Statistical Analyses Plan (SAP)

Bicycling and mortality among individuals with diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)

Version: 2.0

Version date: November 13th 2019

SAP author: Mathias Ried-Larsen (MR-L)

Statistical analyst: Martin G. Rasmussen (MGR)

Statistical advisor: Anders Grøntved (AG)

Trial registration: www.clinicaltrials.gov

Affiliations:

MR-L: The Centre of Inflammation and Metabolism and the Centre for Physical Activity Research, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, 2100 Copenhagen, Denmark.

MGR and AG: Exercise Epidemiology Unit, Institute of Sport Science and Clinical Biomechanics, Faculty of Health Science, University of Sourthern Denmark, Campusvej 55, 5230, Odense M.

BACKGROUND

Premature death from all-causes and cardiovascular (CVD) is increased among diabetic individuals[1]. As only a limited number of antidiabetic medications are effective in preventing all-cause and CVD mortality in this patient group[2,3], it warrants more knowledge on effective means to reduce all-cause and CVD mortality due to diabetes.

Regular physical activity is recommended to diabetes patients as an important part of their diabetes treatment[4]. It is well established that structured exercise can improve glucose control, blood lipids, and reduce blood pressure in diabetes patients[5-8]. Cohort studies have shown an inverse association between overall physical activity, leisure time PA, and walking with outcomes such as mortality and cardiovascular disease in diabetes patients[9]. Although the associations between walking and diabetes complications have been suggested in earlier work, the associations have been inconsistent[9]. Some of this inconsistency may relate to the variation in walking pace and thus intensity[10], as the risk of both all cause and coronary heart disease mortality is lower with faster walking paces compared to slower[11]. Meeting physical activity recommendations both in terms of total physical activity volume and intensity remains a major challenge, especially in diabetes patients[12,13]. Lack of time is often quoted as a barrier to behavior change, so finding ways to incorporate activity with zero net time impact has potential to be more successful. Bicycling is one such potential candidate activity, as it can feasibly replace motor transport for short-to-medium distance trips (REF). Bicycling could therefore help to ensure engagement of diabetes patients in physical activity as a part of the daily routine without requiring time for structured exercise. As moderate-to-high intensities are reached during bicycling with self-selected pace in adults[14,15], and bicycling are among the preferred activities among persons with T2D[16,17], this may be a physical activity that could effectively be adopted by individuals with diabetes as it is manageable and can serve a second purpose as commuting. It is well-established that there is a strong association between increased bicycling and improvements in cardiovascular risk factors, reduced risk of all-cause and cause-specific mortality among the general population[18-20]. There are, however, to our knowledge no studies examining the relationship between bicycling and mortality in a diabetic population.

AIMS

The primary aim of the study is to study the relationship between overall bicycling and all-cause mortality and secondarily cardiovascular disease mortality among individuals with diabetes from

European countries. A secondary aim will be to study the relation of change in bicycling to all-cause mortality and cardiovascular disease mortality.

METHODS

Study design and setting

The study is a nested cohort study in European Prospective Investigation into Cancer and Nutrition (EPIC) cohort (Figure 1). In EPIC, 29 centers in 10 western European countries have collected information on nutrition, lifestyle, anthropometrics and medical history from more than 521 000 study individuals participating in this prospective cohort study[21]. Dietary, medical history and lifestyle were assessed by questionnaires. Blood samples from the participants were taken at the study centers. Baseline (first examination) information was collected 1992-2000 and second examination for exposure has since been obtained at least once in every cohort (Figure 2). Exposure data has subsequently, to the two examinations, been linked to information on vital status and on cause of death were obtained through record linkages with national, regional or local registers, regional health departments, physicians or hospitals, active follow-up or health insurance[22]. The ethical review boards from the International Agency for Research on Cancer (IARC) and all local participating enters approved the study.

