Statistical Analysis Plan

Version 2.0, February 2025

NCT06191848

Authors:

Dr Sharmala Thuraisingam
A/Prof Tim Spelman
Dr Cade Shadbolt
A/Prof Chris Schilling
Prof Michelle Dowsey

Table of Contents

List of Abbreviations

AE	Adverse event
AOANJRR	Australian Orthopaedic Association National Joint Replacement Registry
BMI	Body Mass Index
CI	Confidence interval
FDA	Food and Drug Administration
ICH	International Council for Harmonisation
IQR	Interquartile range
IPTW	Inverse Probability of Treatment Weighting
ITT	Intention to treat
MCS	Mental component score
NDI	National Death Index
OA	Osteoarthritis
PASE	Physical activity scale for the elderly
PCS	Physical component score
RR	Relative risk
SAE	Serious adverse event
SAP	Statistical analysis plan
SF	Short Form
SHR	Subdistribution hazard ratio
VDI	Victorian Death Index
WOMAC	Western Ontario and McMaster Universities Arthritis Index

1. Administrative Information

This statistical analysis plan is based on the study protocol version 2.0 dated 28th November 2024.

1.1 Study identifiers

ClinicalTrials.gov register NCT06191848.

1.2 Revision history

Version	Date	Changes made to document	Authors
1.0	06/12/2024	First draft	Sharmala Thuraisingam
2.0	28/02/2025	Revised following initial feedback	Sharmala Thuraisingam

1.3 Contributors to the statistical analysis plan

Names and ORCID	Affiliation	Role on study	SAP contribution
Dr Sharmala Thuraisingam 0000-0002-9450-1302	University of Melbourne, Department of Surgery, St Vincent's Hospital	Statistician	Statistical concepts and writing of plan
A/Prof Tim Spelman 0000-0001-9204-3216	University of Melbourne, Department of Surgery, St Vincent's Hospital	Senior Statistician	Statistical concepts and writing of plan
Dr Cade Shadbolt 0000-0002-0937-2412	University of Melbourne, Department of Surgery, St Vincent's Hospital	Data Scientist	Statistical concepts and contributions to plan
A/Prof Chris Schilling 0000-0002-1747-7249	University of Melbourne, Department of Surgery, St Vincent's Hospital	Health Economist	Health economic concepts and writing of plan
Prof Michelle Dowsey 0000-0002-9708-5308	University of Melbourne, Department of Surgery, St Vincent's Hospital	Principal Investigator	Statistical concepts and contributions to plan

1.4 Approvals

All authors have reviewed this plan and approve it as final. They find it to be consistent with the requirements of the protocol as it applies to their respective areas. They also find it to be compliant with ICH-E9 principles and, in particular, confirm that this analysis plan was developed in a completely blinded manner (i.e. without knowledge of the effect of the intervention(s) being assessed).

2. Introduction

This statistical analysis plan (SAP) details the intended analyses to be conducted on data collected in the STOP KNEE-OA trial. The SAP includes the primary and secondary analyses, health economic assessments, as well as the proposed layout

of tables and figures for the main outcome paper. All analyses detailed in this plan have been pre-specified prior to the completion of data collection and comply with the FDA's Guidance for Industry ICH E9 Statistical Principles for Clinical Trials.

3. Study synopsis

STOP KNEE-OA is a multicentre, double-blinded, parallel group, randomised controlled trial designed to determine if weekly subcutaneous tirzepatide is superior to placebo in limiting the progression to knee replacement in adults with moderate-to-severe knee osteoarthritis who were eligible for unilateral knee replacement, with a body-mass index (BMI) of at least 30 kg/m^2, without diabetes. Participants will be randomised in a 1:1 ratio to the tirzepatide and placebo study arms, and all participants will receive standardised dietary and physical activity advice. Follow-up of participants will occur every 4 weeks for 72 weeks through on-site and telephone visits. Long term follow-up at 5 and 10 years will be achieved through data linkage with the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) and the National Death Index (NDI)/Victorian Death Index (VDI) (1,2).

4. Study objectives

4.1 Primary objective

The primary efficacy aim is to determine whether weekly tirzepatide added to standardised lifestyle advice compared with standardised lifestyle advice only (placebo), reduces the proportion of participants that progress to knee replacement surgery within 72 weeks. The primary null hypothesis of no difference in the proportion of participants that progress to knee replacement in participants assigned to tirzepatide versus placebo within 72 weeks will be tested.

