Statistical Analysis Plan

Title:

Exercise dosage and the need for glucose-lowering medication in patients with type 2 diabetes: A secondary analysis of the U-TURN randomized clinical trial

Trial registration: NCT02417012

Version 1.0 (July 31st, 2018)

Principal investigator(s):

[1,2]Christopher Scott MacDonald, MSc, PhD student

[1]Mette Yun Johansen, MSc, PhD Student

Senior investigator:

[1]Mathias Ried-Larsen

Statistical analyst:

[2]Sabrina M. Nielsen, MSc; Biostatistics fellow.

Statistical advisor:

[2]Robin Christensen, MSc, PhD; Professor of Biostatistics and Clinical Epidemiology; Head of Musculoskeletal Statistics Unit.

Affiliations:

[1]Centre of Inflammation and Metabolism, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark & Centre for Physical Activity Research, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.
[2]Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Denmark

Background and rationale

Patients with type 2 diabetes (T2D) experience harmful side-effects from their glucose-lowering medication and/or from poorly regulated blood glucose (1-3). Considering the rapid increase in the number of patients with T2D, and given the limited literature indicating the beneficial effect of exercise on diabetic complications, there is a need to more thoroughly explore exercise as a management strategy or maybe even as a recommendable treatment (4).

The dose-response relationship between exercise and diabetes outcome measures is not well-established in patients with T2D. Moreover, the existing data is conflicting. Balducci et al. (5) found that a 12-month aerobic and resistance exercise intervention, corresponding to a volume of 150 min a week (frequency 2x week), led to a reduction in glycated hemoglobin A1c (HbA1c) of 0.3% (95% CI -0.49 to -0.10%) from a baseline value of 7.12% (1.4). However, no dose-response relationship was found. In addition, Vancea et al. did not find any difference in HbA1c reduction between a group of patients with T2D exercising (walking) 5 times per week (150 min/week; baseline: 7.7% [SD 1.8] to 20th week: 7.4% [SD 0.7]) compared to a group exercising 3 times per week (90 min/week; baseline: 8.2% [SD 1.9] to 20th week: 7.4% [SD 1.2]) (6). In the DARE trial, the effect of either aerobic, resistance or a combination of the two was explored in relation to HbA1c. A dose-response relationship was not specifically explored, however the combined group was completing a significantly high volume of total exercise. The aerobic-group reduced HbA1c 0.51 percentage points (95% CI -0.87 to -0.14), while the resistance-group reduced 0.38 percentage points (95% CI – 0.72 to -0.22). The greatest reduction was found in the combined-group with a reduction at approximately 1% (7).

In a post hoc analysis, Di Loreto et al. found an inverse relationship between energy expenditure (voluntary exercise) and body weight, blood pressure, waist circumference, and also HbA1c (8). Di Loreto also reported that an energy expenditure >10 MET/h/week was required to induce benefits in relation to HbA1c. This is in line with a meta-analysis (47 studies) by Umpierre et al. (2011) showing that structured exercise lasting more than 150 min/week was associated with a HbA1c reduction of 0.89 percentage points (95% CI −1.26% to −0.51%), whereas a volume of exercise

below 150 min/week was associated with a HbA1c reduction of 0.36 percentage points (95% CI −0.50% to −0.23%) (9).

In a meta-regression analysis of 26 RCTs, Umpierre et al. (2013) found that frequency of exercise explained nearly 32% of between studies variance, whereas volume explained 15%. The intensity of exercise did not show any statistical significant association with HbA1c changes (10). In contrast, Boulé et al. found that relative exercise intensity was inversely associated with the mean difference in HbA1c (p=0.002) (11). Few studies report the concomitant effect on glucose-lowering medication, making an interpretation of the roles of exercise in relation to HbA1c difficult. In addition, Metformin (first line treatment of T2D) has been shown to blunt the training-induced improvements in insulin sensitivity (12). This finding emphasizes the importance of considering concomitant drug-treatment when interpreting the effect of an exercise intervention on surrogate outcomes.