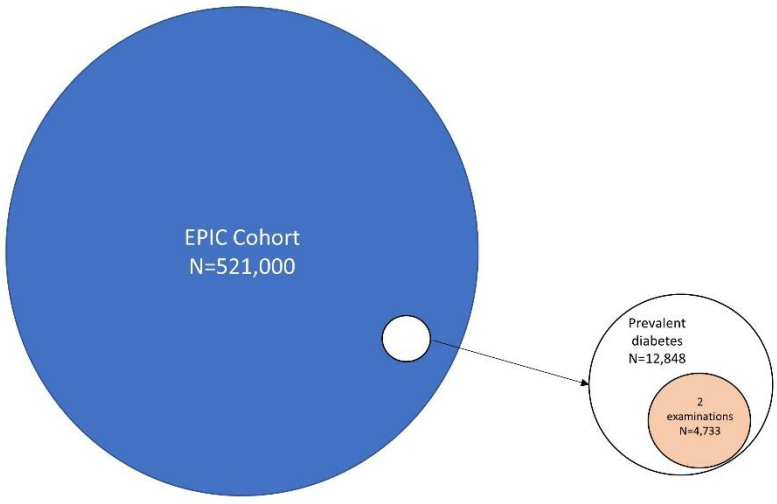

Figure 1 Overview of the sub-study. Participants with prevalent diabetes are sampled from the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. For a secondary analysis, participants with 2 examinations are included.

Study population

This nested prospective cohort study, sampled individuals with prevalent diabetes in the EPIC cohorts at the baseline assessments in the years 1992-2000. Self-reported and confirmed diabetes (validated by a second source (at least 1), including repeated self-report, contact with physician, linkage to register later point, intake of diabetes medicine, registration of diabetic chiropody, baseline glycated hemoglobin>=6.0%, five annual blood glucose measurements or two blood glucose measurements per year for five consecutive years) cases were included in the analyses. No information is available to distinguish between type 1 and type 2 diabetes across the population but type 1 is rare by comparison.

The nested cohort for this study comprised 12,848 individuals with self-reported and/or confirmed diabetes (Figure 1). Mean (SD) follow-up time was 14.6 (4.8) years. Among the 12,848 prevalent diabetes cases, 5,809 cases were confirmed at baseline as described above.

Outcomes

The primary and secondary outcomes are the risk of all-cause and CVD mortality during follow-up, respectively. Mortality data were coded according to the *International Classification of Diseases, Injuries, and Causes of Death, Tenth Revision*, using the codes I00-I99 for CVD mortality.

Exposure data sources

Exposure is assessed through a lifestyle questionnaire comprising information about lifestyle and demographic factors including physical activity, age, educational level and sex at the first (baseline) and second examination (mean (SD) time between examinations; 5.3 (2.3) years, Figure 2). The questionnaires were received by mail. At the first (baseline) examination the participants also underwent a physical examination including blood sampling, measurements of blood pressure and anthropometry.

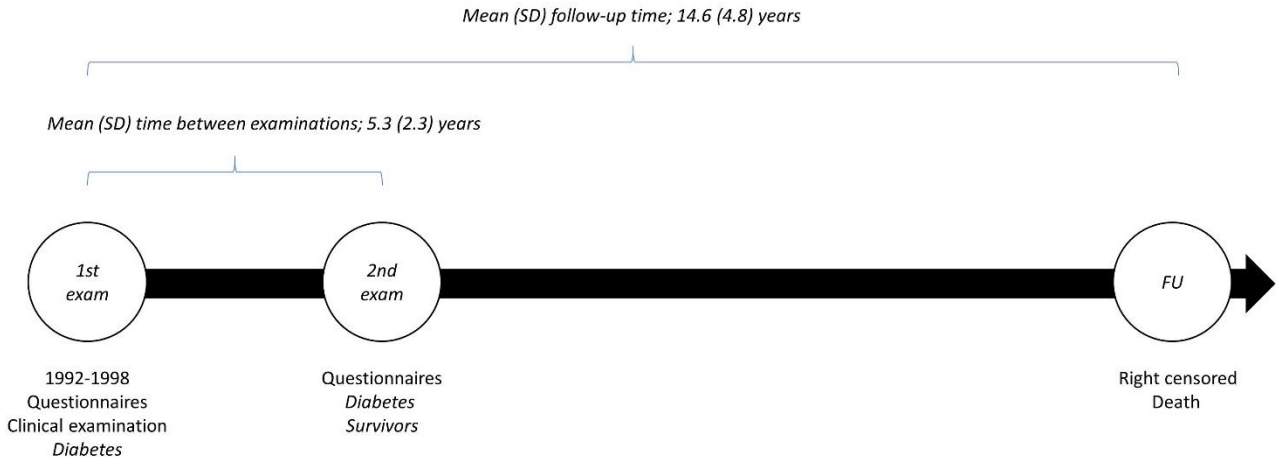

Figure 2 Coherency between examinations in relation to follow-up. Participants from the baseline examinations (1st exam) were invited to participate in a follow-up examination (2nd exam). Participants were followed until death or end of follow-up.

