 26-item questionnaire with a 5-point Likert scale (1= almost never to 5= almost always),[69] comprises three positive sub-scales (self-kindness, common humanity, and mindfulness), three negative sub-scales (self-judgment, isolation, and over-identification), and a global scale. For each score, scores of 1.0 to 2.5, 2.6 to 3.5, and 3.6 to 5.0 indicate low, moderate, and high self-compassion, respectively.

The mean for the overall score and each subscale will be calculated using the score guidance:

- Self-Kindness Items: Questions 5, 12, 19, 23, 26
- Self-Judgment Items (reverse scored): Questions 1, 8, 11, 16, 21
- Common Humanity Items: Questions 3, 7, 10, 15
- Isolation Items (reverse scored): Questions 4, 13, 18, 25
- Mindfulness Items: Questions 9, 14, 17, 22
- Over-identification Items (reverse scored): Questions 2, 6, 20, 24
- To reverse score items (1=5, 2=4, 3=3, 4=2, 5=1).
- To compute a total self-compassion score, first reverse score the negative subscale items - self judgment, isolation, and over-identification. Then take the mean of each subscale, and compute a total mean (the average of the six subscale means).

Researchers can report subscale scores with or without reverse-coding. This study will report reverse-coded subscale scores so that for all SCS scores, higher scores indicate more self-compassion. The guidance does not state how to handle missing data items. For consistency with the DSAS-2 processing, participants with >10% missing responses for a score will be excluded from the calculation of that score.

3.8.6.2 Diabetes-related distress

The Problem Areas in Diabetes (PAID) questionnaire uses a 5-point Likert scale (0 = not a problem to 4 = serious problem).[70] Total scores are calculated by summing the items, dividing by 80, and multiplying by 100, with higher scores indicating higher distress. If there are missing data items, compute the average of the answered items (e.g. if 19/20 are answered, sum the 19 items and divide by 19), then multiply the average score by 20 to get the total score. For consistency with the DSAS-2 processing, if >10% of data items are missing then the PAID score will not be calculated.

3.8.6.3 Depression

The Patient Health Questionnaire (PHQ-9)[71] includes nine statements rated from 0 (Not at all) to 3 (Nearly every day). The total score is calculated by summing the nine responses. The

score ranges from 0-27, with higher scores indicating higher severity of depression. Following the PHQ-9 scoring guidance, if one or two responses are missing then the total score will be calculated by summing the scores of the answered items, dividing by the number of answered items, and then multiplying by 9 to get an estimated total score. If three or more responses are missing then the PHQ-9 score will not be calculated.

3.8.6.4 T2D-related stigma

The T2D Stigma Assessment Scale (DSAS-2)[72] has 19 items scored on a 5-point Likert scale (1 = Strongly disagree to 5 = Strongly agree). The total score is calculated by summing each item score to give a range of 19 to 95, with higher scores indicating higher levels of stigma. The three subscale scores will also be calculated by summing the following items:

- Treated Differently (6 items): Items 1, 4, 7, 10, 14, 17.
- Blame and Judgment (7 items): Items 2, 3, 5, 8, 12, 16, 19.
- Self-stigma (6 items): Items 6, 9, 11, 13, 15, 18.

Following the DSAS-2 scoring guidance, if <10% of the items are missing within a specific subscale/total score then the missing items will be replaced with the average score of the completed items for that participant. If ≥10% of the items are missing within a specific subscale/total score then the subscale/total score will not be calculated.

3.8.6.5 Diabetes self-efficacy

The Diabetes Management Self-Efficacy Scale (DMSES)[73] has 15 items scored from 0 (Cannot do at all) to 10 (Certain can do). The items are summed to produce an overall score, ranging from 0-200, with higher scores indicating higher self-efficacy. In line with the DMSES guidance, if more than two items are missing in total, the scale will not be scored. If there are one or two missing data items, then the missing values will be replaced with the mean of the participant's observed answers.

3.8.7 Physical Activity, Sedentary Behaviour, and Sleep

At all visits, participants can opt in to assessment of physical activity, sedentary behaviour, and sleep objectively using an accelerometer worn for approximately seven full days. Data will be processed using an open-source R programme (GGIR), that applies all required data processing rules to produce analysable endpoints.

3.8.8 Physical Function

At Weeks 52 and 104, the Short Physical Performance Battery (SPPB) is used to objectively measure physical function.[74] It comprises three tests (chair sit-to-stand; standing balance; gait speed), each scored from 0 (unable) to 4 (best) for a maximum total score of 12 (higher scores indicate higher function). In line with SPPB guidance, a score of 0 will be assigned to a test if it was not attempted due to safety or inability. If a test was not completed for other

reasons (e.g. participant refusal) then the value will not be imputed and the total score cannot be calculated.

3.8.9 Habitual Diet Questionnaire

At Weeks 52 and 104, the 20-item UK Diabetes and Diet Questionnaire (UKDDQ)[75] is used to assess recent dietary habits. Each item is scored 0 (healthiest) to 5 (least healthy). The total score is calculated by summing the scores for each of the 20 items, giving a score ranging from 0 to 100, where higher scores indicate a less healthy diet. Following a standard approach for food frequency questionnaires, missing data items will be assumed to be zero consumption. The UKDDQ guidance does not state how many complete items are needed for a valid response. For consistency with the DSAS-2 processing, UKDDQ score will not be calculated for participants with >10% missing responses.

3.8.10 Medical Management and Multimorbidity Status

At each timepoint, changes to the following are recorded:

- The use of glucose-lowering therapies (including addition, removal, or dose-adjustment), with the following definitions:
    - **Glucose-lowering therapies:** Any medication listed under "Diabetes medication" on CRF_9 (Concomitant Medications), which includes alpha-glucosidase inhibitors, biguanides, DPP4 inhibitors, GLP-1 receptor agonists, meglitinides, SGLT2 inhibitors, sulphonylureas, thiazolidinediones, and insulin, as well as tirzepatine.
    - **Addition of glucose-lowering therapies:** Binary (yes/no); "yes" if the number of glucose-lowering therapies is higher at Follow-up than baseline. The number and class of added therapies will also be summarised.
    - **Removal of glucose-lowering therapies:** Binary (yes/no); "yes" if the number of glucose-lowering therapies is lower at Follow-up than baseline. The number and class of removed therapies will also be summarised.
    - For each of the above binary variables, if the visit was completed and no glucose-lowering therapies are reported then it will be assumed that the participant was not using such therapies. Therefore, data will only be missing for these variables if the participant did not complete the study visit.
    - **Dose adjustment for glucose-lowering therapies:** Continuous variable defined as total Defined Daily Dose (DDD) per week calculated by summing the number of DDDs per week (derived using WHO Reference DDD Values) across all glucose-lowering therapies.
- Multimorbidity status (including additional diagnoses, remission, or changes in severity),[76] number, and presence of conditions identified as 'core' conditions in a recent consensus on multimorbidity measurement,[76] with the following definitions:
    - **Core multimorbidity condition:** Any condition that should always or usually be included in a multimorbidity measure[76], except for diabetes as that is an

inclusion criterion (i.e. stroke, coronary artery disease, heart failure, peripheral artery disease, heart valve disorders, arrhythmia, venous thromboembolic disease, aneurysm, hypertension, Addison's disease, cystic fibrosis, thyroid disorders, chronic obstructive pulmonary disease, asthma, bronchiectasis, Parkinson's disease, epilepsy, multiple sclerosis, paralysis, transient ischaemic attack, peripheral neuropathy, chronic primary pain, solid organ cancers, haematological cancers, metastatic cancers, melanoma, benign cerebral tumours that can cause disability, dementia, schizophrenia, depression, anxiety, bipolar disorder, drug or alcohol misuse, eating disorder, autism, post-traumatic stress disorder, connective tissue disease, osteoarthritis, long term musculoskeletal problems due to injury, osteoporosis, gout, chronic liver disease, inflammatory bowel disease, chronic pancreatic disease, peptic ulcer, chronic kidney disease, end stage kidney disease, endometriosis, chronic urinary tract infection, anaemia (including pernicious anaemia, sickle cell anaemia), vision impairment that cannot be corrected, hearing impairment that cannot be corrected, Meniere's disease, HIV/AIDS, chronic Lyme disease, tuberculosis, post-acute covid-19, congenital disease, chromosomal abnormalities).

- **Additional diagnoses of comorbidities:** Binary (yes/no); "yes" if any core multimorbidity conditions are listed as newly diagnosed at Follow-up. The number of diagnoses will also be summarised.
- **Remission of comorbidities:** Binary (yes/no); "yes" if any core multimorbidity conditions are listed as having gone into remission at Follow-up.
- **Improvement in severity of comorbidities:** Binary (yes/no); "yes" if ≥1 core multimorbidity condition is listed as having improved in severity at Follow-up.
- **Worsening in severity of comorbidities:** Binary (yes/no); "yes" if ≥1 core multimorbidity condition is listed as having worsened in severity at Follow-up.
- For each of the above binary variables, if the visit was completed and no change in comorbidities are reported then it will be assumed that the participant did not undergo any changes to comorbidity status. Therefore, data will only be missing for these variables if the participant did not complete the study visit.