The U-turn study

The U-TURN study was a 12-months randomized assessor-blinded two-arm parallel group equivalence trial, designed to test whether an intensive lifestyle intervention was equally effective in maintaining glycemic control compared to standard clinical care in patients with T2D diagnosed < 10 years (13). The study employed a pre-baseline medical titration period in order to limit the potential for overestimating the positive effects of lifestyle on HbA1c. Furthermore, an endocrinologist, blinded to allocation, conducted all medical regulation, during the study, according to a predefined algorithm. The exercise prescription varied in the U-TURN study from baseline to 12-month follow-up. In phase 1 (0-16 weeks), six sessions a week (30-60 min) of aerobic exercise at 60% HRmax were prescribed, hereof two combined session consisting of approximately 30 minutes of aerobic and 30 minutes of resistance exercise. In phase 2 (16-33 weeks) the duration of aerobic sessions increased to 45 to 60 min and the intensity had to be 70% HRmax. In phase 3 (33-52 weeks), two sessions a week (60 min) of aerobic exercise in addition to three combined sessions (aerobic and resistance) were prescribed. The volume of aerobic exercise was approximately double the volume of the current recommendation for patients with T2D, which is 150 min/wk (14). The 12-month follow-up was finalized in September 2016. Based on the adherence results

of the primary outcome article (15), it is reasonable to assume that a lower-than-prescribed dose would be efficient in terms of discontinuation of glucose-lowering medication and secondarily on HbA1c.

As a sub-study, we here present a statistical analysis plan of a secondary analysis where the U-Turn intervention group will be divided into tertiles based on the cumulative volume of aerobic and strength exercise completed from baseline to 12-month follow-up. As all exercise sessions were monitored using Polar V800 heart rate monitors, objective assessment of each exercise session can be made.

Study Aims, Hypothesis and Outcomes

The primary aim of the study is to investigate if a dose-response relationship exists between the exercise adherence and the probability of being discontinued from glucose lowering medications. A key secondary aim is to investigate whether a dose-response relationship also applies to changes in HbA1c.

Hypothesis

We hypothesize a positive relationship between increasing level across tertiles of volume of exercise and the odds of discontinuation of glucose-lowering medication with the standard care group having the lowest odds of discontinuation and the most adherent (3rd) tertile having the highest odds of discontinuing their anti-diabetes drugs (i.e. glucose-lowering medication).

Outcome measures

The primary outcome measure will be the dichotomous event of discontinuation of glucose-lowering medication when assessed at 12-month follow-up.

Additional secondary outcomes when assessed at 12-month follow-up are:

- Change in HbA1c from baseline to 12-month follow-up.
- Reductions, or intensification in glucose-lowering medication (number of individuals).

- Changes in laboratory tests (metabolic markers of glycemic control and lipids)
- Changes in blood pressure
- Changes in body composition (body-mass, BMI, fat-mass, lean body-mass, abdominal fat-mass, body fat percentage)
- Changes in physical fitness, physical activity and diet (VO_2max and relative VO_2max, number of steps pr day, time sitting pr day, energy intake)
- Reduction of lipid-lowering and blood pressure-lowering medications.

The primary outcome, discontinuation of glucose-lowering medication, and a number of the secondary outcomes (HbA1c, intensification and reduction of glucose-lowering medication, blood pressure-lowering and lipid-lowering medications and body composition) was assessed at baseline, 3, 6, 9 and 12-month follow-up. Energy intake, 2-hour OGTT test and VO2max was assessed at baseline and 12-month follow-up.

Primary objective and outcome

The primary objective of the present study is to explore whether odds of discontinuing glucose-lowering medication is associated with the volume of exercise (exercise tertiles) in patients with T2D after the 12-month U-TURN lifestyle intervention.

The primary outcome, discontinuation, is defined as having a glucose-lowering score equal to 0 at 12-months follow-up based on the medical algorithm (13).

We have chosen this outcome as glucose-lowering medications are associated with increased risk of side-effects, decreased well-being and increased financial costs (2, 16). Thus, their use has a major impact on patients with T2D.

Key secondary objectives and outcomes

The key secondary objective of the present study is to investigate the association between the volume of exercise (exercise tertiles and the standard of care [original control group]) and HbA1c.

The key secondary outcome is the change in HbA1c from baseline to 12-month follow-

up.

This outcome is chosen since the reduction of HbA1c is associated with a reduced risk of diabetic complications (17, 18). This makes it a highly relevant outcome measure for patients with T2D. Furthermore, HbA1c is an outcome measure accepted by the Food and Drug Administration (USA) and the European Medicines Agency (Europe), when the efficacy of new glucose-lowering drugs is assessed.