Assessment of physical activity and bicycling

Information about physical activity included duration and frequency of bicycling, walking, gardening, do-it-yourself activities, house-hold work, sports, number of stairs climbed and occupational physical activity. Information on time spend bicycling was reported according to season (winter and summer) and domain (recreational and commuting). Time spent on seasonal (winter and summer) bicycling is collapsed into one variable (mean of summer and winter) to calculate the total bicycling time. If one answer is missing in the seasonal variables for bicycling the value of the missing answer is imputed using sex and age stratified (six categories) using weighted means of cycling according to the ratio of time spend bicycling during summer vs. winter in participants who have answered both questions)[23]. Patterns of missing data will be investigated. *A priori*, the less restrictive missing at random (MAR) assumption is considered more reasonable than the missing data be missing completely at random (MCAR).

Total bicycling is reported in categories of: (0), (>0 to <60), (≤60 to >150), (≤150 to < 300) and (≥300) minutes/week. Changes in total cycling from baseline (1st examination) to the follow-up examination (2nd examination) is calculated as the difference in total time spend between the two examinations. The participants are then stratified according to non-cycling (participants that did not report bicycling at any examinations), ceased to bicycle (participants that bicycled at the first but not at the second examination), initiated bicycling (participants that did not report bicycling at the first

but report bicycling at the second examination), or continued cycling (reported bicycling at the first and second examination as outlined above). Leisure time physical activity (LTPA) energy expenditure (MET-h/week) is calculated by summing energy expenditure from the activities: Gardening, Do-it yourself, stair-climbing and housework activities, walking and vigorous activities. Bicycling is NOT included in this variable. As information on stair-climbing was only available for four study centers at follow-up, this variable is excluded from the calculation of the LTPA energy expenditure in the analyses of change of bicycling habits. Occupational physical activity is reported in categories of sedentary occupation, standing occupation, manual/heavy manual work or non-worker. Persons reporting unknown occupational physical activity are excluded from the analysis.

Other covariates

Dietary intake: The dietary intake, including alcohol consumption was assessed using either a quantitative questionnaire, a semi-quantitative questionnaire or a combination of semi-quantitative questionnaire and 7- or 14-day record. Estimated individual energy and nutrient intakes in this study is based the standardized EPIC Nutrient Database (ENDB)[24]. The Mediterranean diet is associated with improved metabolic control and decreased risk of diabetes[25,26]. Therefore, diet is included as adherence to this type of diet. The relative Mediterranean diet score (rMED) is used to assess adherence to a Mediterranean diet[27]. The score entails consumption of nine components included in the Mediterranean diet[28]: vegetables, legumes, fruits and nuts, cereals, fish and seafood, olive oil, moderate alcohol consumption, meat and meat products and dairy products[25,27]. The rMED score ranges from lowest adherence (0 points) to highest adherence (18 points). The construction of the score has been described in detailed elsewhere[25]. The rMED is further categorized to reflect low (0–6 points), medium (7–10 points), or high (11–18 points) adherence to the Mediterranean diet[27], and energy intake is calculated (kcal/day)[24]. Data is not available for the 2nd.

Smoking: Participants are grouped according self-reported smoking status (never-smoker, former smoker, and current smoker).

Education: Participants are categorized according to the highest self-reported completed educational level (no formal education, primary school, technical school, secondary school, and university degree)

Body composition: Weight, height, waist circumference (WC), and hip circumference were measured with the participants wearing light clothes and without wearing shoes[29]. Body mass index was calculated as weight in kilograms divided by squared height in meters (kg/m^2).

Prevalent co-morbidities: Information on prevalent hypertension, hyperlipidemia, and previous cardiovascular disease.

Diabetes duration: Diabetes duration is calculated based on the age at entry and self-reported age at diagnosis or time at diagnosis based on other sources.

Statistical analysis

Descriptive data for continuous variables is described as medians with interquartile ranges and categorized data were expressed as proportions within strata of bicycling volume.