3.8.11 Habitual Smoking and Alcohol Status

At Weeks 52 and 104, self-reported habitual alcohol consumption and smoking status are determined during clinical review.

3.8.12 Self-Reported Sleepiness

At Weeks 52 and 104, the 8-item Epworth Sleepiness Scale (ESS) is measured.[77] Each item is rated on a scale of 0-3 (three being highly soporific). A total score is obtained by summing the individual item scores, with higher scores indicating higher daytime sleepiness. A score between 0-9 is indicative of no-daytime sleepiness, 10-12 borderline, and >12 describes

excessive daytime sleepiness. The ESS scoring guidance states that if one or more item scores are missing then ESS is invalid. Therefore, the ESS score will not be calculated if any item scores are missing.

3.8.13 Self-Reported Risk of Obstructive Sleep Apnoea (OSA)

At Weeks 52 and 104, the 8-item STOP-BANG Questionnaire is collected. Each question is a yes/no question. The score is calculated by summing the number of "yes" responses. A score of 0-2 indicates low risk, 3-4 indicates medium risk, and 5-8 indicates a high probability of moderate-to-severe OSA.[78] In line with STOP-BANG guidance, if one or two responses are missing then the response will be assumed to be "no". If more than two responses are missing then the STOP-BANG score will not be calculated.

3.8.14 Handgrip Strength

At Weeks 52 and 104, handgrip strength is measured.[79] The maximum of the readings generated is taken as the maximum grip strength.

3.8.15 Disability and Life Participation

At Weeks 52 and 104, the 12-item WHO Disability Assessment Schedule (WHODAS 2.0) is administered.[80] Each item is scored 0 (None) to 4 (Extreme). The total score is obtained by summing all item scores, with higher scores indicating greater disability. Following WHODAS 2.0 guidance, if one item is missing then the answered items will be summed, divided by 11 to find the average, and multiplied by 12 to generate a total score. If more than one item is missing then the total score will not be calculated.

3.8.16 Health Economics Questionnaires

The following questionnaires are collected for the health economics analyses and will not be analysed as part of the main study analyses described in this SAP: Modular Resource Use Measure; Work Productivity and Activity Impairment; EQ-5D-5L questionnaire; and SF-36 questionnaire.

4 ENDPOINTS AND ESTIMANDS FRAMEWORK

Sections 4.1-4.4 describe the estimand framework, with endpoints summarised in Section 0.

4.1 Intercurrent Events and Strategies

Intercurrent events are events that occur after randomisation and affect the interpretation of the endpoint. The below table outlines the intercurrent events identified for this study.

For all endpoints, a treatment policy estimand will be used with intercurrent events handled as per the below table. For the primary endpoint, a hypothetical estimand will also be defined, whereby participants experiencing discontinuation of randomised treatment,

withdrawal from follow-up/loss to follow-up, or death prior to the Week 52 assessment will be excluded from the Per Protocol analysis set (Sections 4.4.1 and 6.1).

Intercurrent event	Definition	Approach for treatment policy estimand	Justification
Discontinuation of randomised treatment	Intervention participants who wish to revert to usual care. Disengagement from usual care is not classified as discontinuation of randomised treatment to reflect real-world practice.	Endpoint data retained and analysed in randomised group.	To align with ITT approach by analysing participants in their allocated arm.
Withdrawal from follow-up/loss to follow-up	Early withdrawal from the study prior to the Week 104 visit.	Missing endpoint data handled using MMRM, with MI applied residually for participants with no post-baseline data (Section 6.4).	Missing data imputation is the missing data method, not the estimand strategy itself.
Death	All-cause mortality prior to endpoint assessment.	Observed pre-death data retained and analysed in randomised group; post-death data will not be imputed.	Post-death there are no HbA1c data to observe, therefore data are truncated, not missing.
Pregnancy	Participants who become pregnant following randomisation will not be withdrawn. Their treatment will be changed based on appropriate treatment during pregnancy, however it will continue to reflect the holistic approach of the intervention.	Endpoint data retained and analysed in randomised group.	To align with ITT approach by analysing participants in their allocated arm.

4.2 Population

A treatment policy estimand will be applied for all endpoints. Therefore, the analysis set for the main analyses will be the Intention-To-Treat (ITT) population.

For the primary endpoint, a hypothetical estimand will also be specified, addressed through the Per Protocol sensitivity analyses (Section 8.1.5).

For optional measures (as per Section 4.4), the analysis set will be the ITT population with at least one post-baseline observation. Unlike non-optional endpoints, residual MI will not be applied for participants with no post-baseline data for these measures, given the higher expected rates of missingness due to their optional nature.

For exploratory endpoints, the analysis set will be the ITT population limited to complete cases as no modelling will be conducted for the exploratory endpoints.

See Section 6.1 for descriptions of the analysis sets and Section 8.1.5 for sensitivity analyses.

4.3 Treatment

The M3 intervention (Section 3.6.1) will be compared with usual care (Section 3.6.2) over the specified timepoint.

4.4 Additional Estimand Details

4.4.1 Primary Endpoint

Attribute	Treatment Policy Estimand	Hypothetical Estimand
Population	ITT	PP
Variable (Endpoint)	Change from baseline to Week 52 in HbA1c. The raw values will be analysed as continuous variables.	Change from baseline to Week 52 in HbA1c. The raw values will be analysed as continuous variables.
Intervention	M3 vs usual care	M3 vs usual care
Intercurrent events (Section 4.1)	Discontinuation of treatment / pregnancy: data retained in randomised group. Withdrawal: data retained where available, missing data handled via MMRM, with MI where no post-baseline values available. Death: data not imputed.	Discontinuation of treatment: intervention participants who revert to usual care included up to point of intercurrent event then excluded. Withdrawal and death prior to Week 52: excluded. Pregnancy: Analysed in allocated arm.
Population Summary	Adjusted mean difference (intervention minus control) plus 95% CI	Adjusted mean difference (intervention minus control) plus 95% CI

4.4.2 Secondary Endpoints

All secondary endpoints follow the treatment policy estimand comparing M3 with usual care, with intercurrent events handled as per Section 4.1. No hypothetical estimand is specified for secondary endpoints.

4.4.2.1 HbA1c

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 in HbA1c 2) Change from baseline to Week 104 in HbA1c The raw values will be analysed as continuous variables.
Population Summary	Adjusted mean difference (intervention minus control) plus 95% CI

4.4.2.2 Multimorbidity Management, Burden, and Health-Related Quality-of-Life

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 in the global MTBQ score 2) Change from baseline to Week 52 in the global MTBQ score 3) Change from baseline to Week 104 in the global MTBQ score The global MTBQ score will be calculated as per Section 3.8.3.
Population Summary	Adjusted mean difference (intervention minus control) plus 95% CI

4.4.2.3 Cardiometabolic and Renal Risk

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26

	2) Change from baseline to Week 52 3) Change from baseline to Week 104 In each of the following: • Body weight • BMI • Lean body mass • Fat mass • Systolic blood pressure • Diastolic blood pressure • Total cholesterol • High-density lipoprotein (HDL) • eGFR The raw values will be analysed as continuous variables.
Population Summary	Adjusted mean difference (intervention minus control) plus 95% CI

4.4.2.4 Psychosocial Wellbeing and Self-Management Support

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 2) Change from baseline to Week 52 3) Change from baseline to Week 104 In each of the following: • Self-compassion (SCS – six subscales and overall score) • Diabetes-related distress (PAID – overall score) • Depression (PHQ-9 – overall score) • T2D-related stigma (DSAS-2 – three subscales and overall score) • Diabetes self-efficacy (DMSES – overall score) All scores will be calculated as per Section 3.8.6.
Population Summary	Adjusted mean difference (intervention minus control) plus 95% CI

4.4.2.5 Physical activity

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 in device-measured habitual daily steps 2) Change from baseline to Week 52 in device-measured habitual daily steps 3) Change from baseline to Week 104 in device-measured habitual daily steps Habitual daily steps (measured and calculated as per Section 3.8.7) is an optional measure and so will likely have a higher amount of missing data. Therefore, residual MI will not be conducted (Section 6.4).
Population Summary	Adjusted mean difference (intervention minus control) plus 95% CI

