Statistical Analysis Plan (SAP)

Physical activity and mortality in individuals with type 2 diabetes: Cross-country comparison in the UK Biobank and China Kadoorie Biobank prospective cohorts

Version 2.1

Version date: April 4th, 2022

SAP author: Jakob Tarp

Statistical analyst: Jakob Tarp

Statistical advisors and co-authors:

Mengyun Luo, Adriano Sanchez-Lastra, Knut Eirik Dalene, Borja Del Pozo Cruz, Mathias Ried-Larsen, Reimar W. Thomsen, Ding Ding, Ulf Ekelund

Affiliations:

JT: Department of Clinical Epidemiology, Department of Clinical Medicine, Aarhus University & University Hospital, Aarhus, Denmark

ML: Prevention Research Collaboration, Sydney School of Public Health, The University of Sydney, Camperdown, NSW Australia

AS-L: Department of Special Didactics, Faculty of Education and Sports Sciences, University of Vigo, Pontevedra, Spain

KED: Department of Sports Medicine, Norwegian School of Sports Sciences, Oslo, Norway

BDPC: Centre for Active and Healthy Ageing, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

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

RWT: Department of Clinical Epidemiology, Department of Clinical Medicine, Aarhus University & University Hospital, Aarhus, Denmark

DD: Prevention Research Collaboration, Sydney School of Public Health, The University of Sydney, Camperdown, NSW Australia

UE: Department of Sports Medicine, Norwegian School of Sports Sciences, Oslo, Norway

BACKGROUND

Engagement in regular physical activity is a cornerstone in lifestyle modifications with type 2 diabetes as it promotes uptake of circulating blood glucose independent of insulin-mediated actions and because physical activity can help maintain a favorable energy balance (1). The seminal LOOK AHEAD randomized controlled trial successfully lowered glycated hemoglobin and body weight and improved cardiorespiratory fitness through intensive lifestyle modifications (2). Yet, the intervention did not lower total mortality or cardiovascular mortality/morbidity, compared with controls offered diabetes support and education. In randomized, controlled trials, physical activity is often embedded within broader combination of lifestyle changes such as diet, weight loss, and smoking cessation (2). Therefore, the isolated effects of physical activity on mortality, and the shape and magnitude of the dose-response pattern, cannot be determined from this data. Physical activity recommendations for individuals with type 2 diabetes are currently identical to those given to the general population (3). This evidence-base is incomplete and the Guidelines Development Group for the WHO 2020 Guidelines on Physical Activity and sedentary behavior called for additional studies investigating dose-response associations with health-outcomes, across different domain of physical activity, to establish minimal effective dose and maximal safety thresholds for individuals with type 2 diabetes (3).

AIMS

To determine dose-response patterns between leisure-time physical activity and all-cause mortality in individuals with type 2 diabetes in the United Kingdom and China. A secondary aim is to study the association between domain-specific physical activity and all-cause mortality and fatal and non-fatal CVD.

METHODS

Study design and setting

The study is a nested cohort study based on the UK Biobank and China Kadoorie Biobank population-based prospective cohort studies. Both cohorts are designed to study the interrelations between environment, lifestyle, and genes, with the aims of improving the prevention, diagnosis, and treatment of chronic diseases. UK Biobank recruited a total of 502,682 participants (approximately 5.5% of 9.2 million invited) aged 37 to 82 years via 22 assessment centers across England, Wales, and Scotland between 2006 and 2010. At the assessment centers participants completed a touch-screen questionnaire, an interview with a nurse, and a wide variety of physical measurements and biological sampling (4). A subsample has attended a repeat assessment of all data collected at the baseline examination. Data has been linked with several electronic registries for ongoing follow-up on health status. Ethical approval to establish the UK Biobank cohort was obtained by the North-West Research Ethics Committee and participants gave written informed consent before data collection. China Kadoorie Biobank recruited 515,420 participants aged 30 to 79 years between 2004 and 2008 from 10 regions of mainland China (5). At the assessment centers participants completed an interviewer-

administered questionnaire, physical measurements and provided blood spot tests and non-fasting blood samples. Data has been linked with several electronic registries for ongoing follow-up on health status. China Kadoorie Biobank was approved by the Ethics Committees at Oxford University, the China National Center for Disease Control and from institutional research boards at the local Centers for Disease Control in the 10 included regions.

Study population

We identified individuals with prevalent type 2 diabetes in the UK Biobank from the baseline assessment (2006-2010) and the 1st repeat assessment (2012-2013), and in China Kadoorie Biobank from the baseline assessment (2004-2008).

UK Biobank: Prevalent type 2 diabetes is determined by the algorithm of Eastwood et al. (6) or from measured Hba1c ≥48 mmol/mol. The algorithm is based on combining information on self-reported diabetes, insulin use, age of diabetes onset, and ethnicity obtained from a questionnaire in addition to self-reported diabetes, self-reported use of medications (Table 1), and age at diabetes diagnosis obtained from an interview with a trained nurse. Both 'probable' and 'possible' type 2 diabetes from the algorithm are included as type 2 diabetes cases. Eastwood et al. (6) showed that 71% of individuals flagged as 'probable' type 2 diabetes and 80% of 'possible' type 2 diabetes had a diabetes related hospital admission code prior to the UKB baseline assessment. Type I diabetes is removed from the sample by combining information on insulin use, time from diagnosis to initiation of insulin use, and age of diagnosis (6). These criteria identify 29,236 individuals with type 2 diabetes.