Primary and secondary outcomes: The risk of and outcome (all-cause (primary outcome) or CVD mortality (secondary outcome)), are expressed as hazard ratios (HRs) with 95% confidence intervals (CIs) according to volume of cycling at baseline (see categorization above) using Cox proportional hazard regression with age as the underlying time-scale. Analyses will be corrected for delayed entry. Participants are considered as 'at-risk' from age at the baseline examination. The data is right-truncated at age of death, emigration, withdrawal from the study or at the end of follow-up. Statistical significance level is set at an $\alpha = 0.05$ (2-sided).

The analyses will be based on the following sequence:

1) A crude model (Model 1) with all-cause or CVD mortality as outcomes, will be fitted with categories of volume of bicycling as exposure and adjusted for sex (male/female), study center and age (continuous in years).
2) Model 1 + educational level, smoking status, diabetes duration, adherence to rMED score (low/medium/high), total energy intake, physical activity (quartiles of LTPA energy expenditure and four categories of occupational PA).
3) Model 2 + prevalent stroke, prevalent myocardia infarct, prevalent cancer, hypertension, hyperlipidemia, central obesity (yes/no according to sex[30]).
4) Effect modification (model 2) by sex is investigated (see below)

The trend across categories of total volume of bicycling is tested by treating the median within each category as a continuous variable in the model.

As a secondary analysis, the association between the all-cause or CVD mortality and change (from the first to the second examination) in total bicycling habits is investigated using the analytic procedure described for the primary. The model is adjusted for covariates in model 2 (baseline variables: sex, study center, educational level, smoking status, adherence to rMED score, total energy intake, physical activity (leisure and occupational). Age and diabetes duration at baseline is NOT included). The model will furthermore be adjusted for age, diabetes duration, smoking status, physical activity (leisure and occupational) at the second examination (to account for residual change)[31]. To the extent they are available the analyses are adjusted for prevalent stroke, myocardial infarction and cancer at follow-up. Dietary variables are not available for this analysis. For the secondary analyses, the participants are considered at risk from age at the second examination.

Effect modification: As the association between physical activity and all-cause and CVD mortality may differ between men and women[9], effect modification by sex (male/female) is evaluated statistically using the likelihood-ratio test when comparing model 2 by entering an interaction term for sex (sex *times* bicycling) with main effects compared to the model including the main effects only.

Model assumptions: Departure from proportional hazards (PH) assumption will be evaluated by tests and graphs of Shoenfeld residuals. If the model variables do not meet the PH assumption, a stratified Cox model will be employed for variables that do not meet the proportional hazards assumption. If the violation of the PH assumption is related to statistical power, the variables in question may be adapted (simplified) to increase the number of cases within each stratum (for categorical variables).

Sensitivity analyses: Several sensitivity analyses are planned to investigate the robustness of the primary (baseline exposure) analyses on the primary and secondary outcomes.

1. To investigate the risk of reverse causation and residual confounding, the following procedures are implemented (separately);
    a. Participants reporting any smoking at baseline are excluded from model 2
    b. Incident cases of death from all causes or CVD during the first 2 years of follow-up and prevalent stroke, myocardial infarction and cancers are excluded from model 2.
    c. Persons reporting any sports are excluded

2. The primary analyses all-cause and CVD mortality are repeated in the participants with a confirmed diabetes diagnosis (model 2).
3. Handling of missing data; To investigate the possibility of selection bias arising from missing data due to lost to follow-up (not participating in follow-up examination) or other reasons, we will conduct a multivariate chained equation imputation (MICE) approach including all covariates to impute missing values on covariates assuming data are missing at random on model 2 (above). Hazard ratios and standard errors will be obtained based on 20 imputed datasets.

Programming plan: All analyses are conducted using STATA IC V.15 (STATA Corp, College Station, Texas, USA). The programming plan for the primary outcome (all-cause mortality) is described below.

```
**********************************************************************************
**** stset D_Endfup, failure(Casemort) id(id) enter(D_Recrui) origin(D_Birth) scale(365.25) ****
**********************************************************************************
stcox i.cyc_kat_b Age_Recr i.Sex i.L_School i.Smoke_Stat Dur_Diab_all i.RMed_Score_C_quar
i.PA_Mets_nocyc_quar i.Pa_Work if analyticsample==1, strata(Center Qe_Energy_quar)
**********************************************************************************
```

The data is designated as survival data (using stset) with all-cause mortality (variable name 'Casemort', 2 levels) as outcome (failure) and age (scaled in years) as the underlying time axis (date variables used: D_Birth, D_Recrui and D_Endfup, all continuous variable). A Cox regression model is then fitted with bicycling (variable name 'cyc_kat_b', 5 levels), sex (variable name 'Sex', 2 levels), study center (variable name 'Center' with 26 levels) and age at recruitment (variable name 'Age_Recr', continuous variable).