4.4.3 Exploratory Endpoints

4.4.3.1 Other Glycaemic Measures

Attribute	Specification

Variable (Endpoint)	1) Change from baseline to Week 26 2) Change from baseline to Week 52 3) Change from baseline to Week 104 In each of the following: • CGM-assessed percentage of time spent within target range • CGM-assessed percentage of time spent above target range • CGM-assessed percentage of time spent below target range • Fasting plasma glucose • Fasting plasma insulin • Fasting circulating C-Peptide CGM data are collected and processed as per Section 3.8.2. For glucose, insulin, and c-peptide raw values will be analysed.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.2 Body Fat Percentage

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 in body fat percentage 2) Change from baseline to Week 52 in body fat percentage 3) Change from baseline to Week 104 in body fat percentage The raw values will be analysed as continuous variables.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.3 Other Lipid Profile Measures

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 2) Change from baseline to Week 52 3) Change from baseline to Week 104 In each of the following: • Low-density lipoprotein (LDL) • Triglycerides The raw values will be analysed as continuous variables.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.4 Liver Function

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 in each attribute of the liver enzyme panel 2) Change from baseline to Week 52 in each attribute of the liver enzyme panel 3) Change from baseline to Week 104 in each attribute of the liver enzyme panel The raw values will be analysed as continuous variables.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.5 Other Renal Function Measures

Attribute	Specification
Variable (Endpoint)	1) Absolute value of UACR at Week 52 2) Absolute value of UACR at Week 104 The raw values will be analysed as continuous variables.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.6 Circulating Inflammatory Profile

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 in CRP 2) Change from baseline to Week 52 in CRP 3) Change from baseline to Week 104 in CRP The raw values will be analysed as continuous variables.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.7 Other Lifestyle Related Measures

Attribute	Specification
Variable (Endpoint)	1) Absolute value at Week 52 2) Absolute value at Week 104 For each of the following: • Self-reported habitual alcohol consumption (units currently drink during a typical week, which will be set to 0 for participants who report they are "Never" or "Ex" alcohol drinkers) • Self-reported sleepiness (ESS – overall score) • Self-reported OSA (STOP-BANG – overall score) • Self-reported habitual dietary intake (UKDDQ questionnaire – overall score) Self-reported alcohol consumption will be obtained as per Section 3.8.11. ESS score will be calculated as per Section 3.8.12. STOP-BANG score will be calculated as per Section 3.8.13. UKDDQ score will be calculated as per Section 3.8.9.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

Attribute	Specification
Variable (Endpoint)	1) Self-reported smoking status at Week 52 2) Self-reported smoking status at Week 104 Smoking status is collected as "Never smoked", "Ex-smoker", and "Current smoker" (Section 3.8.11). The "Never smoked" and "Ex-smoker" categories will be combined to derive a binary smoking status (Non-smoker/Smoker) variable.
Population Summary	Frequency count and percentage

4.4.3.8 Other Measures of Physical Activity, Inactivity, Function, and Disability Risk

Attribute	Specification
Variable (Endpoint)	1) Change from baseline to Week 26 2) Change from baseline to Week 52 3) Change from baseline to Week 104

	In each of the following: • Device-measured time spent performing moderate-to-vigorous intensity physical activity (MVPA) • Device-measured time spent sedentary • Device-measured sleep duration • Device-measured sleep quality Endpoints will be derived following standard processing rules as per Section 7.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

Attribute	Specification
Variable (Endpoint)	1) Absolute value at Week 52 2) Absolute value at Week 104 In each of the following: • SPPB – overall score • Handgrip strength • World Health Organisation Disability Assessment Schedule (WHODAS 2.0 – overall score) SPPB score will be calculated as per Section 3.8.8. Handgrip strength will be measured as per Section 3.8.14, with the maximum value analysed. WHODAS score will be calculated as per Section 3.8.15.
Population Summary	If approximately normally distributed then mean and SD, else median and IQR

4.4.3.9 Changes to Use of Glucose-Lowering Therapies

Attribute	Specification
Variable (Endpoint)	1) At Week 26 2) At Week 52 3) At Week 104 For each of the following: • Addition of glucose-lowering therapies • Removal of glucose-lowering therapies • Dose adjustment for glucose-lowering therapies Each variable will be derived as per Section 3.8.10.
Population Summary	Frequency count and percentage for addition/removal of therapies. For dose adjustment, if approximately normally distributed then mean and SD, else median and IQR

4.4.3.10 Change in Multimorbidity Status

Attribute	Specification
Variable (Endpoint)	1) Status at Week 26 2) Status at Week 52 3) Status at Week 104 For each of the following: • Additional diagnoses of comorbidities

	• Remission of comorbidities • Improvement in severity of comorbidities • Worsening in severity of comorbidities Each binary (yes/no) variable will be derived as per Section 3.8.10.
Population Summary	Frequency count and percentage

4.5 Endpoint Summary

Endpoint	Measure	Week(s)	Classification	Estimand	Missing Data Handling
Glycaemic					
HbA1c	Change from baseline	52	Primary	Treatment policy (primary); Hypothetical (PP sensitivity)	Primary: MMRM under MAR then residual MI.[a] PP set used for hypothetical estimand sensitivity analysis (complete cases).
HbA1c	Change from baseline	26, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
CGM: % of time in target range	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
CGM: % of time above target range	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
CGM: % of time below target range	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Fasting plasma glucose	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Fasting plasma insulin	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Fasting circulating C-peptide	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Cardiometabolic and Renal					
Body weight	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
BMI	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Lean body mass	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Fat mass	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Systolic blood pressure	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Diastolic blood pressure	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Total cholesterol	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
HDL	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
eGFR	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Body fat percentage	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
LDL	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Triglycerides	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
UACR	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
Liver Function					
Liver enzyme panel	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Circulating Inflammatory Profile					
CRP	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Multimorbidity Management, Burden, and Health-Related Quality of Life					
MTBQ global score	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI

Endpoint	Measure	Week(s)	Classification	Estimand	Missing Data Handling
Additional diagnoses of comorbidities	Status	26, 52, 104	Exploratory	Treatment policy	Complete cases
Remission of comorbidities	Status	26, 52, 104	Exploratory	Treatment policy	Complete cases
Improvement in severity of comorbidities	Status	26, 52, 104	Exploratory	Treatment policy	Complete cases
Worsening in severity of comorbidities	Status	26, 52, 104	Exploratory	Treatment policy	Complete cases
Psychosocial Wellbeing and Self-Management Support					
SCS (6 subscales and overall score)	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
PAID (overall score)	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
PHQ-9 (overall score)	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
DSAS-2 (3 subscales and overall score)	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
DMSES (overall score)	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR then residual MI
Physical Activity, Inactivity, Function, and Disability					
Habitual daily steps	Change from baseline	26, 52, 104	Secondary	Treatment policy	MMRM under MAR without residual MI
MVPA	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Sedentary time	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Sleep duration	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
Sleep quality	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases
SPPB (overall score)	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
Handgrip strength (maximum)	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
WHODAS 2.0 (overall score)	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
Lifestyle					
Alcohol consumption	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
Smoking status	Status	52, 104	Exploratory	Treatment policy	Complete cases
ESS (overall score)	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
STOP-BANG (overall score)	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
UKDDQ (overall score)	Absolute value	52, 104	Exploratory	Treatment policy	Complete cases
Glucose-Lowering Therapies					
Addition of glucose-lowering therapies	Status	26, 52, 104	Exploratory	Treatment policy	Complete cases
Removal of glucose-lowering therapies	Status	26, 52, 104	Exploratory	Treatment policy	Complete cases
Dose adjustment for glucose-lowering	Change from baseline	26, 52, 104	Exploratory	Treatment policy	Complete cases

Endpoint	Measure	Week(s)	Classification	Estimand	Missing Data Handling
Therapies					

Abbreviations: BMI, Body Mass Index; CGM, Continuous Glucose Monitoring; CRP, C-Reactive Protein; DMSES, Diabetes Management Self-Efficacy Scale; DSAS-2, Type 2 Diabetes Stigma Assessment Scale; eGFR, estimated Glomerular Filtration Rate; ESS, Epworth Sleepiness Scale; HDL, High-Density Lipoprotein; ITT, Intention-to-Treat; LDL, Low-Density Lipoprotein; MAR, Missing At Random; MTBQ, Multimorbidity Treatment Burden Questionnaire; MVPA, Moderate-to-Vigorous Physical Activity; OSA, Obstructive Sleep Apnoea; PAID, Problem Areas in Diabetes Scale; PHQ-9, Patient Health Questionnaire-9; PP, Per Protocol; SCS, Self-Compassion Scale; SPPB, Short Physical Performance Battery; T2D, Type 2 Diabetes; UACR, Urine Albumin-to-Creatinine Ratio; UKDDQ, UK Diabetes and Diet Questionnaire; WHODAS, WHO Disability Assessment Schedule.