Other secondary objectives and outcomes

Additional secondary objectives of the present study include exploration of the effects of other clinically important outcome variables. The effects on these variables will be provided to support the clinical interpretation of the primary and key secondary outcomes. These measures will be explored based on the three exercise tertiles. The additional secondary outcomes include the between-group difference in:

1. Reductions, or intensification in glucose-lowering medication (number of individuals) from baseline to 12-month follow-up.
2. Metabolic markers of glycemic control (fasting insulin, fasting glucose, 2-hour OGTT glucose) from baseline to 12-months follow-up.
3. Lipids (total cholesterol, LDL and HDL) from baseline to 12-month follow-up.
4. Systolic and diastolic blood pressure from baseline to 12-months follow-up.
5. Body composition (body mass, BMI, fat mass [total, android], lean body mass, body fat percentage) from baseline to 12-months follow-up.
6. Changes in physical fitness, physical activity and diet (VO2max and relative VO2max, number of steps pr day, time sitting pr day, energy intake)

Lipid-lowering and blood pressure-lowering medications from baseline to 12-months follow-up. This will be based on the medical score from the medical algorithm. Every step in the medical algorithm awards 1 point. This point is either added upon progression (if intensifying the treatment) or subtracted (in case of full or partial discontinuation of the treatment). The outcome “reduction” is defined as medical score lower than the individual’s baseline medical score, at 12-month follow-up, whereas the outcome “intensification” is defined as a medical score above baseline medical score,

at 12-month follow-up.

DATA REDUCTION APPROACH, JUSTIFICATION AND SPECIFICATION OF THE TERTILE SPLIT.

The data will be processed according to a protocol approved by the U-Turn steering group. Steps are taken to ensure that the "accepted total volume of exercise" includes registered sessions that can be classified as "exercise" and not physical activity. In addition, steps are taken to ensure that the final data set does not include "false starts" or "long-sessions" where participants eventually forgot to stop their watches after a training. The principles for data reduction are summarized below and in Figure 1.

Observations with a registered duration of less than 10 min are deleted as they are assumed to be "false starts", and 10 minutes is a generally accepted minimal duration for qualification as "structured exercise" or an "exercise session". The same procedure was applied in the primary study.

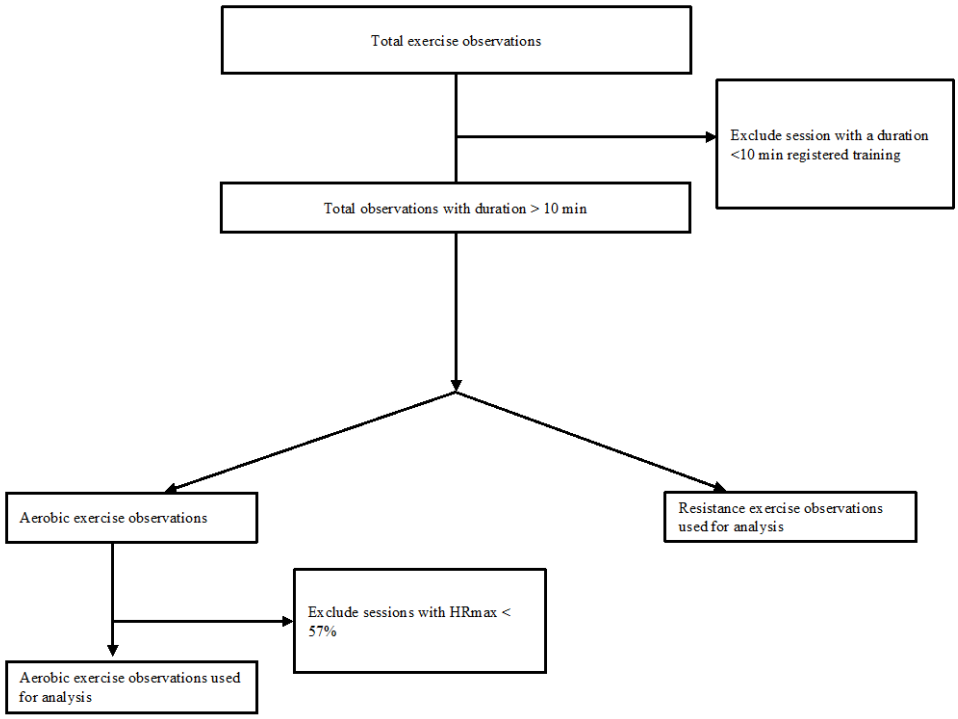
If registrations demonstrate a discrepancy between the "registered duration" of the session on the polar watch and "time spent with the pulse-belt on" (i.e. time registered with an active training heart rate), then the accumulated time spent in the different Polar heart rate zones (corresponding to time spent with the pulse-belt on with a HRmax > 50%) is used as a measure of the duration of a session, as opposed to the duration registered on the Polar-watch by the participant.

In sessions where no heart rate is registered the duration of the training must be based on the registered time and the average heart rate will be imputed. The imputed heart rate used will be based on the mean heart rate of the other exercise sessions (aerobic) during the relevant phase of the intervention (phase 1, 2, or 3).

Sessions with an accumulated time spent in different heart rate-zones of less than 10 min are assumed to be associated with a malfunction in the pulse registration and hence treated as sessions with no heart rate registered. In these cases, the registered time of the session is utilized for duration and a heart rate is imputed.