Implementation of the statistical analysis plan

The SAP will be used as a work chart for the statistical analysis and for drafting and completing the study report (scientific article). The SAP will be implemented using the following steps:

1. The SAP is circulated and approved by all co-authors and subsequently registered at www.clinicaltrials.gov. This is done prior to the commencement of the statistical analyses.
2. Statistical analyses are performed (MG and AG)
3. A preliminary report is drafted (MR-L, MG, AG) and circulated among the key investigators (see below) and finalized
4. According to the standard EPIC publication guidelines, the draft is sent to the EPIC steering group committee for approval and assignment of co-author representatives from the participating EPIC centers. Specifications of authorship rules are specified elsewhere https://epic.iarc.fr/docs/EPIC_Publication%20Guidelines.pdf
5. Upon the approval and appointment of co-authors by the EPIC steering group, the report is circulated among co-authors for further comments and final approval.
6. When agreement about interpretation and conclusion is reached, the report is submitted to a scientific journal (Priority: 1) JAMA/JAMA internal medicine, 2) Annals of Internal Medicine, 3) Diabetes Care, 4) Diabetologia, 5) Mayo Clinic Proceedings.

Key investigators and writing group: Mathias Ried-Larsen (shared first author); Martin Gillies Rasmussen (shared first author); Kim Blond; Lars Bo Andersen; Søren Brage; Nick Wareham; Anders Grøntved

Anticipated outline of the study report (manuscript)

Figure 1: Kaplan-Meier plots of the primary (All-cause mortality, upper panel). Hazard ratios and 95% confidence intervals based on model 2 (see above) are reported.

Table 1 Baseline characteristics of individuals with diabetes across categories of total volume of bicycling					
	Total bicycling (minutes)				
	0 (N=)	*>0 to <60 (N=)*	*≤60 to >150 (N)*	*≤150 to >300 (N=)*	*≥300 (N=)*
Women, (%)					
Age, years					
Education (%)					
None					
Primary school					
Technical/professional school					
Secondary school					
Higher, including university degree					
Not specified					
Smoking					
Never					
Former					
Current					
Unknown					
Body mass index, kg/m^2					
Waist circumference, cm					
Men					
Women					
Central obesity (N (%) - yes/no)					
Leisure time physical activity without bicycling, MET-h/week					
Gardening, MET-h/week					
Do-it-yourself activities, MET-h/week					
Stair-climbing, MET-h/week					
Housework activities, MET-h/week					
Walking, MET-h/week					
Vigorous physical activity, MET-h/week					
Occupational physical activity (%)					
Sedentary occupation					
Standing occupation					
Manual work					
Heavy manual work					
Non-worker					
Unknown					
Energy intake (kcal/day)					

Adherence to the relative Mediterranean diet score (%)					
Low					
Medium					
High					
Prevalent co-morbidities					
Cancer (%)					
Stroke (%)					
Myocardial infarction (%)					
Hyperlipidemia (%)					
Hypertension (%)					
† Mean (standard deviation), ‡ Median (interquartiles), MET-h/week; Metabolic quartile – hours per week					

Table 2 Association between all-cause and cardiovascular disease mortality and total volume of bicycling						
	Total bicycling (minutes)					
	0	*>0 to <60*	*≤60 to >150*	*≤150 to >300*	*≥300*	*P for trend*
All-cause mortality						
Bicycling energy expenditure (METh/week)						
Incidence rate /1000 person-years						
Model 1 (HR and (95% CI))	1 [Reference]					
Model 2 (HR and (95% CI))	1 [Reference]					
Model 3 (HR and (95% CI))	1 [Reference]					
Cardiovascular mortality						
Incidence rate /1000 person-years						
Model 1 (HR and (95% CI))	1 [Reference]					
Model 2 (HR and (95% CI))	1 [Reference]					
Model 3 (HR and (95% CI))	1 [Reference]					

HR; Hazard rate, CI; Confidence interval

Model 1 is adjusted for sex, study center and age

Model 2 is adjusted for sex, study center, age, baseline educational level, smoking status, diabetes duration adherence to rMED score, total energy intake, leisure-time physical activity