Notes: All estimands use a treatment policy strategy for intercurrent events (Section 4.1). The hypothetical estimand for the primary endpoint (HbA1c at Week 52) is addressed by the PP sensitivity analysis (Section 8.1.5). For exploratory endpoints, no inferential testing is conducted; results are presented descriptively only, therefore missing data are not imputed and a complete cases approach is used.

[a] Missing data will predominantly be handled through the MMRM assuming MAR. MI will be used to impute data for participants with no post-baseline data (expected to be a minimal number of participants).

5 QUALITY CONTROL

Quality control procedures will be applied throughout the study. Data will be reviewed on an ongoing basis for completeness, consistency, and plausibility, with queries resolved prior to database lock. The study database contains validation ranges for continuous outcomes that do not prevent data being entered outside of the range, however, if data falls outside of the range, then a warning is generated so the data can be double checked for mistakes.

Where feasible, analysis programmes for key outputs will be verified by a second programmer/statistician. The final SAP will be reviewed and approved by at least one other statistician. Data will be checked for outliers and missing values and validated using the defined score ranges. Any discrepancies identified during quality control will be documented and resolved before results are finalised.

There is no medical coding for medical conditions or medications. Medications will be transcribed as written by the patient/clinician with no interpretation added by the data clerk.

6 GENERAL ANALYSIS METHODS

6.1 Analysis Sets

- **Intention-To-Treat (ITT):** All randomised participants analysed in their allocated arm (regardless of received treatment). The missing data approach will be ensure all randomised participants are analysed (Section 6.4), except for exploratory endpoints and habitual daily steps (an optional secondary endpoint).

- **Per Protocol:** All participants in the ITT population, excluding participants 1) in the intervention arm who revert to usual care, 2) with major IRAS protocol deviations (see below), and 3) with missing primary endpoint data. Participants will be analysed according to allocated arm because of the pragmatic nature of the intervention and control (e.g. GLP-1RAs can be prescribed under both treatment arms, therefore there is not a clear cross-over point between the two groups). The Per Protocol analysis set is used solely for the hypothetical estimand for the primary endpoint (Section 4.5.1).

- **Safety:** Given the holistic and tailored nature of the intervention and control, it is not possible to define a minimum dose. Therefore, a separate Safety population is not defined and all safety analyses will use the ITT population.

The following major IRAS protocol deviations require exclusion from the Per Protocol set:

- Failure to meet all inclusion/exclusion criteria at randomisation, i.e. participant randomised in error.

6.2 Significance Level and Confidence Intervals

A two-sided significance level of alpha = 0.05 will be used for all hypothesis tests. All confidence intervals (CIs) will be two-sided at the 95% level.

6.3 Multiplicity

6.3.1 Primary Multiplicity Control

As there is a single primary endpoint, no multiplicity adjustments will be made.

6.3.2 Secondary and Exploratory Analyses

Secondary endpoints are considered supportive. Results will be presented with nominal p-values and 95% CIs without adjustment for multiplicity. Exploratory endpoints are hypothesis-generating only and only descriptive analyses with no statistical testing will be performed.

6.4 Handling of Missing Data

6.4.1 Expected Patterns and Mechanisms

Missing data may arise from participant withdrawal, missed single items within a visit, missed study visits within an otherwise continuing participant, participant death, loss to follow-up, and insufficient data to derive endpoints (e.g. see Section 3.8 for conditions under which missing data item for PROMs will mean that a score cannot be calculated).

The missing data mechanism is assumed to be MAR for the primary endpoint because if venepuncture is not possible at Week 52 then HbA1c is obtained via electronic medical records where possible. This means that availability of the primary endpoint is less likely to be associated with completion of the trial, making a MAR assumption reasonable.

6.4.2 Missing Baseline Values

Baseline value is required for the MMRMs. If the baseline value is missing then it will imputed using the same MI set up as specified in Section 6.4.4, except that where baseline itself is the variable being imputed, it will be excluded from its own imputation model and replaced by any available pre-baseline or screening values as auxiliary variables.

6.4.3 Missing Covariate Values

The covariates in the MMRMs will be the baseline value and the minimisation factors. See Section 6.4.2 for missing baseline values. The minimisation factors cannot be missing for participants in the ITT analysis set as they are required for randomisation.

6.4.4 Missing Endpoint Data

The primary approach to handling missing outcome data will be the mixed model for repeated measures (MMRM), which uses all observed post-baseline data from each participant under a missing at random (MAR) assumption. Participants with data at one or

more post-baseline visits will contribute to the MMRM regardless of whether data are missing at other timepoints, making explicit imputation unnecessary for these participants.

Multiple imputation (MI) under MAR will be applied as a residual approach for participants with no observed post-baseline data for a given endpoint. These participants are expected to be few, comprising primarily those who withdrew from follow-up, are lost to follow-up or died before their first post-baseline assessment. They would be excluded from the MMRM and would not contribute to the analysis, which is inconsistent with the treatment policy estimand. MI will be used to generate plausible post-baseline values for these participants, who will then be included in the MMRM. The imputation model will include: randomised group, visit, baseline value of the endpoint, and minimisation factors. Fifty imputations will be generated. The MMRM will be run on each imputed dataset and estimates combined using Rubin's rules.

6.4.5 Safety Data

Missing safety data will not be imputed. Only observed data will be presented.

6.4.6 Date Imputation

Partial dates will be imputed as follows:

- Missing day: imputed as the 15th of the month, except for AE start dates which will be imputed as the 1st of the month (conservative, to maximise AE duration).
- Missing month and day: imputed as 1 July of the stated year.
- Fully missing dates: no imputation; flagged in data listings.

6.5 Covariates in Analyses

All models will include the following covariates: randomisation arm (always); baseline value of the outcome variable (continuous outcomes only), and all minimisation factors (as listed in Section 3.4). Continuous covariates will be included as continuous variables. Minimisation factors will be included as categorical fixed effects.

6.6 Multicentre Considerations

The study is conducted at three centres. Centre is one of the minimisation factors and will be included as a fixed effect in all models.

6.7 Visit Windows

Assessments within the defined visit windows (based on the IRAS protocol) will be assigned to that visit. If more than one assessment falls within a window, the one closest to the target date will be used. Assessments outside of a visit window will be assigned to the nearest visit window (unless another assessment also falls within that window) and marked as a deviation.

6.8 Statistical Software

Analyses will be conducted using an appropriate, validated statistical software. All analysis programs will be version-controlled and archived in the Trial Master File (TMF).

7 DESCRIPTIVE STUDY ANALYSES

7.1 Trial Population / Disposition / Analysis Populations

A CONSORT diagram showing the flow of participants will be produced, including the number of people screened, enrolled, randomised, withdrawn, and completed. The number of participants in each analysis set and reasons for exclusion from analysis sets will be reported.

7.2 Baseline Characteristics

Baseline demographics will be summarised for the ITT population by treatment group and overall using number (percentage) for categorical variables and mean (standard deviation) for continuous variables (unless skewed then median and interquartile range will be used). Statistical tests will not be used to compare baseline characteristics by treatment group.

The baseline demographics to be summarised are: the minimisation factors (Section 3.4), age (continuous), ethnicity (categorical), highest level of completed education (categorical), current employment status (categorical), current marital status (categorical), and socioeconomic status (IMD score; continuous).

For categories with <5 participants:

- For ordinal variables the category will be pooled with the next highest category.
- For non-ordinal variables the category will be pooled with the 'Other' category.

7.3 General Medical History

General medical history collected at baseline will be summarised for each of the conditions listed under Part 4 of CRF 3 (Baseline Assessment). Data will be presented by treatment group and overall, using the ITT population. For each condition, the number of participants with the condition of interest will be summarised as count and percentage (assuming missing data are "No" if the baseline visit was completed). If any of the conditions are not present in any of the study participants then this will be noted in the table footnote and the condition will not be included in the table to avoid the potential for null data rows.

Additionally, for conditions reported by ≥5 participants, the following will be summarised:

- Duration in years using mean and SD (based on reported duration, or calculated duration from year of diagnosis/occurrence, depending on which is available; if both are available then the reported duration will be used).

- Number and percentage of participants where the condition of interest is ongoing.

These values will not be summarised for conditions reported by <5 participants to avoid potential identification of participants in the reported data.

7.4 Prior and Concomitant Medications

Prior (reported as Stopped at baseline) and concomitant (reported as Ongoing at baseline) medications collected at baseline will be summarised for each medication reported on CRF 9 (Concomitant Medications). Separate tables will be produced for prior and concomitant medications. Data will be presented by treatment group and overall, using the ITT population. For each drug type, the number of participants using each drug type will be summarised as count and percentage (assuming missing data are "No" if the baseline visit was completed).