China Kadoorie Biobank: prevalent type 2 diabetes is based on self-reported current diabetes with a diagnosis age above 30 years, a random plasma blood glucose ≥11.1 mmol/L, or fasting plasma blood glucose ≥7.0 mmol/L. These criteria identify 30,300 individuals with type 2 diabetes.

Table 1. List of medications used to determine diabetes status in UK Biobank;

	Brand or product as listed in UK Biobank showcase	UK Biobank code
Insulin medication		
	insulin product	1140883066
Metformin medication		
	metformin	1140884600
	glucophage 500mg tablet	1140874686
	rosiglitazone 1mg / metformin 500mg tablet	1141189090
Non-metformin diabetes medications		
	troglitazone	1141153254
	pioglitazone	1141171646
	rosiglitazone	1141177600
	acetohexamide	1140857584
	chlorpropamide	1140874706
	tolazamide	1140874664
	tolbutamide	1140874674

	glibornuride	1140857494
	gliclazide	1140874744
	glipizide	1140874646
	glipizide product	1141157284
	gliquidone	1140874658
	glimepiride	1141152590
	repaglinide	1141168660
	nateglinide	1141173882
	amaryl 1mg tablet	1141156984
	daonil 5mg tablet	1140874724
	semi-daonil 2.5mg tablet	1140874726
	diamicron 80mg tablet	1140874746
	glibenese 5mg tablet	1140874650
	minodiab 2.5mg tablet	1140874652
	repaglinide	1141168660
	nateglinide	1141173882
	starlix 60mg tablet	1141173786
	acarbose	1140868902
	glucobay 50mg tablet	1140868908
	avandia 4mg tablet	1141177606
	actos 15mg tablet	1141171652

Outcomes

The primary outcome is risk of all-cause mortality during follow-up. Secondary outcomes are cardiovascular mortality and major adverse cardiovascular events (MACE, UK Biobank only), coded according to the *International Classification of Diseases, Injuries, and Causes of Death, Tenth Revision.* Cardiovascular mortality will be defined as the primary cause of death from any endpoint from I00 to I99. MACE will be defined as incident episodes of ischemic heart disease (I20-I25) or stroke (I60, I61, I63, and I64) in addition to cardiovascular mortality.

Mortality and morbidity status are obtained through ongoing follow-up by linkage with national registries as provided by the UK Biobank and China Kadoorie Biobank. We will use the longest available follow-up. Current all-cause mortality censoring date is 30 September 2021 (England) in UK Biobank and 31 December 2016 in China Kadoorie Biobank.

Exposure data sources

Physical activity exposures are obtained through cohort-specific lifestyle questionnaires comprising information about lifestyle and demographic factors (https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/TouchscreenQuestionsMainFinal.pdf and https://www.ckbiobank.org/site/binaries/content/assets/resources/pdf/qs_baseline-final-from10june2004.pdf). Questionnaires were completed at study centres along with clinical measurements and, in the UK Biobank, a structured interview conducted by a nurse. For participants satisfying all inclusion criteria (detailed later), data obtained at the first visit will be used.

Assessment of physical activity domains

Our primary exposure variable is leisure-time physical activity (LTPA). We will also consider physical activity during transportation to/from work and during occupation. Leisure-time physical activity (LTPA) is obtained by summarizing information on frequency and duration of activities. Information on physical activity during transportation to work includes transportation mode, distance, time and frequency. Occupational physical activity will be categorized according to the level of standing, walking and heavy manual work performed. The following four categories of will be created as primary LTPA exposure variables (assuming moderate as 3.0 METs);

1. Zero LTPA - reference
2. >0-7.49 MET-hrs/week (below recommendation)
3. 7.5-14.9 MET-hrs/week (recommended levels)
4. ≥15 MET-hrs/week (above recommendation)

UK Biobank: we will use information on walking for pleasure, light DIY, heavy DIY, strenuous sport, and 'other exercises' to calculate total LTPA in MET-hrs/week. We will assign the following MET-values; walking: 3.3 METs, light DIY: 2.25 METs, heavy DIY: 4.5 METs, strenuous sports: 8.0 METs, and 'other exercises': 4.5 METs (7).

Physical activity during transportation to work will be categorized into modes as; passive (car/motor-vehicle/public transportation), walking, or cycling. Participants with mixed modes including passive and walking or passive and cycling will be categorized according to the active mode. Participants with walking and cycling will be categorized as cycling. Occupational physical activity will be based on cross-tabulations of standing/walking (none, sometimes, usually, always) and heavy manual labor (none, sometimes, usually, always). We will create three categories as;

1. No walking/standing, no heavy manual work
2. Sometimes/usual/always walking/standing, no heavy manual work
3. Sometimes/usual/always heavy manual work (irrespective of walking/standing)

China Kadoorie Biobank: participants are asked to report on their *main* type of exercise with the options Taichi/Qiqong, jogging/aerobic exercise, ball games, walking, swimming, other, hence, LTPA in MET-hrs/week refers to one type of activity only. We will assign the following MET-values; Taichi/Qiqong/leisure walking: 3.3 METs (8), jogging/aerobic exercise: 7.4 METs (8), ball games: 5.5 METs (8), brisk walking, gymnastics, folk dancing: 4.2 (8), swimming: 7.2 METs (8), other: 4.5 METs (7).