Model 3 is adjusted for sex, study center, follow-up age, baseline educational level, smoking status, prevalent stroke, prevalent myocardia infarct, prevalent cancer, diabetes duration adherence to rMED score, total energy intake, leisure-time physical activity

Table 3 Association between all-cause or cardiovascular mortality and changes in bicycling from baseline to follow-up				
	Changes in total bicycling			
	No bicycling	*Cessation*	*Initiation*	*Continuation*
Median (interquartiles) time spend bicycling at follow-up (minutes)				
All-cause mortality				
Incidence rate /1000 person-years				
Model 1	1 [Reference]			
Model 2	1 [Reference]			
Cardiovascular mortality				
Incidence rate /1000 person-years				
Model 1	1 [Reference]			
Model 2	1 [Reference]			

HR; Hazard rate, CI; Confidence interval

Model 1 is adjusted for sex, study center and age at follow-up

Model 2 is adjusted for sex, study center, follow-up age, baseline educational level, smoking status at follow-up, diabetes duration adherence to rMED score (baseline and follow-up), total energy intake (baseline and follow-up), leisure-time physical activity (baseline and follow-up)

Supplementary material

sTable 1 – Sensitivity analyses; Association between all-cause and cardiovascular disease mortality and total volume of bicycling among persons with diabetes						
	Total bicycling (minutes)					
	0	*>0 to <60*	*≤60 to >150*	*≤150 to >300*	*≥300*	*P for trend*
All-cause mortality						
Incidence rate /1000 person-years						
Ever-smokers excluded (HR and (95% CI))	1 [Reference]					
Incident cases with the first 2 years of follow-up AND prevalent stroke, myocardial infarction and cancer excluded (HR and (95% CI))	1 [Reference]					
Confirmed diabetes only (HR and (95% CI))	1 [Reference]					
Missing data imputed (HR and (95% CI))	1 [Reference]					
Cardiovascular mortality						
Incidence rate /1000 person-years						
Ever-smokers excluded (HR and (95% CI))	1 [Reference]					
Incident cases with the first 2 years of follow-up AND prevalent stroke, myocardial infarction and cancer excluded (HR and (95% CI))	1 [Reference]					
Confirmed diabetes only (HR and (95% CI))	1 [Reference]					
Missing data imputed (HR and (95% CI))	1 [Reference]					

HR; Hazard rate, CI; Confidence interval
The models (based on model 2 in the main analysis) are adjusted for sex, study center, follow-up age, baseline educational level, smoking status, diabetes duration adherence to rMED score, total energy intake, leisure-time physical activity

sTable 2 Baseline characteristics of individuals with self-reported and confirmed diabetes		
	Self-reported diabetes only	*Confirmed diabetes*
N		
Women, (%)		
Age, years		
Education (%)		
None		
Primary school		
Technical/professional school		
Secondary school		
Higher, including university degree		
Not specified		
Smoking		
Never		
Former		
Current		
Unknown		
Body mass index, kg/m^2		
Waist circumference, cm		
Men		
Women		
Central obesity (N (%) - yes/no)		
Leisure time physical activity without bicycling, MET-h/week		
Gardening, MET-h/week		
Do-it-yourself activities, MET-h/week		
Stair-climbing, MET-h/week		
Housework activities, MET-h/week		
Walking, MET-h/week		
Vigorous physical activity, MET-h/week		
Occupational physical activity (%)		
Sedentary occupation		
Standing occupation		
Manual work		
Heavy manual work		
Non-worker		
Unknown		
Energy intake (kcal/day)		
Adherence to the relative Mediterranean diet score (%)		
Low		
Medium		
High		
Co-morbidities		
Cancer (%)		
Stroke (%)		
Myocardial infarction (%)		
Hyperlipidemia (%)		
Hypertension (%)		
† Mean (standard deviation), ‡ Median (interquartiles), MET-h/week; Metabolic quartile – hours per week		