4.2 Secondary objective

The secondary efficacy aims are to evaluate whether weekly tirzepatide is superior to placebo in: (1) improving osteoarthritis symptoms, (2) reducing patient's willingness to undergo knee replacement, (3) reducing bodyweight, (4) improving physical and mental health status, (5) improving physical activity levels, (6) reducing the use of prescription pain medication at 72 weeks and (7) reducing the proportion of patients that progress to knee replacement within 5 and 10 years. In each case, the null hypothesis of no difference between study arms will be tested.

4.3 Safety objective

The safety aims are to determine whether weekly tirzepatide are associated with adverse events, serious adverse events and adverse events that require treatment discontinuation.

5. Patient population

5.1 Inclusion criteria

Participants will be eligible to participate in the trial if they provide informed consent and meet all the following criteria:

- Age ≥ 18 years
- BMI ≥30 kg/m^2

- Been unsuccessful in losing body weight through lifestyle modification.
- Deemed eligible by an orthopaedic surgeon to be on the waiting list at one of the participating sites for primary knee replacement surgery in the target joint due to osteoarthritis.
- Moderate-to-severe osteoarthritis in the target joint as defined by Kellgren-Lawrence grade ≥2.
- Willing to and capable of self-injecting tirzepatide and following study procedures for the duration of the trial.
- Female participants must not be pregnant, breastfeeding nor of reproductive potential (further details in study protocol)

5.2 Exclusion criteria

Participants will be ineligible to participate in the trial if they meet any of the following criteria:

- Deemed eligible by an orthopaedic surgeon to be on the waiting list at one of the participating sites for knee replacement surgery in the contralateral knee.
- Used prescription medication for weight loss in the three months prior to screening.
- Undergone surgical or endoscopic procedures for weight loss in the past.
- Diagnosed with type 1 or 2 diabetes.
- Have laboratory evidence of diabetes mellitus during screening.
- Have a personal or family history of medullary thyroid carcinoma (MTC) or multiple endocrine neoplasia syndrome type 2 (MEN 2).
- Have an active malignancy excluding basal or squamous cell skin cancer.
- Awaiting or have a transplanted organ.
- Have received chronic systemic glucocorticoid therapy (for more than 14 days) in prior 3 months or have a significant, active autoimmune abnormality (e.g., lupus or rheumatoid arthritis) that the study doctor deems likely to require systemic glucocorticoid therapy during the next 18 months.
- Have any other medical condition, abnormal laboratory tests or medications as outlined in the protocol section 5.3.2 that makes them unsuitable to participate in the trial.
- Are part of the study site team/personnel or family member of the study site team/personnel.
- Currently enrolled in any other study of an investigational product or previously enrolled in another study within 90 days of study visit 1.

6. Study outcomes

6.1 Primary outcome

The primary outcome is the proportion of patients that progress to knee replacement surgery at 72 weeks. The effect estimate is the relative risk of knee replacement surgery at 72 weeks (tirzepatide vs placebo).

6.2 Secondary outcomes

Secondary outcomes for the main analysis and health economic assessment are summarised in Table 1.

6.3 Safety outcomes

Safety outcomes have been defined in the study protocol and include adverse events (AE), serious adverse events (SAE) and adverse events requiring treatment discontinuation that occur any time throughout the 72-week trial period.

Table 1. Secondary outcome measures

Main Analysis	Variable type	Effect estimates	Assessment time points	Analysis set	Statistical model
Change in WOMAC pain	Continuous	Mean difference in change in WOMAC pain	72 weeks	FAS	Linear mixed model
Change in WOMAC function	Continuous	Mean difference in change in WOMAC function	72 weeks	FAS	Linear mixed model
Change in WOMAC stiffness	Continuous	Mean difference in change in WOMAC stiffness	72 weeks	FAS	Linear mixed model
Willingness for surgery*	Binary	Relative risk	72 weeks	FAS	Binomial regression
Percentage change in body weight	Continuous	Mean difference in percentage change in body weight	72 weeks	FAS	Linear mixed model
≥5% reduction in body weight	Binary	Relative risk	72 weeks	FAS	Binomial regression
≥10% reduction in body weight	Binary	Relative risk	72 weeks	FAS	Binomial regression
≥20% reduction in bodyweight	Binary	Relative risk	72 weeks	FAS	Binomial regression
Change in SF-36 MCS	Continuous	Mean difference in change in SF-36 MCS	72 weeks	FAS	Linear mixed model
Change in SF-36 PCS	Continuous	Mean difference in change in SF-36 PCS	72 weeks	FAS	Linear mixed model
Change in PASE	Continuous	Mean difference in change in PASE	72 weeks	FAS	Linear mixed model
Use of non-opioid prescription analgesics	Binary	Relative risk	68-72 weeks	FAS	Binomial regression
Use of opioid prescription analgesics	Binary	Relative risk	68-72 weeks	FAS	Binomial regression
Opioid prescription analgesic dose	Continuous	Mean difference in oral morphine equivalent (OME) per day (3)	72 weeks	FAS	Linear mixed model
Knee replacement surgery (long term)	Time to event with death as	Subdistribution hazard ratio	5 and 10 years	FAS	Fine & Gray competing risk model