Aerobic observations with a HRmax of less than 57 % are excluded as they are classified as very light intensity (physical activity) rather than moderate to vigorous intensity exercise (19). This intensity cutoff is also accepted as a cutoff between exercise and physical activity.

Figure 1: Flow-chart of the data-reduction process



Tertile split

The volume of aerobic and resistance training for each month (frequency * mean duration of sessions from the specific month) will be calculated for every participant based on sessions with a HRmax above 57%. The participants are then stratified into tertiles based on the "total volume of exercise" (aerobic + resistance) during the intervention across all 12 months.

Analysis of Objectives and Outcomes

Primary endpoint

The analysis of the primary endpoint is based on a sequential analytic approach in order to maintain the type 1 error rate.

First, we investigate whether a trend in the odds ratio in glucose-lowering medication between StC and each of the exercise tertiles (1st Low; 2nd Intermediate, and 3rd High, respectively) discontinued their anti-diabetics at 12-months follow-up. If present ($p < .10$ will be considered indicative), then between group comparisons for effect size estimation is initiated in the following order;

1) StC vs. T3 (highest exercise volume). If a difference is present ($p<.05$, 2-sided) then the next between group comparison is performed. If not – then the statistical inference sequence will be terminated
2) StC vs. T2 (moderate exercise volume). If a difference is present ($p<.05$, 2-sided) then the next between group comparison is performed. If not – then the statistical inference sequence is terminated
3) StC vs. T1 (Lowest exercise volume). If a difference is present ($p<.05$, 2-sided) then the next between group comparison is performed. If not – then the statistical inference sequence is terminated
4) T3 vs. T1. If a difference is present ($p<.05$, 2-sided) then the next between group comparison is performed. If not – then the statistical inference sequence is terminated
5) T3 vs. T2. If a difference is present ($p<.05$, 2-sided) then the next between group comparison is performed. If not – then the statistical inference sequence is terminated
6) T2 vs. T1. If a difference is present ($p<.05$, 2-sided) then the next between group comparison is performed.

The analysis will be based on the *as-observed population* (missing data will not be imputed in the primary analyses).

Key secondary and other endpoints

The glycemic secondary outcome, i.e. reduction in HbA1c, will be analyzed using repeated-measures linear mixed modeling to explore the trajectories (interaction) between exercise-tertiles and time (months); this is recorded with assessments done at 3, 6, 9, and 12 months follow up.

Along with the analysis of discontinuation, the intensification or reduction of

the prescribed glucose-lowering medication for each time point after baseline (3, 6, 9, and 12 months) for each exercise-tertile will be reported.

Other exploratory endpoints, blood pressure- and lipid-lowering medication, body composition and markers of glycemic control etc., will be reported as within and between group changes from baseline to follow-up.

Sample size considerations

As previously described (13), the sample size in this study was based on what was considered feasible, within the local context, enabling up to 120 participants to be enrolled in the trial period (April 29th 2015 to August 17th 2017). The sample size of the main-study was truncated at 120 participants or the N reached at the end of the recruitment period. At the end of the recruitment period, 64 patients were randomized to the intervention group and 34 patients to the standard care group. No formal statistical power analysis was performed for the present sub-study outcomes.

Statistical methods

All data were collected longitudinally. All data were entered into the database twice, by two separate staff members. Hereafter, a third staff member checked the data entries for inconsistencies. Inconsistencies were corrected using the original registrations and the participant journals.

Causal inference from observational studies almost always have bias because prognostic factors are unequally distributed between patients exposed or not exposed to an intervention. Due to the post-hoc allocation into four apparent “treatment groups” which differ from the random allocation in the original study, this study framework represents comparative effectiveness research, with secondary aims and objective testing for superiority rather than equivalence like for the primary outcome; thus, the inference from this will remain observational (i.e. not a causal design).

To account for observed confounding all participants are assigned a propensity of being allocated to an “exercise volume” group. As the original design included a standard care group, participants will be assigned per default a propensity of 0, as they did not exercise; i.e. this group will maintain to be a randomized control group.

One alternative to the standard approach is propensity analysis, in which groups are covariate adjusted according to the likelihood of membership in exposed (1st, 2nd, 3rd tertile) or unexposed groups. Propensity methods can deal with multiple data-driven prognostic factors, even if there are relatively few patients having outcome events (i.e. individuals getting off their anti-diabetics). All participants in the original intervention group will be assigned a propensity score of being allocated to each of the three tertiles (T1 – low, T2 – moderate, T3 – high). The propensity of being allocated to a tertile will be data-driven based on baseline variables (pre-randomization [i.e. not influenced post hoc elements]).