Medications will be reported based on drug type. The exception will be diabetes medications which will be reported at a more granular level (i.e. alpha-glucosidase inhibitors, biguanides, DPP4 inhibitors, GLp1 receptor agonists, meglitinides, SGLT2 inhibitors, sulphonylureas, thiazolidinediones, insulin), with drug trade names assigned to the category of diabetes medication based on the drug reference table in CRF 9.

Additionally, for drug types reported by ≥5 participants, the following will be summarised:

- Dose using mean and SD.
- Duration of use (Stop date – Start date for prior medications; Baseline date – Start date for concomitant medications).
- Number and percentage of participants reporting each frequency. If <5 participants are in a category then the category will be pooled with the 'Other' category.
- Number and percentage of participants reporting each route. If <5 participants are in a category then the category will be pooled with the 'Other' category.

These values will not be summarised for drug types reported by <5 participants to avoid potential identification of participants in the reported data.

7.5 Measurements of Treatment Compliance

The intervention targets ≥12 hours of annual contact time per year. However, given the participant-led nature of M3, there will be participants who prefer/require less than this. Therefore, the mean (and other summary statistics) for contact time will be reported for the intervention arm, by year and overall. Due to the participant-led nature of the intervention, there will not be a formal comparison against the target of ≥12 hours of annual contact time.

Contact time will be recorded through appointment ledger bookings on SystmOne (including quarterly reviews and care coordinator check ins) and time spent using digital tools (e.g. MyDesmond and MyHealthMapp). These systems, combined with the appointment ledger, provide reports showing the number/type/duration of appointments per participant.

7.6 IRAS Protocol Deviations

A listing of IRAS protocol deviations will be produced and ordered by treatment arm then deviation type then participant ID. Additionally, IRAS protocol deviations will be summarised by treatment group and for the total population. Both tables will use the ITT population.

No electronic records of IRAS protocol deviations have been captured. The physical records will be collected from sites.

7.7 Safety Data

A listing of AEs will be produced and ordered by treatment arm then AE type then participant ID. Additionally, AEs will be summarised by treatment group and overall. Both tables will use the ITT population, which is equivalent to the Safety population (Section 6.1). No formal hypothesis tests will be conducted for the safety data.

Only AEs that occur from the point of randomisation, are deemed directly related to study activities, and are observed by the investigator or reported by the participant are recorded.

The following information will be included in the listing: AE type, start and stop dates, frequency, outcome, treatment, severity, action, and seriousness.

The same information will be included in the summary table, except that Start and Stop dates will be replaced with length in days (Stop date – Start date, or Final Visit date – Start date for ongoing AEs). The table will summarise serious AEs and all AEs separately.

8 EFFICACY ANALYSES

8.1 Primary Endpoint

8.1.1 Analysis Population

The primary analysis will use the ITT set.

8.1.2 Model Specification

The primary analysis will compare the primary endpoint (HbA1c at Week 52; Section 4.4.1) between the intervention and control groups using a mixed model for repeated measures (MMRM). The MMRM structure will be:

- Fixed effects: randomised group (categorical), visit (categorical), group-by-visit

interaction, baseline HbA1c (continuous), baseline-by-visit interaction, and the minimisation factors (categorical).

- Covariance structure: unstructured (if this fails to converge, an appropriate alternative covariance structure will be selected).

The primary inferential estimate is the group-by-visit interaction contrast at Week 52.

Analyses will be conducted following database lock after 1) the Week 52 visit and 2) the Week 104 visit. At each analysis, the MMRM will be fitted using all data available up to and including that timepoint. Results reported from each analysis will be as follows:

- Week 52 analysis: Week 26 and Week 52 estimates.
- Week 104 analysis: Week 104 estimates only.
- Week 26 and Week 52 estimates from the Week 104 analysis will not be reported, as the primary results for these timepoints are taken from the Week 52 analysis.

Missing outcome data will be handled as described in Section 6.4. The MMRM uses all observed post-baseline HbA1c values under MAR. For participants with no observed post-baseline HbA1c, values will be imputed using MI, as described in Section 6.4.

8.1.3 Hypothesis

$$H_0: \mu_intervention - \mu_control = 0 \text{ (no treatment difference)}$$

$$H_1: \mu_intervention - \mu_control \neq 0 \text{ (two-sided alternative)}$$

The null hypothesis will be rejected if the two-sided p-value is < 0.05.

8.1.4 Presentation

The adjusted mean difference will be reported, with 95% CI and two-sided p-value.

8.1.5 Sensitivity Analyses

Two pre-specified sensitivity analyses of the primary endpoint will be conducted, namely:

- **Per protocol analysis:** The primary analyses will be repeated in the Per Protocol set to address the hypothetical estimand for the primary endpoint, estimating the effect of M3 on HbA1c at Week 52 under the assumption that all participants adhered to their randomised treatment. This complements the primary treatment policy estimand by providing an estimate of efficacy under conditions of full compliance. As participants with missing primary endpoint data are excluded from the Per Protocol set by definition (Section 6.1), no missing data imputation is required for this analysis.
- **Complete case analysis**: The primary analyses will be repeated excluding participants

in the ITT set with missing data for any of the variables required for the primary analysis. This will test the robustness of the primary treatment policy estimand to the MAR assumption underlying the MMRM and MI. The conclusion from the primary treatment policy estimand will be considered robust if directionally consistent results of a similar magnitude are observed for the complete case analysis.

Results from both sensitivity analyses will be tabulated alongside the primary analysis result.

8.1.6 Subgroup Analyses

Subgroup analyses for the primary endpoint are pre-specified as exploratory, intended to assess consistency of the primary analysis result across clinically relevant subgroups. The pre-specified subgroups are baseline HbA1c, age, duration of T2D, and ethnicity. This will be assessed by adding interaction terms between group allocation and the potential effect modifiers (plus the main term if not already included) to the MMRM (Section 8.1.2), one at a time. A forest plot of treatment effects by subgroup will be presented. Subgroup analyses are not powered for formal inference; p-values for treatment-by-subgroup interaction terms will be reported.

Subgroup	Categories	Interaction Test
Baseline HbA1c	≤/>8.5 % [≤/>69 mmol/mol]	Treatment × HbA1c
Baseline age	≤/>25 years	Treatment × Age
Duration of T2D at baseline	Data-driven (two categories to be based on the median value)	Treatment × Duration
Ethnicity	White/Non-White	Treatment × Ethnicity

8.2 Secondary Endpoints

All secondary endpoints (Section 4.4.2) are continuous and so will be analysed using the same approach as the primary endpoint (Section 8.1). The same reporting plan will also be used (i.e. Week 52 analysis: Week 26 and Week 52 estimates; Week 104 analysis: Week 104 estimates only). No sensitivity or subgroup analyses are planned for the secondary endpoints, therefore if any such analyses are conducted then they will be clearly identified as post-hoc exploratory analyses.

8.3 Exploratory Endpoints

To limit multiplicity, for all exploratory endpoints (Section 4.4.3), only descriptive analyses with no statistical testing will be performed. Continuous endpoints will be summarised using mean, standard deviation, median, and interquartile range. Binary endpoints will be summarised using frequency count and percentage. All exploratory analyses will be considered hypothesis generating, rather than providing firm conclusions. No sensitivity or subgroup analyses are planned for the exploratory endpoints, therefore if any such analyses are conducted then they will be clearly identified as post-hoc analyses.

9 TABLES, LISTINGS, AND FIGURES (TLF) PLAN

Given that this is a non-regulatory study, shell TLFs have not been created. However, all TLFs will align with the following standards:

- All table footnotes will use a standard footer to include abbreviations, key definitions, model details and missing data handling (if applicable), analysis programme name and version, and data cutoff date.
- All figures will include axis labels, units, sample size, and a clear legend.

10 SAP DEVIATIONS AND AMENDMENTS

10.1 Changes from the IRAS Protocol

Section	IRAS Protocol Text	SAP Text	Rationale for Change
7.7 Safety Data	All serious adverse events will be tabulated and summarised by treatment group; no formal statistical testing will be performed. All events will be summarised by seriousness, expectedness and relatedness.	Removes relatedness from the AE summary.	Only related AEs are recorded in the study, therefore relatedness is redundant.
8.1.2. Model Specification	The protocol specified that a linear regression model would be used.	A MMRM will be used for primary and secondary analyses. The fixed effects are unchanged.	The MMRM approach makes better use of all available longitudinal data under the MAR assumption and more appropriately accounts for the correlation between repeated measures within participants.
8.2 Secondary and Exploratory Endpoints	Presence of new-onset albuminuria at Week 52 and Week 104 and progression of albuminuria at Week 52 and Week 104 are included as secondary endpoints.	These endpoints are not included in the planned analyses.	Albuminuria is not measured at baseline, therefore it is not possible to derive new-onset albuminuria or progression of albuminuria.
10.8 Adjusted analysis	To maintain randomisation the per-protocol analysis will use instrumental variable methodology (using randomisation as the instrument) to assess how the effectiveness is mediated by compliance to the intervention.	Participants will be analysed according to allocated intervention arm because of the pragmatic nature of both the intervention and control (e.g. GLP-1RAs can be prescribed under both treatment arms, therefore there is not a clear cross-over point between the two groups).	The instrumental variable methodology is not feasible to implement as there is not a treatment compliance definition in the IRAS protocol.