	competing risk				
Health Economic Analysis					
Change in health-related quality of life using SF-36	Continuous	Mean difference in Quality Adjusted Life Years (QALY)	3 monthly until 72 weeks, 5 and 10 years	FAS	Linear and generalized linear models
Cost-effectiveness	Ratio	Incremental cost-effectiveness ratio (cost per QALY)	72 weeks, 5 and 10 years	FAS	Linear and generalized linear models

*Patients will be considered willing to undergo surgery at 72 weeks if they had a knee replacement by 72 weeks or remained on the surgery waiting list.
Abbreviations: WOMAC=Western Ontario and McMaster Universities Osteoarthritis Index; MCS=mental component score; PCS=Physical Component Score; PASE=Physical Activity Scale for the Elderly; SF-36=36-Item Short Form Health Survey; QALY=Quality-adjusted life-year; FAS=Full Analysis Set.

7. Intervention

Participants will be randomised to receive either tirzepatide or placebo and self-administer the pre-filled syringes subcutaneously into the thigh or abdomen on a weekly basis for 72 weeks. Administration of the drug/placebo will occur on the same day, at approximately the same time, each week. Intervention participants will initially receive 2.5mg of tirzepatide which will be increased by 2.5mg every four weeks over the 20-week dose escalation period until the target weekly dose of 15mg is reached. Participants who are unable to reach the target weekly dose will remain at the highest tolerable dose of either 5mg or 10mg. Dose modifications will only be permitted for participants experiencing gastrointestinal symptoms and will be addressed as outlined in the protocol (section 5.5.3 Dose escalation and modification). Participants that require their dose of tirzepatide reduced below 5mg will discontinue the study drug.

All participants will receive dietary and physical activity counselling throughout the study, including advice on making healthy food choices and achieving calculated calorie deficits.

8. Randomisation and blinding

Randomisation of participants will occur at the second study visit, after participant screening and consent have occurred. The study's Interactive Web Response Systems (IWRS) will randomly allocate participants to either the tirzepatide or placebo group in a 1:1 ratio using randomly permuted block sizes of 2 and 4. Randomisation will be stratified by site, gender (male or female) and body mass index (BMI) category (30-39.9 kg/m^2 or 40+ kg/m^2).

All investigators, site personnel and statisticians will be blinded to participant study group allocations. Only the study pharmacist at the central pharmacy will be unblinded to ensure the correct medication and dosages are dispensed to participants. Individual site pharmacists will remain blinded.

9. Sample size

To detect a 10% reduction in progression to knee replacement surgery at 72 weeks from 95% in the placebo group to 85% in the tirzepatide group (with 80% power, two-tailed α=0.05), and 10% attrition rate, a sample size of 352 participants (176 per group) is required. The sample size is based on the ABS Study (4), a randomized controlled trial that demonstrated a 24.4% (95% CI 9.0% to 39.8%) reduction in progression to knee replacement surgery in surgical weight loss (gastric banding) participants compared with participants who received lifestyle modification advice. The effect size for the sample size calculation is based on the lower limit of the 95% confidence interval for the effect size observed in the ABS study. The estimated attrition rate of 10% and assumed 95% progression to knee replacement in the placebo arm is based on the ABS study and other trials investigating the progression to orthopaedic surgery ((4–6). The sample size calculation was performed in Stata version 18 (StataCorp. 2023. Stata Statistical Software: Release 18. College Station, TX: StataCorp LLC.) using the "sampsi" command, which allows for continuity correction (7).

10. Statistical analysis

10.1 General principles

The following sections provide further detail on the primary and secondary analyses stated in the study protocol.

All data cleaning and statistical analyses will be conducted using Stata version 18 (StataCorp. 2023. Stata Statistical Software: Release 18. College Station, TX: StataCorp LLC.)(7). Statistical tests will be two-tailed and a p-value<0.05 used to determine statistical significance. Continuous data will be summarised using means, standard deviations (SD) and range, and medians and interquartile ranges (IQR) for skewed data. Categorical data will be summarised using counts and percentages. Transformations will be tested when modelling skewed continuous outcomes. Results from analyses will be accompanied by 95% confidence intervals and p-values.