Physical activity during transportation to work will be categorized into modes as; passive (bus/car/ferry/train/motor bike), walking, or cycling. MET-hours/week from walking or cycling for transportation will be used to describe the dose-response patterns between active transportation and mortality, assigning walking for transportation as 4.0 METs and cycling as 6.8 METs (8). Occupational physical activity will be obtained directly from the questionnaire as;

1. Mainly sedentary
2. Standing occupation
3. Manual work + heavy manual work.

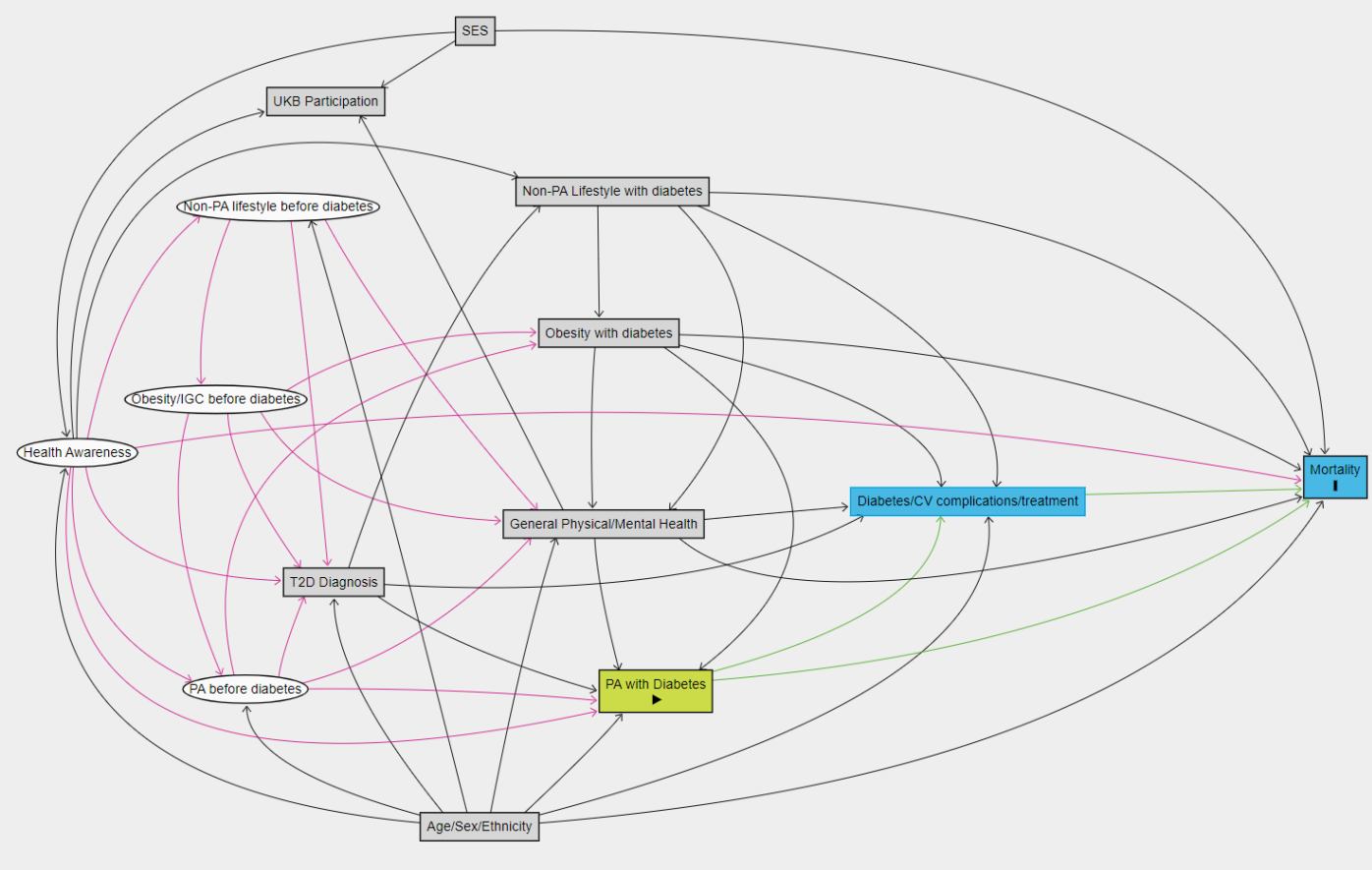
Quality control of physical activity data: Individuals reporting implausible high values will be excluded. In UK Biobank, implausible high is defined as >24 hrs/day of physical activity (from IPAQ which asks about activity during all waking hours), TV-viewing, leisure time pc usage, and sleep. Sleep-time was assumed 8 hrs/day if missing. For transportation physical activity, individuals with <1 commute/week will be categorized as working from home, >10 outwards journeys/week are deemed implausible and

their transportation mode is set to missing. In China Kadoorie Biobank, implausible high values are defined as >24 hrs/day of physical activity, housework, watching TV/reading, and sleep.