Pre-specified covariates of prognostic value

Selection of covariates in traditional risk-adjusted (multivariable) regression analyses should be based on the research question at hand and on substantial knowledge such as what is biologically plausible; i.e. not data-driven (20). From this perspective the following covariates were suggested a priori to all subsequent models:

1) Sex – *Males having a higher chance of exercise adherence vs. females* (21, 22).
2) BMI – *Lower BMI is associated with higher adherence to exercise* (21, 22).
3) Educational level – *Higher educational level is associated with higher adherence to exercise* (23).
4) Fitness level at baseline – *Baseline exercise level is associated with higher adherence to exercise* (21, 24).
5) Age – *lower age is associated with higher adherence to exercise* (21).

Propensity score analysis vs Traditional multivariable regression analysis

We will explore whether these prespecified covariates have an impact on the research question by applying traditional (multivariable) regression analysis. In this we will examine the association between treatment selection (1st, 2nd, 3rd tertile vs control) and the outcome of interest (off the anti-diabetic drugs). Based on these potentially important prognostic variables (assessed at baseline), it levels the playing field across prognostic factors between groups by creating apparently "prognostically homogenous strata" and combining results across these strata. By applying multivariable regression, we will simultaneously include a number of the proposed prognostic factors and yield

a single estimate of “exercise treatment effect” adjusted for these factors (i.e. possible confounders).

Rather than only applying these prespecified covariates in our primary models we will attempt to adjust for a single propensity covariate due to the limited number of events (25) by using propensity score adjustment we will compute and assign a propensity score to each participant (i.e. a unique figure indicative of the likelihood of being allocated to one of the three subsequent tertiles or control group - conditioning on that participants status on measured prognostic factors.

We will apply the propensity score, computed as a single continuous variable ranging from 0 to 1, in our primary analyses. First, we will conduct a regression analysis with propensity score and group allocation as the independent variables. As for a multivariable adjusted regression, this approach will attempt to create prognostic variable balance by comparing the outcome between the intervention and the control within groups of patcipants that are homogeneous, in this case for their propensity to receive the allocated group (1^{st}, 2^{nd}, or 3^{rd} tertile vs control).

By applying both adjustment strategies (with propensity score adjustment being our primary approach) we will conduct a simple regression analysis adjusting for propensity score and compare it to the findings from the traditional multivariable approach (26). As recently emphasized by Agoritsas and colleagues, a major caveat to propensity scores, like traditional regression, is that it does not address residual confounding (e.g. if important baseline measures not collected to model from in the first place); empirically, results are often very similar to traditional regression (26).

Statistical computation

Following the estimation of the propensity score, the outcome is analyzed using multivariable logistic regression adjusted for the propensity score.

For the primary analysis, group (StC, T1, T2, and T3) is treated as a categorical exposure variable generating a p-value for trend. For the between-group comparison, the exposure is treated as an ordinal categorical variable (expected increases in odds of discontinuation on anti-diabetics) from StC (lowest odds) to T3 (highest odds).

Dichotomous outcomes assessed during 12 months (repeated measures):

- Discontinuation of glucose-lowering medication
- Reduction in glucose-lowering medication
- Intensification in glucose-lowering medication
- Reduction in lipid-lowering medication
- Reduction in blood pressure-lowering medication

Continuous outcomes assessed after 12 months (not repeated measures):

- Fasting insulin
- Fasting glucose
- 2-h OGTT glucose
- VO2max
- Relative VO2max
- Energy intake

Continuous outcomes assessed during 12 months (repeated measures):

- Hemoglobin A1c%
- Body mass
- BMI
- Fat mass
- Lean body mass
- Abdominal fat mass
- Steps daily
- Sitting time
- Systolic blood pressure
- Diastolic blood pressure
- Total cholesterol
- LDL
- HDL
- Triglycerides

For the statistical analyses, we will primarily use the statistical software R (version

3.3.3 or newer) (27). The following codes will be used for the main analyses:

```
#Primary analysis of primary outcome:
  glmer(MedDiscont ~ TIMEfac + GROUP + TIMEfac*GROUP + GlucoseLoweringMedScore_0 + PS + (1 |
  PtID), data = dA, family = binomial)

#Primary analysis of secondary outcomes:
  glmer(OutcomeY ~ TIMEfac + GROUP + TIMEfac*GROUP + OutcomeY_0 + PS + (1 | PtID), data = dA,
  family = binomial)
  lmer(OutcomeZ ~ TIMEfac + GROUP + TIMEfac*GROUP + OutcomeZ_0 + PS + (1 | PtID), data = dA)

#For secondary outcomes only measures available at baseline and 12 months:
  lm(OutcomeZ ~ GROUP + OutcomeZ_0 + PS, data = dA)
```

Handling of missing data and sensitivity analysis.