10.2 Interim analyses

No interim analyses for efficacy, futility and/or sample size revision are pre-planned for this trial. The Data Monitoring Committee may request and conduct unplanned interim analyses if required.

10.3 Multiplicity adjustment

No formal adjustments for multiplicity are planned given there is only one primary outcome. However, the multiple comparisons, strength of evidence and consistency of effects across the outcomes will be considered when interpreting the secondary analyses results.

10.4 Data sets to be analysed

All primary and secondary analyses will be performed using the full analysis set. This data set will include all data obtained from all participants randomised in the trial regardless of whether they adhered to or discontinued the study drug. Outcome data for knee replacement at 72 weeks will be obtained through data linkage with the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) (1). To facilitate the analysis of long-term outcomes, knee replacement data (from AOANJRR) will be linked at 5 and 10 years, along with dates of death from the National Death Index (NDI)/Victorian Death Index (VDI) (2). The safety analysis data set will include all participants randomised in the trial and any adverse event data.

10.5 Participant disposition

A CONSORT flowchart (Appendix 1, Figure 1) will be prepared to summarise the number of participants who were screened for eligibility, excluded, randomised and withdrew. Reasons for exclusion and withdrawal will be listed.

10.6 Patient characteristics and baseline comparisons

Participant characteristics at baseline will be summarised by trial arm and presented as outlined in Appendix 1 Table 1. Descriptive statistics will be used to compare baseline characteristics between the study arms.

10.7 Analysis of the primary outcome

10.7.1 Main analysis

The "treatment-regimen" estimand (8) will compare the proportion of participants in each of the assigned treatment groups that progress to knee replacement within 72-weeks regardless of adherence to study drug. Binomial regression will be used to estimate the relative risk of knee replacement at 72 weeks in the intervention arm compared with the placebo arm. Results will be presented as shown in Appendix 1 Table 4. Point estimate relative risks and risk differences will be presented with associated 95% confidence intervals. A graph depicting the cumulative incidence of knee replacement with time will be presented (Appendix 1, Figure 5).

10.7.2 Sensitivity analyses

Three sensitivity analyses will be conducted on the primary outcome:

(1) An adjusted analysis which extends the main analysis model to include stratification variables. Two different models will be fit. The first will include site, gender and BMI as fixed effects, the second model will include these stratification variables as random effects.

(2) Time to knee replacement will be analysed using a Cox proportional hazards model with treatment included as a covariate. The treatment effect will be estimated by the hazard ratio (HR) and 95% confidence interval.

(3) The "trial product" estimand (hypothetical strategy) (8) provides an estimate of the treatment effect had participants remained on their treatment for the planned 72-week treatment duration. The estimand is based on a hypothetical scenario in which intercurrent events do not occur (8). In this study, intercurrent events are defined as discontinuation of treatment due to adverse events or for other reasons outlined in the protocol (8). Either mixed models or Inverse Probability of Treatment Weighting (IPTW) with or without multiple imputation may be used for the "trial product" estimand (9,10).

Treatment effect estimates obtained from the sensitivity analyses will be included in a forest plot (Appendix 1, Figure 6) alongside the "treatment-regimen" estimand.

10.7.3 Subgroup analyses

The primary analysis as defined in section 10.7.1, will be conducted according to the following subgroups defined by baseline characteristics:

1. Sex (male, female)
2. Body Mass Index (30 to 40, ≥40 kg/m^2)
3. Kellgren-Lawrence Grade (2, 3 or 4)

For each subgroup analysis, the subgroup variable and an interaction between subgroup and treatment will be included as fixed effects in the model. The significance of the interaction term (P<0.05) will be used to assess whether treatment effects differ across the levels of the subgroup. Treatment effect estimates from the subgroup analyses will be included in a forest plot with estimates from the main and sensitivity analyses (Appendix 1, Figure 6).

10.8 Analysis of secondary outcomes

The "treatment-regimen" estimand (8) will be estimated for all secondary outcomes. Where mixed models are used, random intercepts will be included to account for clustering of the outcome within participants and an unstructured variance-covariance structure assumed. Treatment effect estimates will be adjusted by stratification factors. Appendix 1, Tables 5-7 show the intended table layouts for the results, and Figures 7-13 the planned plots.

10.8.1 Continuous secondary outcomes

For the continuous secondary outcomes listed in Table 1, linear mixed models using restricted maximum likelihood estimation will be used to estimate the difference in mean outcome between the arms at 72 weeks. The outcome at 3-monthly intervals will be included as dependent variables in the model. The outcome at baseline, time and an interaction between time and treatment group will be included as fixed effects in the model.