1. Roglic G, Unwin N. Mortality attributable to diabetes: estimates for the year 2010. *Diabetes research and clinical practice.* 2010;87(1):15-19.
2. Dalan R. Sodium-Glucose Cotransporter-2 Inhibition in Type 2 Diabetes Mellitus: A Review of Large-Scale Cardiovascular Outcome Studies and Possible Mechanisms of Benefit. *Cardiology in review.* 2018;26(6):312-320.
3. Schnell O, Ryden L, Standl E, Ceriello A, Group CES. Updates on cardiovascular outcome trials in diabetes. *Cardiovascular diabetology.* 2017;16(1):128.
4. Colberg SR, Sigal RJ, Yardley JE, Riddell MC, Dunstan DW, Dempsey PC, Horton ES, Castorino K, Tate DF. Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association. *Diabetes care.* 2016;39(11):2065-2079.
5. Umpierre D, Ribeiro PA, Kramer CK, Leitao CB, Zucatti AT, Azevedo MJ, Gross JL, Ribeiro JP, Schaan BD. Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis. *JAMA : the journal of the American Medical Association.* 2011;305(17):1790-1799.
6. Boule NG, Haddad E, Kenny GP, Wells GA, Sigal RJ. Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials. *JAMA : the journal of the American Medical Association.* 2001;286(10):1218-1227.
7. Chudyk A, Petrella RJ. Effects of exercise on cardiovascular risk factors in type 2 diabetes: a meta-analysis. *Diabetes care.* 2011;34(5):1228-1237.
8. Hayashino Y, Jackson JL, Fukumori N, Nakamura F, Fukuhara S. Effects of supervised exercise on lipid profiles and blood pressure control in people with type 2 diabetes mellitus: a meta-analysis of randomized controlled trials. *Diabetes research and clinical practice.* 2012;98(3):349-360.
9. Sluik D, Buijsse B, Muckelbauer R, Kaaks R, Teucher B, Johnsen NF, Tjonneland A, Overvad K, Ostergaard JN, Amiano P, Ardanaz E, Bendinelli B, Pala V, Tumino R, Ricceri F, Mattiello A, Spijkerman AM, Monninkhof EM, May AM, Franks PW, et al. Physical Activity and Mortality in Individuals With Diabetes Mellitus: A Prospective Study and Meta-analysis. *Archives of internal medicine.* 2012;172(17):1285-1295.
10. Jakobsen I, Solomon TP, Karstoft K. The Acute Effects of Interval-Type Exercise on Glycemic Control in Type 2 Diabetes Subjects: Importance of Interval Length. A Controlled, Counterbalanced, Crossover Study. *PloS one.* 2016;11(10):e0163562.
11. Batty GD, Shipley MJ, Marmot M, Smith GD. Physical activity and cause-specific mortality in men with Type 2 diabetes/impaired glucose tolerance: evidence from the Whitehall study. *Diabetic medicine : a journal of the British Diabetic Association.* 2002;19(7):580-588.
12. Zhao G, Ford ES, Li C, Mokdad AH. Compliance with physical activity recommendations in US adults with diabetes. *Diabetic medicine : a journal of the British Diabetic Association.* 2008;25(2):221-227.
13. Morrato EH, Hill JO, Wyatt HR, Ghushchyan V, Sullivan PW. Physical activity in U.S. adults with diabetes and at risk for developing diabetes, 2003. *Diabetes care.* 2007;30(2):203-209.
14. Blond MB, Rosenkilde M, Gram AS, Tindborg M, Christensen AN, Quist JS, Stallknecht BM. How does 6 months of active bike commuting or leisure-time exercise affect insulin sensitivity, cardiorespiratory fitness and intra-abdominal fat? A randomised controlled trial in individuals with overweight and obesity. *British journal of sports medicine.* 2019.
15. Berntsen S, Malnes L, Langaker A, Bere E. Physical activity when riding an electric assisted bicycle. *The international journal of behavioral nutrition and physical activity.* 2017;14(1):55.