10.2 Changes from the SAP

10.2.1 Planned vs. Unplanned Deviations

A planned deviation is a prospective amendment to this SAP, approved before the affected analysis is performed. An unplanned deviation occurs when it is discovered post-hoc that the analysis differed from the SAP. Both must be documented.

10.2.2 Amendment Process

- Any proposed amendment to the SAP after database lock will be drafted by the Trial Statistician with written justification.
- The amendment will be reviewed by the Principal Investigator and Lead Statistician at a minimum.
- The approved amendment will be issued as a numbered SAP document, version-controlled, and filed in the TMF.
- All deviations from the SAP, whether or not a formal amendment is issued, will be listed in the Statistical Analysis Report.

11 APPENDIX A: Study Within a Trial (SWAT)

11.1 Rationale

Poor retention in trials gives rise to missing data, which can undermine the validity, reliability, and generalisability of RCTs.[81] Poor retention and subsequent challenges with missing data appear to be particularly problematic in trials gathering patient reported outcomes (PROMs).[82,83] Barriers to face-to-face, paper-based outcome completion include lack of time and burden associated with both administering the outcomes (for researchers and healthcare professionals) and completing the measure (for participants).[84]

Using technology to streamline and maximise the collection of outcome data, particularly PROMs, may mitigate the problem of poor retention, and has become increasingly possible with widespread access to Smart phones and tablets over the last decade.[81,84-86] These benefits may be particularly relevant to people with early-onset T2D given the burden of care experienced by this group and the complex lives they lead. Despite this, data that could enhance key trial processes like outcome completion and retention are lacking in this group, because they are poorly represented in clinical trials overall.[87]

There is a paucity of high certainty evidence that can be applied to trials to improve retention, and limited recent data specifically examining the impact of using remote, app-based electronic outcome data collection.[81,85] Consequently, this area is a priority for future methodology research[88,89] and the subject of several ongoing SWATs.[90,91]

11.2 SWAT Aim and Objectives

A self-contained SWAT is embedded in the M3 study to compare face-to-face, paper-based outcome completion with pragmatic outcome completion via a web-based app.

11.3 SWAT Methods

The SWAT was prospectively registered on the Northern Ireland Network for Trials Methodology Research SWAT store as SWAT 215.[91]

The M3 patient-facing app offers a secure method through which participants can share health data. Participation in the SWAT has no impact on the M3 intervention delivery. Individuals randomised to pragmatic (app-based) data collection do not have access to any intervention features provided through the app.

Pragmatic outcome collection involves the flexible use of remote methods of assessment (e.g. via an app) for trial outcomes that can be gathered in this manner appropriately, specifically PROMs, i.e. MTBQ, SCS, PAID, PHQ-9, DSAS-2, DMSES, ESS, STOP-BANG, WHODAS, UKDDQ, self-reported habitual alcohol consumption, and smoking status.

11.4 SWAT Inclusion and Exclusion Criteria

All participants recruited to the control arm of the M3 host trial are eligible for the embedded SWAT. The 3-item Digital Health Care Literacy Scale will be used at screening to determine participants confidence, familiarity, and comfort related to digital solutions. Participants with low scores are supported to use the app as required.[92] See the IRAS protocol for full eligibility details.

11.5 Sample Size Calculation

The SWAT sample size is determined by the number of participants in the control group of the host M3 trial. It is not powered to detect a statistically significant difference between groups. The results will contribute to future meta-analysis examining data collection methods in trials of similar complex interventions and in similar populations.

11.6 Methods of Allocation

All participants randomised to the control group in the main trial are then instantly randomised 1:1 to either:

- Pragmatic data collection with remote collection of appropriate outcomes via the app

Or

- Face-to-face, paper-based data collection of all outcomes during trial visits

Randomisation is separate to that of the minimisation procedures for the main trial (Section 3.4). This second randomisation is not stratified and uses random permuted block randomisation.

Participants retain their allocation to app-enabled data collection or face-to-face data collection at all subsequent follow-ups. Due to the nature of the SWAT, it is not possible to blind outcome assessors or participants to allocation.

11.7 Outcome Measures

The primary outcome for the SWAT is the **proportion of participants concordant with outcome completion at 52 weeks** (defined as returning a fully-completed questionnaire). This outcome will be defined for each questionnaire (i.e. regardless of whether the other questionnaires on the CRF are complete or not).

Secondary outcomes are:

- **Proportion of participants concordant with outcome completion** (returning a fully-completed questionnaire) at 26 (for those questionnaires collected) and 104 weeks.
- **Proportion of participants concordant with at least partial completion** (returning at least a partially-completed questionnaire) at 26 (for those questionnaires collected), 52, and 104 weeks.
- **Number and proportion of missing complete measures:** Each questionnaire will be scored according to the published scoring algorithms. Where the score cannot be calculated due to missing data, it will be classed as a missing complete measure. The mean number of missing complete measures per participant will be summarised, as well as the total number (proportion) of missing complete measures across the SWAT.
- **Number and proportion of missing data items.** The mean number of missing data items per participant will be summarised, as well as the total number (proportion) of missing data items across the SWAT.

11.8 SWAT analysis

To compare paper-based and web-based outcome completion, primary and secondary SWAT outcomes will be reported descriptively by arm. No formal statistical testing will be performed. An equivalent completion rate between the two methods will be considered beneficial, because of the reduced burden and increased flexibility for participants and staff that is associated with the web-based method.[84,93,94]

12 PUBLICATION

It is intended that the primary peer-reviewed manuscript for this trial will report all primary and secondary endpoints. Exploratory endpoints will only be reported in the manuscript if they have baseline data and are part of a grouping that would be insufficient to report as a

secondary manuscript (e.g. glycaemic endpoints) or if removal of the exploratory endpoints from the primary manuscript affect interpretation of the primary or secondary endpoint findings.

13 REFERENCES

1. Lascar N, Brown J, Pattison H, Barnett AH, Bailey CJ, Bellary S. Type 2 diabetes in adolescents and young adults. *The lancet Diabetes & endocrinology*. 2018;6(1):69–80.
2. Wilmot EG, Davies MJ, Yates T, Benhalima K, Lawrence IG, Khunti K. Type 2 diabetes in younger adults: the emerging UK epidemic. *Postgraduate Medical Journal*. 2010;86(1022):711–718.
3. Steinarsson AO, Rawshani A, Gudbjörnsdottir S, Franzén S, Svensson A-M, Sattar N. Short-term progression of cardiometabolic risk factors in relation to age at type 2 diabetes diagnosis: a longitudinal observational study of 100,606 individuals from the Swedish National Diabetes Register. *Diabetologia*. 2018;61(3):599–606.
4. Arslanian S, El Ghormli L, Young Kim J, et al. The shape of the glucose response curve during an oral glucose tolerance test: forerunner of heightened glycemic failure rates and accelerated decline in β-cell function in TODAY. *Diabetes Care*. 2019;42(1):164–172.
5. Luk A, Kong A, Basu A. Young-onset diabetes, nutritional therapy and novel insulin delivery systems: a report from the 21st Hong Kong Diabetes and Cardiovascular Risk Factors–East Meets West Symposium. *Diabetic Medicine*. 2020;37(8):1234–1243.
6. Yeung RO, Zhang Y, Luk A, et al. Metabolic profiles and treatment gaps in young-onset type 2 diabetes in Asia (the JADE programme): a cross-sectional study of a prospective cohort. *The lancet Diabetes & endocrinology*. 2014;2(12):935–943.
7. Chan JC, Lau ES, Luk AO, et al. Premature mortality and comorbidities in young-onset diabetes: a 7-year prospective analysis. *The American journal of medicine*. 2014;127(7):616–624.
8. Hillier TA, Pedula KL. Complications in young adults with early-onset type 2 diabetes: losing the relative protection of youth. *Diabetes care*. 2003;26(11):2999–3005.
9. Sattar N, Rawshani A, Franzén S, et al. Age at diagnosis of type 2 diabetes mellitus and associations with cardiovascular and mortality risks: findings from the Swedish National Diabetes Registry. *Circulation*. 2019;139(19):2228–2237.
10. Wilmot E, Idris I. Early onset type 2 diabetes: risk factors, clinical impact and management. *Therapeutic advances in chronic disease*. 2014;5(6):234–244.
11. Gunathilake W, Song S, Sridharan S, Fernando D, Idris I. Cardiovascular and metabolic risk profiles in young and old patients with type 2 diabetes. *QJM: An International Journal of Medicine*. 2010;103(11):881–884.
12. Carnethon MR, Gidding SS, Nehgme R, Sidney S, Jacobs Jr DR, Liu K. Cardiorespiratory fitness in young adulthood and the development of cardiovascular disease risk factors. *Jama*. 2003;290(23):3092–3100.
13. Lascar N, Altaf Q-A, Raymond NT, et al. Phenotypic characteristics and risk factors in a multi-ethnic cohort of young adults with type 2 diabetes. *Current medical research and opinion*. 2019;35(11):1893–1900.
14. Browne JL, Nefs G, Pouwer F, Speight J. Depression, anxiety and self-care behaviours of young adults with Type 2 diabetes: results from the International Diabetes Management