10.8.2 Binary secondary outcomes

Binomial regression models will be used to estimate the relative risk of willingness for surgery, reduction in body weight (≥5%, ≥10% and ≥20%) and use of opioid and non-opioid prescription analgesics (Table 1) in the tirzepatide arm compared with the placebo arm at 72 weeks.

10.8.3 Time to event secondary outcomes

Fine & Gray competing risk regression models will be used to estimate the rate of knee replacement (subdistribution hazard rate) in the tirzepatide arm compared with the placebo arm at 5 and 10 years. Death will be treated as a competing risk given the age of the study population.

10.9 Treatment of missing data

It is anticipated that there will not be any missing outcome data for the primary outcome given the occurrence of knee replacement at 72 weeks will be determined through in person study visits, review of hospital records and linkage with the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) which has near complete capture of all knee replacements in Australia (1). However, should data linkage be delayed, and/or there be less than 40% missing outcome data, multiple imputation may be used, subject to approval by the Data Safety Monitoring Board. Imputation models will include the outcome, stratification factors and any auxiliary variables available in the data set that may explain the missing outcome data. The validity of the imputed data will be assessed via graphical comparisons of imputed and observed data, including kernel density plots and histograms, and comparisons of goodness-of-fit between imputed and non-imputed datasets. The same approach will be used for secondary outcomes where there are less than 40% missing data.

10.10 Health economic assessment

A cost-utility analysis will be completed to assess the cost-effectiveness of treatment relative to the status quo standardised lifestyle advice. First, the trial results will be considered without any further extrapolation. Outcomes will be Quality Adjusted Life Years (QALYs), calculated via linear interpolation of health-related quality of life measures from

the SF-36 at baseline and each follow-up measurement. A health system perspective will be adopted, with costs including the intervention and follow-up treatment, in particular joint replacement. A cost-effectiveness plane and cost-acceptability curve will be presented. A subsequent step may also be included, extrapolating the results of the longer-term impact of the trial results for a further 5 and 10 years using the Australian Osteoarthritis (AUS-OA) validated health economic model (12), using a discount rate of 3.5%. A probabilistic sensitivity analysis would be incorporated to assess the robustness of the model-based results to changes in the input parameters and assumptions. Results will be presented as shown in Appendix 1, Table 8 and Figures 14-15.

10.11 Analysis of safety outcomes

For each event type (AE and SAE), the number and proportion of participants experiencing the event, median number of events and rate of events will be reported by study group. The proportion of participants that experience a fatal event or an adverse event resulting in treatment discontinuation will also be reported. Adverse events occurring in over 5% of participants in any of the trial arms will be summarised. Results will be presented as shown in Appendix 1, Table 9.

11. References

1. South Australian Health and Medical Research Institute. Home - AOANJRR [Internet]. 2019 [cited 2020 Apr 24]. Available from: https://aoanjrr.sahmri.com/

2. Australian Institute of Health and Welfare. National Death Index (NDI) - Australian Institute of Health and Welfare [Internet]. 2020 [cited 2020 Apr 24]. Available from: https://www.aihw.gov.au/about-our-data/our-data-collections/national-death-index

3. Nielsen S, Degenhardt L, Hoban B, Gisev N. A synthesis of oral morphine equivalents (OME) for opioid utilisation studies. Pharmacoepidemiol Drug Saf [Internet]. 2016 Jun 1 [cited 2024 Dec 3];25(6):733–7. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/pds.3945

4. Dowsey MM, Brown WA, Cochrane A, Burton PR, Liew D, Choong PF. Effect of Bariatric Surgery on Risk of Complications After Total Knee Arthroplasty: A Randomized Clinical Trial. JAMA Netw Open [Internet]. 2022 Apr 1 [cited 2024 Jul 10];5(4):e226722–e226722. Available from: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2791035

5. Dowsey MM, Gould DJ, Spelman T, Pandy MG, Choong PF. A Randomized Controlled Trial Comparing a Medial Stabilized Total Knee Prosthesis to a Cruciate Retaining and Posterior Stabilized Design: A Report of the Clinical and Functional Outcomes Following Total Knee Replacement. J Arthroplasty [Internet]. 2020 Jun 1 [cited 2024 Jul 10];35(6):1583-1590.e2. Available from: https://pubmed.ncbi.nlm.nih.gov/32139194/