Other variables

Variables used to mitigate confounding were selected based on a directed acyclic graph (DAG) (Figure 1, example for UK Biobank) and cohort-specific availability of the relevant variables.

Figure 1. Directed acyclic graph for UK Biobank



SES; socioeconomic status, PA; physical activity, IGT; impaired glucose control, UKB; UK Biobank, CV; cardiovascular

The directed acyclic graph (DAG) depicts how lifestyle factors impacts the risk of developing type 2 diabetes as well as how lifestyle factors track over time (physical activity before type 2 diabetes predicts physical activity with type 2 diabetes). We have not included each physical activity domain separately, but each domain will be adjusted for physical activity in other domains.

A particular concern for causal inference in this context is the potential reciprocal association between physical activity and current comorbidities (General Physical/Mental Health + Diabetes/CV complications/treatment). As we do not follow individuals from before their diabetes diagnosis, we do not know the causal sequence of factors observed at study baseline were both diabetes status, physical activity, other lifestyle factors, and current medications are assessed. It is possible that physical activity may promote better health, but it is also possible that poor health may cause low physical activity. Therefore, the primary analysis will be based on the following;

- Remove from the analysis those individuals with comorbidities/physical limitations where there is a high risk of limitations to be physically active.
- Multivariable-adjustment for duration of type 2 diabetes because complications tend to increase with time. We consider this appropriate as physical activity does not affect duration of diabetes, but duration of diabetes may impact physical activity levels.
- No adjustment for use of glucose-, blood pressure-, or cholesterol lowering drugs because physical activity may affect use of these drugs. They are therefore potential mediators of the effects of physical activity on mortality.
- Under this DAG, the physical activity-mortality association is confounded by a backdoor path from Diabetes PA-Health Awareness-Mortality. To address this confounding, we have included an indicator for family history of major non-communicable diseases as a marker for health awareness (which may also be a marker of genetic susceptibility to those conditions).

In addition to the nodes included in the DAG above, the DAG for China Kadoorie Biobank also included experiences of famine with weight loss. Famine with weight loss could impact long-term health and make individuals more susceptible to metabolic diseases when resources are no longer limited (9).

List of confounding variables identified from the DAG and their operationalization

UK Biobank

Acquired through a combination of self-report from an electronic questionnaire and reported during interview with a nurse

Age: used as time-scale

Sex: male/female

Ethnicity: European, South Asian, African Caribbean, other

Non-PA lifestyle:

Smoking: never, previous, current

Diet: meeting 2 of 3 dietary targets (1) ≤3 weekly servings of red meat and ≤1 serving/week of processed meat, 2) ≥2 servings/week of fish including one with oily fish, and 3) ≥5 servings/day of fruits or vegetables

Alcohol: never, previous, current <3 times/week, current ≥3 times/week

BMI: continuous

SES:

Living with partner (yes/no)

Education: No qualifications, Other qualifications than college/university degree, University degree

Townsend index: continuous

Employment: unemployed, employed, retired

General Physical/Mental health: depression (yes/no), loneliness (yes/no), years since T2D diagnosis (calculated as assessment data minus self-reported age of diabetes diagnosis. Individuals with T2 diabetes flagged from measured Hba1c are assigned 0 years since diagnosis).

T2D diagnosis: T2D status ascertained from self-reported diabetes ("Has a doctor ever told you that you have diabetes?") or use of diabetes medication ('Medication for cholesterol, blood pressure or diabetes' [UK Biobank Fields 6177 or 6153).

Health awareness: family history of CVD, cancer or diabetes in biological or adoptive parents or siblings. This variable may also capture genetic risk of these conditions.

China Kadoorie Biobank

Age: used as time-scale

Sex: male/female

Non-PA lifestyle:

Smoking: never, occasional, former regular, current

Diet: Intake of fresh fruit and meat (<4 times/week vs ≥4 times/week)

Alcohol: Never/occasionally, former weekly, current, <3 days/week, current ≥3 days/week

BMI: continuous

SES:

Education: Primary school or below, middle school, high school or higher

Employment: Unemployed, employed, retired

Marital status: Married, separated/divorced/widowed/never married

Household income: <10,000 yuan/y, 10,000-19,999 yuan/y, 20,000-34,999 yuan/y, >=35,000 yuan/y

Health care cover: yes/no

General Physical/Mental health:

Experience of food shortage with weight loss: yes/no

Major depressive episode in last 12 months: yes/no

Years since T2D diagnosis (calculated as assessment data minus self-reported age of diabetes diagnosis. Individuals with T2 diabetes flagged from measured blood glucose are assigned 0 years since diagnosis).

T2D diagnosis: T2D status ascertained from self-reported diabetes ("Has a doctor ever told you that you have diabetes?")

Health awareness: family history of CVD, cancer or diabetes in biological parents or siblings. This variable may also capture genetic risk of these conditions.

Statistical analysis

Descriptive data for continuous variables is described as means with standard deviations. Categorical or categorized variables are presented as proportions within strata of LTPA. Descriptive data across strata of transportation and occupational physical activity are presented in a supplement.