and Impact for Long-term Empowerment and Success (MILES) Study. *Diabetic Medicine*. 2015;32(1):133–140.
15. Barker MM, Zaccardi F, Brady EM, et al. Age at diagnosis of type 2 diabetes and cardiovascular risk factor profile: A pooled analysis. *World Journal of Diabetes*. 2022;13(3):260.
16. Nanayakkara N, Pease A, Ranasinha S, et al. Younger people with Type 2 diabetes have poorer self-care practices compared with older people: results from the Australian National Diabetes Audit. *Diabetic Medicine*. 2018;35(8):1087–1095.
17. Al-Saeed AH, Constantino MI, Molyneaux L, et al. An inverse relationship between age of type 2 diabetes onset and complication risk and mortality: the impact of youth-onset type 2 diabetes. *Diabetes care*. 2016;39(5):823–829.
18. Song SH, Gray TA. Early-onset type 2 diabetes: high risk for premature diabetic retinopathy. *Diabetes Research and Clinical Practice*. 2011;94(2):207–211.
19. Middleton T, Constantino M, Molyneaux L, et al. Young-onset type 2 diabetes and younger current age: increased susceptibility to retinopathy in contrast to other complications. *Diabetic Medicine*. 2020;37(6):991–999.
20. Lv X, Ran X, Chen X, et al. Early-onset type 2 diabetes: A high-risk factor for proliferative diabetic retinopathy (PDR) in patients with microalbuminuria. *Medicine*. 2020;99(19)
21. Nanayakkara N, Curtis AJ, Heritier S, et al. Impact of age at type 2 diabetes mellitus diagnosis on mortality and vascular complications: systematic review and meta-analyses. *Diabetologia*. 2021;64(2):275–287.
22. Urbina EM, Kimball TR, McCoy CE, Khoury PR, Daniels SR, Dolan LM. Youth with obesity and obesity-related type 2 diabetes mellitus demonstrate abnormalities in carotid structure and function. *Circulation*. 2009;119(22):2913–2919.
23. Huo X, Gao L, Guo L, et al. Risk of non-fatal cardiovascular diseases in early-onset versus late-onset type 2 diabetes in China: a cross-sectional study. *The lancet Diabetes & endocrinology*. 2016;4(2):115–124.
24. Rhodes ET, Prosser LA, Hoerger T, Lieu T, Ludwig DS, Laffel LM. Estimated morbidity and mortality in adolescents and young adults diagnosed with type 2 diabetes mellitus. *Diabetic medicine*. 2012;29(4):453–463.
25. Huo L, Magliano DJ, Rancière F, et al. Impact of age at diagnosis and duration of type 2 diabetes on mortality in Australia 1997–2011. *Diabetologia*. 2018;61(5):1055–1063.
26. Echouffo-Tcheugui JB, Niiranen TJ, McCabe EL, et al. An early-onset subgroup of type 2 diabetes: a multigenerational, prospective analysis in the Framingham heart study. *Diabetes care*. 2020;43(12):3086–3093.
27. Liu J-J, Liu S, Gurung RL, et al. Risk of progressive chronic kidney disease in individuals with early-onset type 2 diabetes: a prospective cohort study. *Nephrology Dialysis Transplantation*. 2020;35(1):115–121.
28. Dabelea D, Stafford JM, Mayer-Davis EJ, et al. Association of type 1 diabetes vs type 2 diabetes diagnosed during childhood and adolescence with complications during teenage years and young adulthood. *Jama*. 2017;317(8):825–835.
29. Luk AO, Lau ES, So W-Y, et al. Prospective study on the incidences of cardiovascular-renal complications in Chinese patients with young-onset type 1 and type 2 diabetes. *Diabetes care*. 2014;37(1):149–157.
30. Svane J, Lynge TH, Pedersen-Bjergaard U, et al. Cause-specific mortality in children and young adults with diabetes mellitus: a Danish nationwide cohort study. *European Journal of Preventive Cardiology*. 2021;28(2):159–165.

31. Wilmot EG, Leggate M, Khan JN, et al. Type 2 diabetes mellitus and obesity in young adults: the extreme phenotype with early cardiovascular dysfunction. *Diabetic Medicine*. 2014;31(7):794–798.
32. Saydah SH, Siegel KR, Imperatore G, Mercado C, Gregg EW. The cardiometabolic risk profile of young adults with diabetes in the US. *Diabetes Care*. 2019;42(10):1895–1902.
33. Zheng L, Chen X, Luo T, et al. Peer Reviewed: Early-Onset Type 2 Diabetes as a Risk Factor for End-Stage Renal Disease in Patients With Diabetic Kidney Disease. *Preventing Chronic Disease*. 2020;17
34. Davidson AJ, Park AL, Berger H, et al. Association of improved periconception hemoglobin A1c with pregnancy outcomes in women with diabetes. *JAMA network open*. 2020;3(12):e2030207–e2030207.
35. Constantino MI, Molyneaux L, Limacher-Gisler F, et al. Long-term complications and mortality in young-onset diabetes: type 2 diabetes is more hazardous and lethal than type 1 diabetes. *Diabetes care*. 2013;36(12):3863–3869.
36. Gregg EW, Cheng YJ, Srinivasan M, et al. Trends in cause-specific mortality among adults with and without diagnosed diabetes in the USA: an epidemiological analysis of linked national survey and vital statistics data. *The Lancet*. 2018;391(10138):2430–2440.
37. Ke C, Shah BR, Luk AO, Di Ruggiero E, Chan JC. Cardiovascular outcomes trials in type 2 diabetes: time to include young adults. *Diabetes Obes Metab*. 2020;22(1):3–5.
38. Ke C, Shah BR, Stukel TA, Chan JC, Luk A. Excess burden of mental illness and hospitalization in young-onset type 2 diabetes. *Annals of internal medicine*. 2019;171(1):78–79.
39. Browne JL, Scibilia R, Speight J. The needs, concerns, and characteristics of younger Australian adults with type 2 diabetes. *Diabetic Medicine*. 2013;30(5):620–626.
40. Silverstein J, Cheng P, Ruedy KJ, et al. Depressive symptoms in youth with type 1 or type 2 diabetes: results of the pediatric diabetes consortium screening assessment of depression in diabetes study. *Diabetes Care*. 2015;38(12):2341–2343.
41. McVoy M, Hardin H, Fulchiero E, et al. Mental health comorbidity and youth onset type 2 diabetes: A systematic review of the literature. *The International Journal of Psychiatry in Medicine*. 2022:00912174211067335.
42. Tönnies T, Hoyer A, Brinks R. Productivity-adjusted life years lost due to type 2 diabetes in Germany in 2020 and 2040. *Diabetologia*. 2021;64(6):1288–1297.
43. National Diabetes Audit, 2015-2016 (2017).
44. Selvin E, Parrinello CM, Sacks DB, Coresh J. Trends in prevalence and control of diabetes in the United States, 1988–1994 and 1999–2010. *Annals of internal medicine*. 2014;160(8):517–525.
45. Gopalan A, Mishra P, Alexeeff SE, et al. Initial glycemic control and care among younger adults diagnosed with type 2 diabetes. *Diabetes Care*. 2020;43(5):975–981.
46. Magliano DJ, Martin VJ, Owen AJ, Zomer E, Liew D. The productivity burden of diabetes at a population level. *Diabetes care*. 2018;41(5):979–984.
47. Benhalima K, Song SH, Wilmot EG, et al. Characteristics, complications and management of a large multiethnic cohort of younger adults with type 2 diabetes. *Primary care diabetes*. 2011;5(4):245–250.
48. Ranchagoda JD. Prevalence and BMI of early-onset adult type 2 diabetes in a multiethnic population. *Diabetologia*. 2020;63((SUPPL 1)):S77–S78.
49. Goff LM. Ethnicity and type 2 diabetes in the UK. *Diabetic Medicine*. 2019;36(8):927–938.