6. Dowsey M, Castle D, Knowles S, Monshat K, Salzberg M, Nelson E, et al. The effect of mindfulness training prior to total joint arthroplasty on post-operative pain and physical function: A randomised controlled trial. Complement Ther Med [Internet]. 2019 Oct 1 [cited 2024 Jul 10];46:195–201. Available from: https://pubmed.ncbi.nlm.nih.gov/31519279/

7. StataCorp LLC. Stata | Data Analysis and Statistical Software [Internet]. 2018 [cited 2018 Mar 12]. Available from: https://www.stata.com/products/

8. European Medicines Agency - Committee for Medicinal Products for Human Use. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2020 [cited 2024 Sep 23]; Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf

9. Jastreboff AM, Aronne LJ, Ahmad NN, Wharton S, Connery L, Alves B, et al. Tirzepatide Once Weekly for the Treatment of Obesity. New England Journal of Medicine [Internet]. 2022 Jul 21 [cited 2024 Dec 3];387(3):205–16. Available from: https://www.nejm.org/doi/full/10.1056/NEJMoa2206038

10. Parra CO, Daniel RM, Wright D, Bartlett JW. Estimating hypothetical estimands with causal inference and missing data estimators in a diabetes trial. 2023 Aug 24 [cited 2024 Dec 6]; Available from: https://arxiv.org/abs/2308.13085v2

11. Bond S, Allison A. rpsftm: rank-preserving structural failure time models for survival data. 2024.

12. Schilling C, Zhou Y, Rele S, Shadbolt C, Thuraisingam S, O'Brien P, et al. The development of AUS-OA - a population health economic model of osteoarthritis in Australia. Int J Microsimul. 2024;