Substantial instances of missing data are a serious problem that could undermines the scientific credibility of causal conclusions from clinical trials. The assumption that sophisticated statistical analysis methods can compensate for missing data are not justified, which is why aspects of the U-Turn trial design limiting the likelihood of missing data was an important objective. In the study project (i.e. SAP) we analyze and report on the basis of data that has already been collected.

There is no universal method for handling missing data in a clinical trial, since each trial has its own set of design and measurement characteristics. The range of approaches to modeling and inference is extremely broad, and no single method or class of methods is suitable for all situations. For this particular study our primary analysis will be based on the Intention-treat population (patients will remain in the original randomized group independent of what happened post-randomization). However, different types of adjustment methods for missing data were considered:
complete-case analysis, single imputation methods, estimating-equation methods, and ‘Data as observed’ methods based on a statistical model.

As an example of a statistical model, continuous measures (repeated over time) will be assumed to have a normal distribution with a specified form of mean and covariance matrix. Methods that are based on a statistical model such as this include (restricted) maximum likelihood, in which estimates and standard errors are based on the likelihood function given the observed data

For the purpose of sensitivity analyses we will assess the robustness of the primary analyses by using multiple imputation techniques, in which multiple sets of plausible values for missing data are created from their model-based predictive distribution, and estimates and standard errors are obtained with the use of multiple-imputation combining rules.

Furthermore, we will in a sensitivity analysis investigate the impact of adjusting for energy intake on the results of our primary analyses.

Dose-response modeling

A key objective of exploratory studies like this is to adequately characterize the dose response relationship of exercise adherence (dose) and clinical outcome (off the glucose-lowering medication). An important decision here will be on the choice of a suitable dose response function to support dose selection for the subsequent confirmatory studies. As a tertiary goal of this exercise dosage project, we will compare different approaches for model selection and model averaging using mathematical properties as well as simulations.

Prognostic value

We will use univariate logistic regression analyses to investigate if some of the independent variables appear significantly associated with clinically outcome (discontinuation off of the glucose-lowering drugs). By use of simple logistics regression for several variables, the odds, will indicate some association. Subsequently we will construct a multiple logistic regression model, to explore whether any of these remain statistically significantly and independently associated with the clinical outcome.

Anticipated outline of the study report

Figure 1: Flow diagram illustrating how many patients adhere (= no attrition) to the program over time, illustrated for each of the three tertiles + the control group.

Table 1: Baseline characteristics in the "three exercise groups": The characteristics will be presented per high, moderate and low group and in total as means (standard

deviation) or medians (interquartile range). Dichotomous and categorical data will be presented as actual numbers (%).

	Standard Care (n=)	**Lower tertile (n=)**	**Intermediate tertile (n=)**	**Upper tertile (n=)**
Demographics				
Age, years				
Female, no (%)				
Type 2 diabetes duration, years				
Glycemic control				
Hemoglobin A_{1c}, %				
Fasting glucose, mg/dL				
Fasting insulin, µIU/mL				
2-h OGTT glucose, mg/dL				
Lipids				
Total cholesterol, mg/dl				
LDL, mg/dl				
HDL, mg/dl				
Triglycerides, mg/dl				
Blood pressure				
Systolic, mm Hg				
Diastolic, mm Hg				
Body composition				
Body mass, kg				
BMI, kg/m^2				
Fat mass, kg				
Lean body mass, kg				
Abdominal fat mass, kg				
Body fat percentage, %				
Physical fitness, physical activity and diet				
VO_2max, ml O_2/min				
Relative VO_2max, ml O_2/kg/min				
Steps, daily steps				
Sitting, min/day				
Energy intake, kcal/d				

Table 2.

Phase	**Exercise**	**Prescribed/implemented**	**Low tertile, mean (sd) (n=)**	**Intermediate tertile, mean (sd) (n=)**	**High tertile, mean (sd) (n=)**

1	Aerobic volume	180-360 min/week *Min to max*			
	Strength volume	60 min/week *Min to max*			
	Combined volume	240-300 min/week *Min to max*			
	Intensity	60-80% of HRmax			
2	Aerobic volume	240-300 min/week *Min to max*			
	Strength volume	60 min/week *Min to max*			
	Combined volume	300-360 min/week *Min to max*			
	Intensity	70-90% of HRmax			
3	Aerobic volume	210 min/week *Min to max*			
	Strength volume	90 min/week *Min to max*			
	Combined volume	300 min/week *Min to max*			
	Intensity	70-90% of HRmax			
12 months					
	Aerobic volume	mean based on 12 months			
	Strength volume	mean based on 12 months			

Figure 2: Bar-chart illustrating the proportion of participants (with 95% CIs) with discontinuation of anti-diabetes drugs at 12-month on the y-axis for each tertile. This plot will also reveal the absolute numbers as well as the results of the primary statistical model.