Inclusion criteria: The target population is individuals with type 2 diabetes. Because several conditions connected to type 2 diabetes may obstruct or prevent participation in physical activity, there is a high risk of reverse causation bias, i.e. physical activity-mortality associations are biased away from the null because these conditions prevent physical activity and increase risk of death. Thus, several restrictions will be applied to reduce the potential influence of reverse causation bias. More lenient approaches will be pursued in sensitivity analyses to explore generalizability of results to individuals with prevalent CVD or cancer.

Full information on covariates is needed for inclusion under the assumption that data is missing at random. We will include a standard flowchart detailing number of exclusions with reasons, including; missing exposure data, missing other data, excluded due to pre-existing conditions.

UK Biobank: The primary analysis is based on participants without a history of CVD (myocardial infarction, stroke, ischemic stroke, intracerebral or subarachonid haemorrhage, angina, or heart failure), cancer (excluding non-melanoma skin cancers), chronic degenerative neurological problems, chronic widespread pain, chronic respiratory diseases, chronic immunological/systemic diseases, renal/kidney failure, liver failure/cirrhosis, psychological /psychiatric problems, substance abuse/dependency, anorexia/bulimia/other eating disorders, or COPD. We are also excluding those who were pregnant, underweight, unable to walk, living in a care home or requiring attendance, disability or mobility allowance. These data are acquired using a combination of self-report (baseline questionnaire), clinical measurement, interview with a nurse, and electronic data linkage.

China Kadoorie Biobank: Restricted to individuals not currently working in agricultural or related occupations because these individuals were asked separate questions on occupational and transport-related physical activity during the baseline assessment. The primary analysis is based on participants without a history of CVD (coronary heart disease, stroke, transient ischaemic attack), rheumatic heart disease, emphysema/bronchitis, or cancer. We also excluded individuals currently on treatment for tuberculosis, cirrhosis/chronic hepatitis, kidney disease, rheumatic arthritis, psychiatric disorders, or neurasthenia and those with generalized anxiety disorder, continuous pain/discomfort, or underweight. These data are obtained from the baseline questionnaire or clinical measurements.

Primary and secondary outcomes: The risk of all-cause mortality (primary outcome), cardiovascular mortality, and MACE (secondary outcomes), are expressed as hazard ratios (HRs) from Cox proportional hazards regression models with age as the time scale and 95% confidence intervals (CI). The primary exposure is categories of leisure-time physical activity. Associations for transport and occupational PA will be analyzed as secondary exposures among participants in employment/self-employed at the time of baseline assessment, following the same procedures as for leisure-time PA. Each physical activity domain will be adjusted for other physical activity domains (i.e. leisure-time PA adjusted for transportation and occupation and vice versa).

Analyses will be corrected for delayed entry. Participants are considered ‘at-risk’ from 3 years after attending a study centre. The data is right-censored at age of death, emigration, loss to follow-up, withdrawal from the study or end of observation time, whichever occurs first. For CVD mortality and

MACE, competing risks will be handled using Fine-Gray models with death from other causes as a competing event. Statistical significance is set at $\alpha = 0.05$ (two-sided).

Analysis will be performed separately for the UK Biobank and China Kadoorie Biobank cohorts, based on four models (identical for all-cause mortality, CVD mortality, and MACE):

1) A crude model (model 1), will be fitted with categories of leisure time, transport and occupational PA, separately, as exposures and adjusted for sex and age.
2) Model 2, Model 1 + covariates listed in section 'other variables' + mutual adjustment for other physical activity domains.
3) Model 3, Model 2 but leaving out adjustment for BMI to examine potential mediation through control of body mass with physical activity
4) Model 4, Model 2 + adjustment for glucose-lowering drugs and use of other pharmaceutical treatment.

Continuous dose-response: All continuous dose-response analyses will be winsorized at the 95th percentile of the exposure distribution.

The shape of the dose-response association between MET-hrs/week of LTPA and all-cause mortality, CVD-mortality and MACE will be explored using restricted cubic splines with 3 knots placed at the 10th, 50th and 90th percentiles of exposure distribution among individuals with non-zero LTPA. The reference category will be zero LTPA.

Effect modification: Effect modification by sex (male/female), T2D diagnosis by medical doctor or on treatment with blood-glucose lowering drug(s) (yes/no), age (<60 vs ≥ 60 years of age), diabetes duration (<5 vs ≥ 5 years) will be evaluated statistically using the likelihood-ratio test (physical activity exposure-by-effect modifier interaction term + main effects compared to model 2 including the main effects only).

Model assumptions: Departure from proportional hazards assumption will be evaluated by tests and graphs of Shoenfeld residuals. If the model variables do not meet the proportional hazards assumption, a stratified Cox model will be employed for the relevant variables. If the violation of the proportional hazards 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: The following sensitivity analyses are planned for the primary and secondary outcomes.

1. To examine robustness against residual confounding (all based on model 2)
    a. Excluding ever smokers
    b. Adjustment for total energy intake, dietary intake of fibers, and the ratio of polyunsaturated to saturated fat from online 24-hr dietary recalls performed between April 2009 and June 2012 (UKB, not available in all participants).
2. To examine generalizability of results
    a. Repeat model 2 with inclusion of individuals with prevalent CVD or cancer
    b. Repeat model 2 excluding individuals classified as 'possibleT2D' (UK Biobank only)

Implementation of the SAP

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 commencement of the statistical analysis
2. Statistical analyses are performed (JT)
3. A preliminary report is drafted (JT) and circulated among co-authors
4. The report is revised and circulated among co-authors for further comments and final approval
5. When agreement about interpretation and conclusion is reached, the report is submitted to a scientific journal (Priority: 1) JAMA, 2) Lancet Diabetes & Endocrinology, 3), BMJ, 4) Annals of Internal Medicine, 5) Diabetes Care, 6) Diabetologia).