Appendix 1: Proposed tables and figures

Figure 1: Consort flowchart

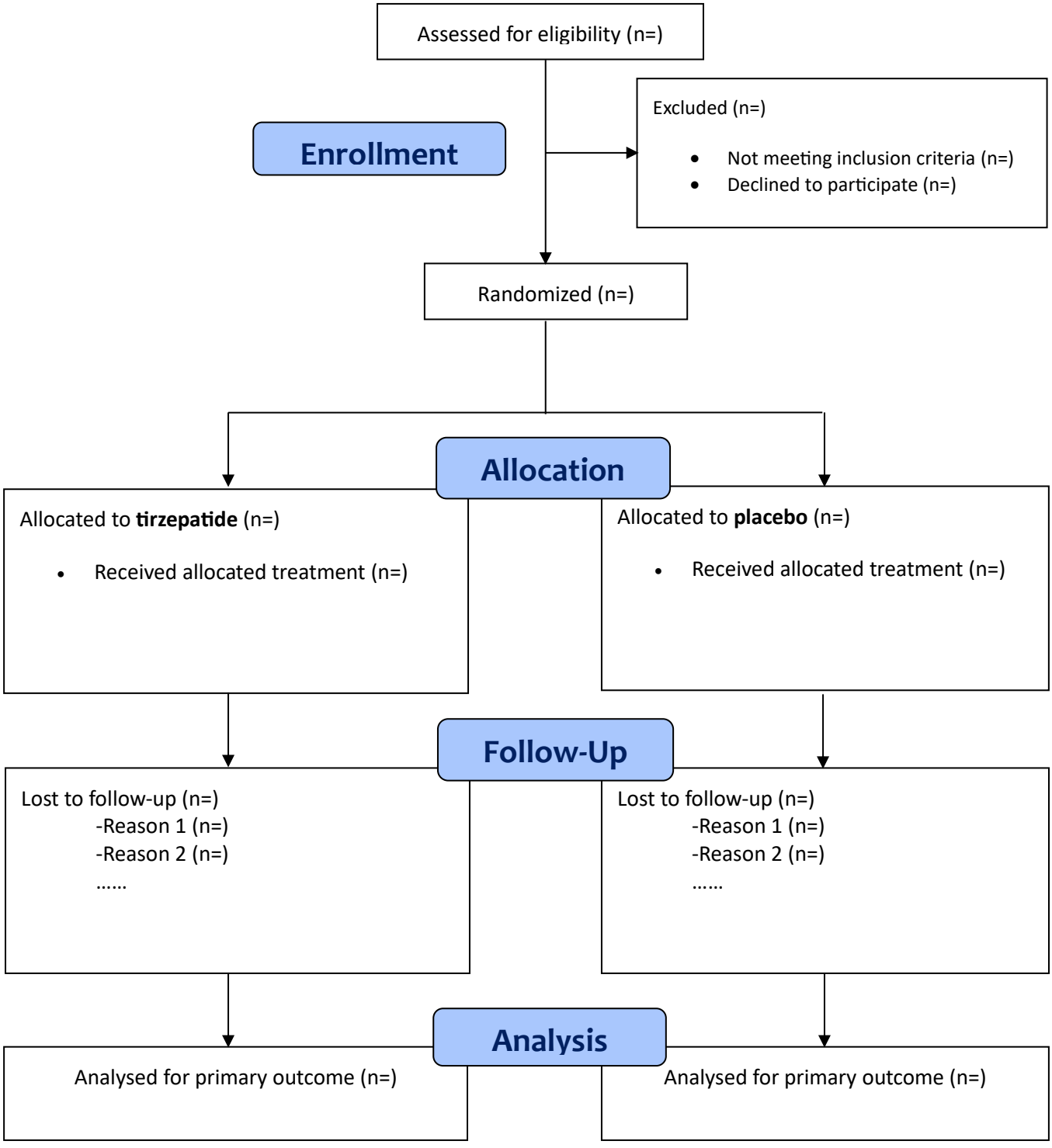

Figure 2a: Data linkage reporting flowchart at 72 weeks

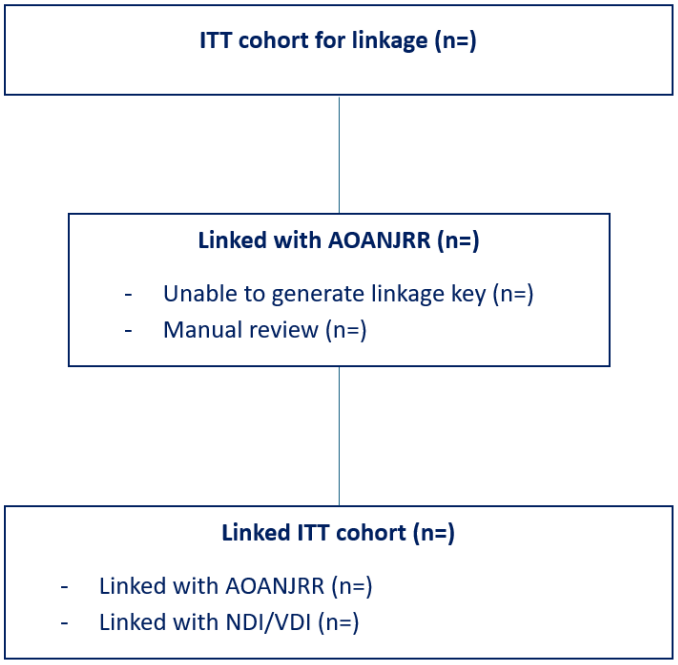

Figure 2b: Data linkage reporting flowchart at 5 and 10 years

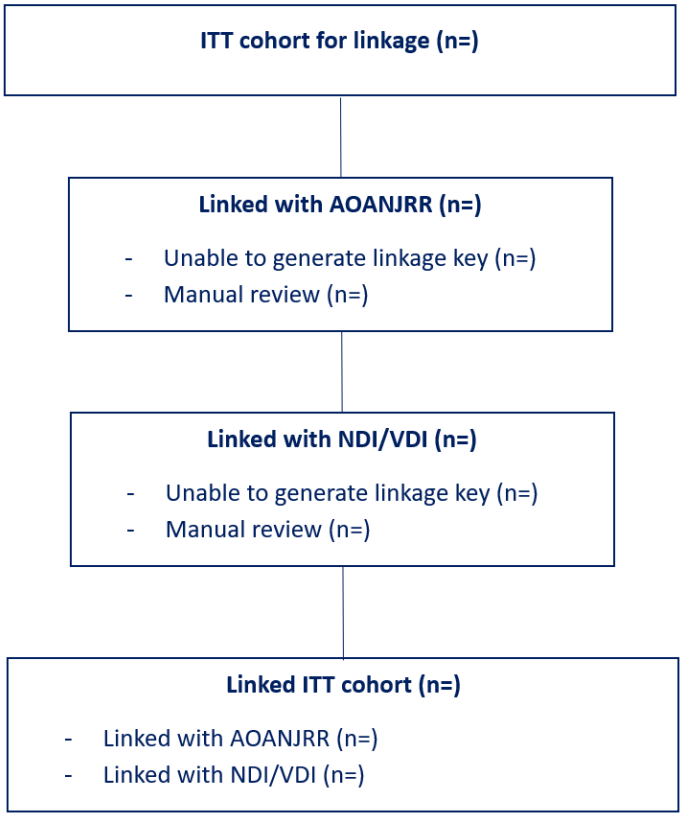

Table 1: Baseline characteristics

	Total (N=)	**Tirzepatide (N=)**	**Placebo (N=)**
	n (%)	**n (%)**	**n (%)**
Demographics			
Age, *median [IQR]*			
Female			
etc.			
Clinical measures			
BMI (kg/m^2)			
30-39.9			
≥40			
etc.			
Sites			
etc.			