16. Forbes CC, Plotnikoff RC, Courneya KS, Boule NG. Physical activity preferences and type 2 diabetes: exploring demographic, cognitive, and behavioral differences. *The Diabetes educator.* 2010;36(5):801-815.
17. Wanko NS, Brazier CW, Young-Rogers D, Dunbar VG, Boyd B, George CD, Rhee MK, el-Kebbi IM, Cook CB. Exercise preferences and barriers in urban African Americans with type 2 diabetes. *The Diabetes educator.* 2004;30(3):502-513.
18. Kelly P, Kahlmeier S, Gotschi T, Orsini N, Richards J, Roberts N, Scarborough P, Foster C. Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship. *The international journal of behavioral nutrition and physical activity.* 2014;11:132.
19. Nordengen S, Andersen LB, Solbraa AK, Riiser A. Cycling and cardiovascular disease risk factors including body composition, blood lipids and cardiorespiratory fitness analysed as continuous variables: Part 2-systematic review with meta-analysis. *British journal of sports medicine.* 2019.
20. Nordengen S, Andersen LB, Solbraa AK, Riiser A. Cycling is associated with a lower incidence of cardiovascular diseases and death: Part 1 - systematic review of cohort studies with meta-analysis. *British journal of sports medicine.* 2019.
21. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, Overvad K, Tjonneland A, Clavel-Chapelon F, Thiebaut A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public health nutrition.* 2002;5(6B):1113-1124.
22. Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, van der Schouw YT, Spencer E, Moons KG, Tjonneland A, Halkjaer J, Jensen MK, Stegger J, Clavel-Chapelon F, Boutron-Ruault MC, Chajes V, Linseisen J, Kaaks R, Trichopoulou A, Trichopoulos D, et al. General and abdominal adiposity and risk of death in Europe. *The New England journal of medicine.* 2008;359(20):2105-2120.
23. Rasmussen MG, Grontved A, Blond K, Overvad K, Tjonneland A, Jensen MK, Ostergaard L. Associations between Recreational and Commuter Cycling, Changes in Cycling, and Type 2 Diabetes Risk: A Cohort Study of Danish Men and Women. *PLoS medicine.* 2016;13(7):e1002076.
24. Slimani N, Deharveng G, Unwin I, Southgate DA, Vignat J, Skeie G, Salvini S, Parpinel M, Moller A, Ireland J, Becker W, Farran A, Westenbrink S, Vasilopoulou E, Unwin J, Borgejordet A, Rohrmann S, Church S, Gnagnarella P, Casagrande C, et al. The EPIC nutrient database project (ENDB): a first attempt to standardize nutrient databases across the 10 European countries participating in the EPIC study. *European journal of clinical nutrition.* 2007;61(9):1037-1056.
25. InterAct C, Romaguera D, Guevara M, Norat T, Langenberg C, Forouhi NG, Sharp S, Slimani N, Schulze MB, Buijsse B, Buckland G, Molina-Montes E, Sanchez MJ, Moreno-Iribas MC, Bendinelli B, Grioni S, van der Schouw YT, Arriola L, Beulens JW, Boeing H, et al. Mediterranean diet and type 2 diabetes risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study: the InterAct project. *Diabetes care.* 2011;34(9):1913-1918.
26. Esposito K, Maiorino MI, Ciotola M, Di Palo C, Scognamiglio P, Gicchino M, Petrizzo M, Saccomanno F, Beneduce F, Ceriello A, Giugliano D. Effects of a Mediterranean-style diet on the need for antihyperglycemic drug therapy in patients with newly diagnosed type 2 diabetes: a randomized trial. *Annals of internal medicine.* 2009;151(5):306-314.

27. Buckland G, Agudo A, Lujan L, Jakszyn P, Bueno-de-Mesquita HB, Palli D, Boeing H, Carneiro F, Krogh V, Sacerdote C, Tumino R, Panico S, Nesi G, Manjer J, Regner S, Johansson I, Stenling R, Sanchez MJ, Dorronsoro M, Barricarte A, et al. Adherence to a Mediterranean diet and risk of gastric adenocarcinoma within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study. *The American journal of clinical nutrition.* 2010;91(2):381-390.
28. Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. *The New England journal of medicine.* 2003;348(26):2599-2608.
29. Haftenberger M, Lahmann PH, Panico S, Gonzalez CA, Seidell JC, Boeing H, Giurdanella MC, Krogh V, Bueno-de-Mesquita HB, Peeters PH, Skeie G, Hjartaker A, Rodriguez M, Quiros JR, Berglund G, Janlert U, Khaw KT, Spencer EA, Overvad K, Tjonneland A, et al. Overweight, obesity and fat distribution in 50- to 64-year-old participants in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public health nutrition.* 2002;5(6B):1147-1162.
30. Alberti KG, Zimmet P, Shaw J, Group IDFETFC. The metabolic syndrome--a new worldwide definition. *Lancet.* 2005;366(9491):1059-1062.
31. Twisk J, Proper K. Evaluation of the results of a randomized controlled trial: how to define changes between baseline and follow-up. *J Clin Epidemiol.* 2004;57(3):223-228.