Anticipated outline of the study report (manuscript)

Table 1. Descriptive characteristics.

	Ref (0 leisure time PA)	Ref-7.49 MET-hrs/week	7.5-14.9 MET-hrs/week	≥15 MET-hrs/week
	No LTPA	*Below recommendation*	*Recommended levels*	Above recommendation
UK Biobank				
N (% Women)				
Age (years), mean (SD)				
Body mass index (kg/m^2), mean (SD)				
LTPA (MET-hours/wk), mean (SD)				
Participation in sports, No. (%)				
Duration of diabetes (years), mean (SD)				
Education, No. (%)				
No qualifications				
Other qualifications than college/university degree				
College/University degree				
Smoking, No. (%)				
Never				
Former				
Current				
Family history of CVD, cancer or diabetes (yes), No. (%)				
Statins (yes), No. (%)				
Use of blood-pressure lowering drugs, No (%)				
0				
1				
2				
3 or more				
Doctor diagnosis or on treatment for type 2 diabetes (yes), No. (%)*				
Use of blood-glucose lowering drugs, No (%)**				
None				
Insulin only				
Non-insulin only				

Insulin and non-insulin				
China Kadoorie Biobank				
% Women				
Age				
Body mass index (kg/m^2), mean (SD)				
LTPA (MET-hours/wk), mean (SD)				
Participation in heavy physical activity/exercise, No. (%)**				
Duration of diabetes (years), mean (SD)				
Education, No. (%)				
No school or primary school				
Middle school				
High school or higher				
Smoking, No. (%)				
Never				
Occasional				
Former				
Current				
Family history of CVD, cancer or diabetes (yes), No. (%)				
Statins (yes), No. (%)				
Use of blood-pressure lowering drugs, No (%)				
0				
1				
2 or more				
Doctor diagnosis or on treatment for type 2 diabetes (yes), No. (%)*				
Use of blood-glucose lowering drugs, No (%)**				
None				
Insulin only				
Chlorpropamide or metformin only				
Insulin and Chlorpropamide or metformin				

*Individuals with type 2 diabetes identified from self-report or use of glucose-lowering drugs Eastwood et al., 2016, UKB Biobank) or from self-reported diagnosis of diabetes from a doctor (China Kadoorie Biobank).

**Individuals identified with type 2 diabetes solely from measured Hba1c (UK Biobank) or random blood glucose (China Kadoorie Biobank) are not included in the denominator.

***Includes work, transportation, domestic and leisure activities

Table 2. Leisure Time Physical Activity, all-cause and cardiovascular mortality

	Ref (0 leisure time PA)	Ref-7.49 MET-hrs/week	7.5-14.9 MET-hrs/week	≥15 MET-hrs/week
	No LTPA	*Below recommendation*	*Recommended levels*	Above recommendation
All-Cause Mortality				
UK Biobank				
N = , *deaths* =				
Crude incidence rate/1000 person-years				
Model 1 (HR [95%CI])	1 [reference]			
Model 2 (HR [95%CI])	1 [reference]			
Model 3 (HR [95%CI])	1 [reference]			
Model 4 (HR [95%CI])	1 [reference]			
China Kadoorie Biobank				
N = , *deaths* =				
Crude incidence rate/1000 person-years				
Model 1 (HR [95%CI])	1 [reference]			
Model 2 (HR [95%CI])	1 [reference]			
Model 3 (HR [95%CI])	1 [reference]			
Model 4 (HR [95%CI])	1 [reference]			
CVD Mortality				
UK Biobank				
N =, *deaths* =				
Crude incidence rate/1000 person-years				
Model 1 (sHR [95%CI])	1 [reference]			
Model 2 (sHR [95%CI])	1 [reference]			
Model 3 (sHR [95%CI])	1 [reference]			
Model 4 (sHR [95%CI])	1 [reference]			
China Kadoorie Biobank				
N =, *deaths* =				
Crude incidence rate/1000 person-years				
Model 1 (sHR [95%CI])	1 [reference]			
Model 2 (sHR [95%CI])	1 [reference]			

Model 3 (sHR [95%CI])	1 [reference]			
Model 4 (sHR [95%CI])	1 [reference]			

Figure 1. Continuous dose-response curves LTPA (MET-hrs/week) and all-cause mortality. Hazard ratios and 95% confidence intervals.

Figure 2. Continuous dose-response curves LTPA (MET-hrs/week) and CVD-mortality. Hazard ratios and 95% confidence intervals.