Figure 3. Dosing of tirzepatide and placebo over trial period

Table 2: Reasons for discontinuing allocated treatment by study group

Figure 4. Cumulative proportion of patients who discontinued treatment by study group

Table 3: Concomitant therapies by study group

Table 4: Analysis of primary outcome (treatment-regimen estimand)

	Tirzepatide (N=)	**Placebo (N=)**	**Unadjusted**			
	n (%)	**n (%)**	**RR (95% CI)**	**P-value**	**RD (95% CI)**	**P-value**
Knee replacement at 72 weeks						

Abbreviations: RR= relative risk, RD=risk difference, CI=confidence interval

Figure 5. Cumulative incidence (%) of knee replacement with time

Figure 6. Forest plot of treatment effect estimates at 72 weeks from primary, sensitivity and subgroup analyses

Table 5: Analysis of continuous secondary outcomes (treatment-regimen estimand)

	Tirzepatide (N=)		**Placebo (N=)**		**Adjusted**	
	N	**Estimated mean (95% CI)**	**N**	**Estimated mean (95% CI)**	**Tirzepatide vs placebo between-group differences (95% CI)**	**P-value**
Change in WOMAC pain at 72 weeks						

Change in WOMAC function at 72 weeks						
Change in WOMAC stiffness at 72 weeks						
Percentage change in body weight at 72 weeks						
Change in SF-36 MCS at 72 weeks						
Change in SF-36 PCS at 72 weeks						
Change in PASE at 72 weeks						
OME per day at 68-72 weeks						

Note: All change estimates are from baseline.
Abbreviations: CI=confidence interval, WOMAC=Western Ontario and McMaster Universities Arthritis Index, SF-36=36-Item short form survey, PCS=physical component summary, MCS=mental component summary, PASE=physical activity scale for the elderly, OME=oral morphine equivalent.

Figures 7-13: Graphs for longitudinal secondary outcomes (WOMAC pain, WOMAC function, WOMAC stiffness, body weight, SF-36 MCS, SF-36 PCS, PASE) (a) effect estimates at 3-monthly time intervals and (b) cumulative frequency of change from baseline.

Table 6: Analysis of binary secondary outcomes

	Tirzepatide (N=)	**Placebo (N=)**	**Adjusted**			
	n (%)	**n (%)**	**RR (95% CI)**	**P-value**	**RD (95% CI)**	**P-value**
≥5% reduction in body weight at 72 weeks						
≥10% reduction in body weight at 72 weeks						
≥20% reduction in body weight at 72 weeks						
Willingness for surgery at 72 weeks						
Use of non-opioid prescription analgesics at 68-72 weeks						
Use of opioid prescription analgesics at 68-72 weeks						

Abbreviations: RR=relative risk, RD=risk difference, CI=confidence interval

Table 7. Analysis of secondary time to event outcomes

	Tirzepatide (N=)	Placebo (N=)	Adjusted	
Knee replacement	**n (%)**	**n (%)**	**SHR (95% CI)**	**P-value**
5 years				
10 years				

Abbreviations: SHR=subdistribution hazard ratio, CI=confidence interval

Table 8: Health economics analysis

	Tirzepatide (N=)	Placebo (N=)	Difference
QALYs			
Costs			
Incremental cost effectiveness ratio (ICER)			

Abbreviations: QALYs = Quality Adjusted Life Years

Figure 14: Cost-effectiveness plane

Figure 15: Cost acceptability curve

Table 9: **Adverse drug events**

	Total (N=)	Tirzepatide (N=)	Placebo (N=)
Any adverse event			
At least one adverse event, *n(%)*			
Median number of adverse events, *median [IQR]*			
Rate of adverse events, *(events per person)*			
Event type			
Adverse event (AE)			
At least one, *n(%)*			
Median number, *median [IQR]*			
Rate, *(events per person)*			
Serious adverse event (SAE)			
At least, *n(%)*			
Median number, *median [IQR]*			
Rate, *(events per person)*			
Adverse event requiring treatment discontinuation, n(%)			
Fatal event, n(%)			
Specific adverse events reported in more than 5% of participants in any group, n(%)			

Abbreviations: AE=adverse event, SAE=serious adverse event