Table 3: Between-group comparisons (high, moderate, low) of the changes in the primary, key secondary and exploratory outcomes from baseline to 12-month follow-up. The changes will be reported by tertile. The results will be reported as the difference in change between the lower tertile and the moderate and high tertile, respectively.

Outcome variable	**Standard Care (n=)**	**Lower tertile (n=)**	**Intermediate tertile (n=)**	**Upper tertile (n=)**	**Between-group difference: *P*-value**	**Between-group difference: Estimate and 95% CIs**
Secondary Outcome						
Hemoglobin A1c%					U vs. StC: XX I vs. StC: XX L vs. StC: XX	

Proportion of participants with reduction in glucose-lowering medication						
Proportion of participants with intensification in glucose-lowering medication						
Glycemic control						
Fasting insulin, μIU/mL						
Fasting glucose, mg/dL						
2-h OGTT glucose, mg/dL						
Lipids						
Total cholesterol, mg/dL						
LDL, mg/dL						
HDL, mg/dL						
Triglycerides, mg/dL						
Blood pressure						
Systolic, mmgHg						
Diastolic, mmHg						
Body composition						
Body mass, kg						
BMI, kg/m^2						
Fat mass, kg						
Lean body mass, kg						
Abdominal fat mass, kg						
Body fat percentage, %						
Physical fitness, physical activity and diet						
VO_2max, ml O_2/min						
Relative VO_2max, ml O_2/kg/min						

Steps, daily steps						
Sitting, min/day						
Energy intake, kcal/d						
Medication						
Proportion of participants with reduction in lipid-lowering medication						
Proportion of participants with reduction in blood pressure-lowering medication						

Table 4: Prognostic value of various baseline assessments to predict who can discontinue glucose-lowering drugs after 12 months

	Odds Ratio (95% CI)		
	Simple logistic regression	Multiple logistic regression (prespecified covariates)†	Multiple logistic regression (post hoc covariates)††
Exercise dose (0= control; 1=low, 2= intermediate, 3=high)			
Demographics			
Age at consent, years			
Female, no (%)			
Type 2 diabetes duration, years			
Glycemic control			
Hemoglobin A_{1c}, %			
Fasting glucose, median (IQR), mg/dL			
Fasting insulin, µIU/mL			
2-h glucose, mg/dL			
Lipids			
Total cholesterol, mg/dl			
LDL, mg/dl			
HDL, mg/dl			
Triglycerides, mg/dl			
Blood pressure			
Systolic, mm Hg			
Diastolic, mm Hg			
Body composition			
Body mass, kg			

BMI			
Fat mass, kg			
Lean body mass, kg			
Abdominal fat mass, kg			
Body fat percentage, %			
Physical fitness, physical activity and diet			
VO_2max, ml O_2/min			
Relative VO_2max, ml O_2/kg/min			
Steps, daily steps			
Sitting, min/day			
Energy intake, kcal/d			

† Sex – *Males having a higher chance of exercise adherence vs. females;* BMI – *Lower BMI is associated with higher adherence to exercise;* Educational level – *Higher educational level is associated with higher adherence to exercise;* Fitness level at baseline – *Baseline exercise level is associated with higher adherence to exercise;* Age – *lower age is associated with higher adherence to exercise*.

††: *Simultaneous adjustment for all statistically significant covariates (derived from the univariate approach).*

Reference list:

1. Huang ES, Brown SE, Ewigman BG, Foley EC, Meltzer DO. Patient perceptions of quality of life with diabetes-related complications and treatments. Diabetes Care. 2007;30(10):2478-83.
2. Espeland MA, Glick HA, Bertoni A, Brancati FL, Bray GA, Clark JM, et al. Impact of an intensive lifestyle intervention on use and cost of medical services among overweight and obese adults with type 2 diabetes: the action for health in diabetes. Diabetes Care. 2014;37(9):2548-56.
3. Prasad-Reddy L, Isaacs D. A clinical review of GLP-1 receptor agonists: efficacy and safety in diabetes and beyond. Drugs Context. 2015;4:212283.
4. Ried-Larsen M, MacDonald CS, Johansen MY, Hansen KB, Christensen R, Almdal TP, et al. Why prescribe exercise as therapy in type 2 diabetes? We have a pill for that! Diabetes Metab Res Rev. 2018:e2999.
5. Balducci S, Zanuso S, Nicolucci A, De Feo P, Cavallo S, Cardelli P, et al. Effect of an intensive exercise intervention strategy on modifiable cardiovascular risk factors in subjects with type 2 diabetes mellitus: a randomized controlled trial: the Italian Diabetes and Exercise Study (IDES). Arch Intern Med. 2010;170(20):1794-803.
6. Vancea DM, Vancea JN, Pires MI, Reis MA, Moura RB, Dib SA. Effect of frequency of physical exercise on glycemic control and body composition in type 2 diabetic patients. Arq Bras Cardiol. 2009;92(1):23-30.
7. Sigal RJ, Kenny GP, Boule NG, Wells GA, Prud'homme D, Fortier M, et al. Effects of aerobic training, resistance training, or both on glycemic control in type 2 diabetes: a randomized trial. Ann Intern Med. 2007;147(6):357-69.
8. Di Loreto C, Fanelli C, Lucidi P, Murdolo G, De Cicco A, Parlanti N, et al. Make your diabetic patients walk: long-term impact of different amounts of physical activity on type 2 diabetes. Diabetes Care. 2005;28(6):1295-302.
9. Umpierre D, Ribeiro PA, Kramer CK, Leitao CB, Zucatti AT, Azevedo MJ, et al. Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis. JAMA. 2011;305(17):1790-9.
10. Umpierre D, Ribeiro PA, Schaan BD, Ribeiro JP. Volume of supervised exercise training impacts glycaemic control in patients with type 2 diabetes: a systematic review with meta-regression analysis. Diabetologia. 2013;56(2):242-51.
11. Boule NG, Kenny GP, Haddad E, Wells GA, Sigal RJ. Meta-analysis of the effect of structured exercise training on cardiorespiratory fitness in Type 2 diabetes mellitus. Diabetologia. 2003;46(8):1071-81.
12. Malin SK, Gerber R, Chipkin SR, Braun B. Independent and combined effects of exercise training and metformin on insulin sensitivity in individuals with prediabetes. Diabetes Care. 2012;35(1):131-6.
13. Ried-Larsen M, Christensen R, Hansen KB, Johansen MY, Pedersen M, Zacho M, et al. Head-to-head comparison of intensive lifestyle intervention (U-TURN)

versus conventional multifactorial care in patients with type 2 diabetes: protocol and rationale for an assessor-blinded, parallel group and randomised trial. BMJ Open. 2015;5(12):e009764.
14. Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, et al. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement executive summary. Diabetes Care. 2010;33(12):2692-6.
15. Johansen MY, MacDonald CS, Hansen KB, Karstoft K, Christensen R, Pedersen M, et al. Effect of an Intensive Lifestyle Intervention on Glycemic Control in Patients With Type 2 Diabetes: A Randomized Clinical Trial. JAMA. 2017;318(7):637-46.
16. Cramer JA. A systematic review of adherence with medications for diabetes. Diabetes Care. 2004;27(5):1218-24.
17. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA. 10-year follow-up of intensive glucose control in type 2 diabetes. N Engl J Med. 2008;359(15):1577-89.
18. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ. 2000;321(7258):405-12.
19. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc. 2011;43(7):1334-59.
20. Lydersen S. Statistical review: frequently given comments. Ann Rheum Dis. 2015;74(2):323-5.
21. Unick JL, Gaussoin SA, Hill JO, Jakicic JM, Bond DS, Hellgren M, et al. Four-Year Physical Activity Levels among Intervention Participants with Type 2 Diabetes. Med Sci Sports Exerc. 2016;48(12):2437-45.
22. Wing RR, Hamman RF, Bray GA, Delahanty L, Edelstein SL, Hill JO, et al. Achieving weight and activity goals among diabetes prevention program lifestyle participants. Obes Res. 2004;12(9):1426-34.
23. Forechi L, Mill JG, Griep RH, Santos I, Pitanga F, Molina M. Adherence to physical activity in adults with chronic diseases: ELSA-Brasil. Rev Saude Publica. 2018;52:31.
24. Eaglehouse YL, Venditti EM, Kramer MK, Arena VC, Vanderwood KK, Rockette-Wagner B, et al. Factors related to lifestyle goal achievement in a diabetes prevention program dissemination study. Transl Behav Med. 2017;7(4):873-80.
25. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000;19(4):453-73.
26. Agoritsas T, Merglen A, Shah ND, O'Donnell M, Guyatt GH. Adjusted Analyses in Studies Addressing Therapy and Harm: Users' Guides to the Medical Literature. JAMA. 2017;317(7):748-59.
27. Team RC. R: A language and environment for statistical

computing: R Foundation for Statistical Computing, Vienna, Austria.; 2013 [Available from: http://www.R-project.org/.