Supplementary Material

eTable 1. Descriptive characteristics UK Biobank, detailed

	Ref (0 leisure time PA)	Ref-7.49 MET-hrs/week	7.5-14.9 MET-hrs/week	≥15 MET-hrs/week
	No LTPA	*Below recommendation*	*Recommended levels*	Above recommendation
UK Biobank				
N (% Women)				
Age (years), mean (SD)				
Body mass index (kg/m^2), mean (SD)				
LTPA (MET-hours/wk), mean (SD)				
Participation in sports (yes), No. (%)				
Duration of diabetes (years), mean (SD)				
Education, No. (%)				
No qualifications				
Other qualifications than college/university degree				
College/University degree				
Townsend Index, mean (SD)				
Marital Status (living with partner), No. (%)				
Ethnicity, No. (%)				
European				

South Asian				
African Caribbean				
Other				
Occupational Physical activity, No. (%)				
Sedentary				
Some standing, No heavy				
Heavy manual work				
Not in employment				
Retired				
Transportation, No. (%)				
Passive				
Walking				
Cycling				
Working from home				
Not in employment				
Smoking, No. (%)				
Never				
Former				
Current				
Alcohol intake, No. (%)				
Never				
Former				
Current,<3 times/week				
Current, ≥3 times/week				
Diet Pattern (meeting ≥2 recommendations), No. (%)				
Body mass index categories (kg/m^2), No. (%)				
18.5-25				
25-30				
30-35				
≥35				
Family history of CVD, cancer or diabetes (yes), No. (%)				
Depression (yes), No. (%)				
Loneliness (yes), No. (%)				
Beta-blockers (yes), No. (%)				
Calcium-channel blockers (yes), No. (%)				

ACE-inhibitors (yes), No. (%)				
Thiazide diuretics (yes), No. (%)				
Loop diuretics (yes), No. (%)				
Potassium-sparring diuretics (yes), No. (%)				
Statins (yes), No. (%)				
Hba1c (mmol/mol), mean (SD)*				
Doctor diagnosis or on treatment for type 2 diabetes (yes), No. (%)**				
Insulin medication (yes), No. (%)***				
Metformin medication (yes), No. (%)***				
Other glucose-lowering medication (yes), No. (%)***				
Any non-insulin glucose-lowering drug (yes), No. (%)***				

*n=

**Individuals with type 2 diabetes identified from self-report or use of glucose-lowering drugs (Eastwood et al., 2016)

***Reported use at nurse interview, individuals identified with type 2 diabetes solely from measured Hba1c are not included in the percentage denominator.

eTable 2. Descriptive characteristics China Kadoorie Biobank, detailed

	Ref (0 leisure time PA)	Ref-7.49 MET-hrs/week	7.5-14.9 MET-hrs/week	≥15 MET-hrs/week
	No LTPA	*Below recommendation*	*Recommended levels*	Above recommendation
China Kadoorie Biobank				
% Women				
Age				
Body mass index (kg/m^2), mean (SD)				
LTPA (MET-hours/wk), mean (SD)				
Participation in heavy physical activity/exercise, No. (%)*				
Duration of diabetes (years), mean (SD)				
Education, No. (%)				
No school or primary school				
Middle school				
High school or higher				
Household income, No. (%)				
<10,000 yuan/year				
10,000-19,999 yuan/year				
20,000-34,999 yuan/year				

≥35,000 yuan/year				
Marital Status (Married), No. (%)				
Health cover (yes), No. (%)				
Occupational Physical activity, No. (%)				
Sedentary				
Standing				
Manual work				
Not in employment				
Retired				
Transportation, No. (%)				
Passive				
Walking				
Cycling				
Working from home				
Not in employed/retired				
Smoking, No. (%)				
Never				
Occasional				
Former				
Current				
Alcohol intake, No. (%)				
Never/rarely				
Former weekly				
Current, <3 times/week				
Current, ≥3 times/week				
Regular fruit consumption (≥4 days/week), No. (%)				
Regular meat consumption (≥4 days/week), No. (%)				
Body mass index categories (kg/m^2), No. (%)				
18.5-25				
25-30				
30-35				
≥35				
Family history of CVD, cancer or diabetes (yes), No. (%)				
Experienced food shortage with weight loss (yes), No. (%)				
Major depression in last 12 months (yes), No. (%)				

Beta-blockers (yes), No. (%)				
ACE-inhibitors (yes), No. (%)				
Diuretics (yes), No. (%)				
Calcium-antagonists (yes), No. (%)				
Statins (yes), No. (%)				
Random glucose (mmol/L), mean (SD)**				
Doctor diagnosis of type 2 diabetes (yes), No. (%)***				
Insulin medication (yes), No. (%)****				
Chlorpropamide or metformin medication (yes), No. (%)****				

*Includes work, transportation, domestic and leisure activities

**n=

***Self-reported diagnosis of diabetes from a doctor

****The denominator in the percentage is individuals with a doctor diagnosis of T2D.

eTable 3. Leisure Time Physical Activity and risk of major adverse cardiovascular events (MACE)