50. Benhalima K, Wilmot E, Khunti K, Gray LJ, Lawrence I, Davies M. Type 2 diabetes in younger adults: clinical characteristics, diabetes-related complications and management of risk factors. *Primary care diabetes*. 2011;5(1):57–62.
51. UK D. State of the nation (England 2016): time to take control of diabetes. https://www.diabetes.org.uk/professionals/position-statements-reports/statistics/state-of-the-nation-2016-time-to-take-control-of-diabetes
52. Gregg EW, Karter AJ, Gerzoff RB, et al. Characteristics of insured patients with persistent gaps in diabetes care services: the Translating Research into Action for Diabetes (TRIAD) study. *Medical care*. 2010;48(1):31.
53. Bo A, Jensen NH, Bro F, Nicolaisen SK, Maindal HT. Higher patient assessed quality of chronic care is associated with lower diabetes distress among adults with early-onset type 2 diabetes: cross-sectional survey results from the Danish DD2-study. *Primary Care Diabetes*. 2020;14(5):522–528.
54. Gæde P, Oellgaard J, Carstensen B, et al. Years of life gained by multifactorial intervention in patients with type 2 diabetes mellitus and microalbuminuria: 21 years follow-up on the Steno-2 randomised trial. *Diabetologia*. 2016;59(11):2298–2307.
55. Gæde P, Vedel P, Larsen N, Jensen GV, Parving H-H, Pedersen O. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. *New England Journal of Medicine*. 2003;348(5):383–393.
56. Gæde P, Vedel P, Parving H-H, Pedersen O. Intensified multifactorial intervention in patients with type 2 diabetes mellitus and microalbuminuria: the Steno type 2 randomised study. *The Lancet*. 1999;353(9153):617–622.
57. Gæde P, Valentine WJ, Palmer AJ, et al. Cost-effectiveness of intensified versus conventional multifactorial intervention in type 2 diabetes: results and projections from the Steno-2 study. *Diabetes care*. 2008;31(8):1510–1515.
58. Oellgaard J, Gæde P, Rossing P, et al. Reduced risk of heart failure with intensified multifactorial intervention in individuals with type 2 diabetes and microalbuminuria: 21 years of follow-up in the randomised Steno-2 study. *Diabetologia*. 2018;61(8):1724–1733.
59. Taheri S, Zaghloul H, Chagoury O, et al. Effect of intensive lifestyle intervention on bodyweight and glycaemia in early type 2 diabetes (DIADEM-I): an open-label, parallel-group, randomised controlled trial. *The lancet Diabetes & endocrinology*. 2020;8(6):477–489.
60. May C, Montori VM, Mair FS. We need minimally disruptive medicine. *Bmj*. 2009;339
61. May CR, Eton DT, Boehmer K, et al. Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness. *BMC health services research*. 2014;14(1):1–11.
62. Gulsin GS, Brady EM, Swarbrick DJ, et al. Rationale, design and study protocol of the randomised controlled trial: Diabetes Interventional Assessment of Slimming or Training tO Lessen Inconspicuous Cardiovascular Dysfunction (the DIASTOLIC study). *BMJ open*. 2019;9(3):e023207.
63. Htike Z, Yates T, Brady E, et al. Rationale and design of the randomised controlled trial to assess the impact of liraglutide on cardiac function and structure in young adults with type 2 diabetes (the LYDIA study). *Cardiovascular diabetology*. 2016;15(1):1–9.
64. Chamnan P, Simmons RK, Khaw K, Wareham N, Griffin S. Change in HbA1c over 3 years does not improve the prediction of cardiovascular disease over and above HbA1c measured at a single time point. *Diabetologia*. 2013;56(5):1004–1011.

65. Battelino T, Danne T, Bergenstal RM, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. *Diabetes care*. 2019;42(8):1593–1603.
66. Xing D, Kollman C, Beck RW, et al. Optimal sampling intervals to assess long-term glycemic control using continuous glucose monitoring. *Diabetes technology & therapeutics*. 2011;13(3):351–358.
67. Riddlesworth TD, Beck RW, Gal RL, et al. Optimal sampling duration for continuous glucose monitoring to determine long-term glycemic control. *Diabetes technology & therapeutics*. 2018;20(4):314–316.
68. Duncan P, Murphy M, Man M-S, Chaplin K, Gaunt D, Salisbury C. Development and validation of the multimorbidity treatment burden questionnaire (MTBQ). *BMJ open*. 2020;8(4):e019413.
69. Neff KD. The development and validation of a scale to measure self-compassion. *Self and identity*. 2003;2(3):223–250.
70. Polonsky WH, Anderson BJ, Lohrer PA, et al. Assessment of diabetes-related distress. *Diabetes care*. 1995;18(6):754–760.
71. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine*. 2001;16(9):606–613.
72. Browne JL, Ventura AD, Mosely K, Speight J. Measuring the stigma surrounding type 2 diabetes: development and validation of the Type 2 Diabetes Stigma Assessment Scale (DSAS-2). *Diabetes Care*. 2016;39(12):2141–2148.
73. Sturt J, Hearnshaw H, Wakelin M. Validity and reliability of the DMSES UK: a measure of self-efficacy for type 2 diabetes self-management. *Primary Health Care Research & Development*. 2010;11(4):374–381.
74. Puthoff ML. Outcome measures in cardiopulmonary physical therapy: short physical performance battery. *Cardiopulmonary physical therapy journal*. 2008;19(1):17.
75. England CY, Thompson JL, Jago R, Cooper AR, Andrews RC. Development of a brief, reliable and valid diet assessment tool for impaired glucose tolerance and diabetes: the UK Diabetes and Diet Questionnaire. *Public health nutrition*. 2017;20(2):191–199.
76. Ho IS, Azcoaga-Lorenzo A, Akbari A, et al. Measuring multimorbidity in research: Delphi consensus study. *BMJ Medicine*. 2022;
77. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *sleep*. 1991;14(6):540–545.
78. Chung F, Abdullah HR, Liao P. STOP-Bang questionnaire: a practical approach to screen for obstructive sleep apnea. *Chest*. 2016;149(3):631–638.
79. Innes E. Handgrip strength testing: a review of the literature. *Australian Occupational Therapy Journal*. 1999;46(3):120–140.
80. Organisation WH. Measuring health and disability: Manual for WHO Disability Assessment Schedule
WHODAS 2.0. https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health/who-disability-assessment-schedule
81. Gillies K, Kearney A, Keenan C, et al. Strategies to improve retention in randomised trials. *Cochrane Database of Systematic Reviews*. 2021;(3)
82. Food U. Drug Administration: Guidance for industry: Patient-reported outcome measures—Use in medical product development to support labeling claims. *Food and Drug Administration, Center for Drug Evaluation and Research*. 2009;6

83. Rombach I, Rivero-Arias O, Gray AM, Jenkinson C, Burke O. The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. *Quality of Life Research*. 2016;25(7):1613–1623.
84. Meirte J, Hellemans N, Anthonissen M, et al. Benefits and disadvantages of electronic patient-reported outcome measures: systematic review. *JMIR perioperative medicine*. 2020;3(1):e15588.
85. Belisario JSM, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. *Cochrane database of systematic reviews*. 2015;(7)
86. Aiyegbusi OL. Key methodological considerations for usability testing of electronic patient-reported outcome (ePRO) systems. *Quality of life research*. 2020;29(2):325–333.
87. Sargeant JA, Brady EM, Zaccardi F, et al. Adults with early-onset type 2 diabetes (aged 18–39 years) are severely underrepresented in diabetes clinical research trials. *Diabetologia*. 2020;63(8):1516–1520.
88. Tudur Smith C, Hickey H, Clarke M, Blazeby J, Williamson P. The trials methodological research agenda: results from a priority setting exercise. *Trials*. 2014;15(1):1–7.
89. Priority. Prioritising Recruitment in Randomised Trials study. https://priorityresearch.ie/priority-two-questions/
90. Treweek S, Bevan S, Bower P, et al. Trial Forge Guidance 2: how to decide if a further Study Within A Trial (SWAT) is needed. *Trials*. 2020;21(1):1–9.
91. Research NIHfTM. SWAT repository store. https://www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/SWATSWARInformation/Repositories/SWATStore/
92. Karnoe A, Furstrand D, Christensen KB, Norgaard O, Kayser L. Assessing Competencies Needed to Engage With Digital Health Services: Development of the eHealth Literacy Assessment Toolkit. *J Med Internet Res*. May 10 2018;20(5):e178. doi:10.2196/jmir.8347
93. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value in Health*. 2009;12(4):419–429.
94. Bernstein DN, McIntyre AW, Baumhauer JF. Effect of assessment administration method and timing on patient-reported outcome measures completion and scores: Overview and recommendations. *Musculoskeletal Care*. 2020;18(4):535–540.