	Ref (0 leisure time PA)	Ref-7.49 MET-hrs/week	7.5-14.9 MET-hrs/week	≥15 MET-hrs/week
	No LTPA	*Below recommendation*	*Recommended levels*	Above recommendation
MACE				
UK Biobank				
N = , *cases* =				
Crude incidence rate/1000 person-years				
Model 1 (sHR [95%CI])	1 [reference]			
Model 2 (sHR [95%CI])	1 [reference]			
Model 3 (sHR [95%CI])	1 [reference]			
Model 4 (sHR [95%CI])	1 [reference]			

eTable 4. Descriptive characteristics by transportation mode

eTable 5. Descriptive characteristics by occupational physical activity

eTable 6. Physical Activity for Transportation to work

	Ref (Passive transportation)	Walking	Cycling
All-cause mortality			
UK Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (HR [95%CI])	1 [reference]		
Model 2 (HR [95%CI])	1 [reference]		
Model 3 (HR [95%CI])	1 [reference]		
Model 4 (HR [95%CI])	1 [reference]		
China Kadoorie Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (HR [95%CI])	1 [reference]		
Model 2 (HR [95%CI])	1 [reference]		
Model 3 (HR [95%CI])	1 [reference]		
Model 4 (HR [95%CI])	1 [reference]		
CVD mortality			
UK Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (sHR [95%CI])	1 [reference]		
Model 2 (sHR [95%CI])	1 [reference]		
Model 3 (sHR [95%CI])	1 [reference]		
Model 4 (sHR [95%CI])	1 [reference]		
China Kadoorie Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (sHR [95%CI])	1 [reference]		
Model 2 (sHR [95%CI])	1 [reference]		
Model 3 (sHR [95%CI])	1 [reference]		
Model 4 (sHR [95%CI])	1 [reference]		

MACE			
UK Biobank			
N = , cases =			
Crude incidence rate/1000 person-years			
Model 1 (sHR [95%CI])	1 [reference]		
Model 2 (sHR [95%CI])	1 [reference]		
Model 3 (sHR [95%CI])	1 [reference]		
Model 4 (sHR [95%CI])	1 [reference]		

In UK Biobank, participants using mixed modes are coded according to their active transportation. Participants reporting any cycling are coded as cyclist. Only the primary mode of transportation is reported in China Kadoorie Biobank.

eTable 7. Occupational Physical Activity

	Sedentary	Standing/walking	Manual work*
All-cause mortality			
UK Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (HR [95%CI])	1 [reference]		
Model 2 (HR [95%CI])	1 [reference]		
Model 3 (HR [95%CI])	1 [reference]		
Model 4 (HR [95%CI])	1 [reference]		
China Kadoorie Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (HR [95%CI])	1 [reference]		
Model 2 (HR [95%CI])	1 [reference]		
Model 3 (HR [95%CI])	1 [reference]		
Model 4 (HR [95%CI])	1 [reference]		
CVD mortality			
UK Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (sHR [95%CI])	1 [reference]		

Model 2 (sHR [95%CI])	1 [reference]		
Model 3 (sHR [95%CI])	1 [reference]		
Model 4 (sHR [95%CI])	1 [reference]		
China Kadoorie Biobank			
N = , *deaths* =			
Crude incidence rate/1000 person-years			
Model 1 (sHR [95%CI])	1 [reference]		
Model 2 (sHR [95%CI])	1 [reference]		
Model 3 (sHR [95%CI])	1 [reference]		
Model 4 (sHR [95%CI])	1 [reference]		
MACE			
UK Biobank			
N = , cases =			
Crude incidence rate/1000 person-years			
Model 1 (sHR [95%CI])	1 [reference]		
Model 2 (sHR [95%CI])	1 [reference]		
Model 3 (sHR [95%CI])	1 [reference]		
Model 4 (sHR [95%CI])	1 [reference]		

*Derived from a combination of walking/standing and heavy manual work in UK Biobank. Combining ‘manual work’ and ‘heavy manual work’ in China Kadoorie Biobank

eTable 8. Sensitivity analyses, all-cause mortality, CVD-mortality and MACE based on model 2.

eFigure 1. Participant flowcharts

eFigure 2. Continuous dose-response curves LTPA (MET-hrs/week) and risk of MACE. Hazard ratios and 95% confidence intervals.

1. Colberg SR, Sigal RJ, Yardley JE, Riddell MC, Dunstan DW, Dempsey PC, et al. Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association. Diabetes care. 2016;39(11):2065-79.
2. Look Ahead Research Group. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. The New England journal of medicine. 2013;369(2):145-54.
3. DiPietro L, Al-Ansari SS, Biddle SJH, Borodulin K, Bull FC, Buman MP, et al. Advancing the global physical activity agenda: recommendations for future research by the 2020 WHO physical activity and sedentary behavior guidelines development group. Int J Behav Nutr Phys Act. 2020;17(1):143.
4. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.
5. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol. 2011;40(6):1652-66.
6. Eastwood SV, Mathur R, Atkinson M, Brophy S, Sudlow C, Flaig R, et al. Algorithms for the Capture and Adjudication of Prevalent and Incident Diabetes in UK Biobank. PLoS One. 2016;11(9):e0162388.
7. Pearce M, Strain T, Kim Y, Sharp SJ, Westgate K, Wijndaele K, et al. Estimating physical activity from self-reported behaviours in large-scale population studies using network harmonisation: findings from UK Biobank and associations with disease outcomes. Int J Behav Nutr Phys Act. 2020;17(1):40.
8. Bennett DA, Du H, Clarke R, Guo Y, Yang L, Bian Z, et al. Association of Physical Activity With Risk of Major Cardiovascular Diseases in Chinese Men and Women. JAMA Cardiol. 2017;2(12):1349-58.
9. Roseboom TJ. Epidemiological evidence for the developmental origins of health and disease: effects of prenatal undernutrition in humans. J Endocrinol. 2019;242(1):T135-T44.