








OPEN ACCESS

# Accelerometer-measured intensity-specific physical activity, genetic risk and incident type 2 diabetes: a prospective cohort study

Mengyun Luo ,<sup>1,2,3</sup> Chenhao Yu ,<sup>3</sup> Borja Del Pozo Cruz ,<sup>4,5,6</sup>  
Liangkai Chen ,<sup>7</sup> Ding Ding ,<sup>1,2</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bjsports-2022-106653>).

For numbered affiliations see end of article.

## Correspondence to

Dr Ding Ding; [melody.ding@sydney.edu.au](mailto:melody.ding@sydney.edu.au)  
Dr Liangkai Chen; [clk@hust.edu.cn](mailto:clk@hust.edu.cn)

LC and DD contributed equally.

Accepted 26 April 2023

## ABSTRACT

**Objective** Although 30 min/day of moderate-intensity physical activity is suggested for preventing type 2 diabetes (T2D), the current recommendations exclusively rely on self-reports and rarely consider the genetic risk. We examined the prospective dose-response relationships between total/intensity-specific physical activity and incident T2D accounting for and stratified by different levels of genetic risk.

**Methods** This prospective cohort study was based on 59 325 participants in the UK Biobank (mean age=61.1 years in 2013–2015). Total/intensity-specific physical activity was collected using accelerometers and linked to national registries until 30 September 2021. We examined the shape of the dose-response association between physical activity and T2D incidence using restricted cubic splines adjusted for and stratified by a polygenic risk score (based on 424 selected single nucleotide polymorphisms) using Cox proportional hazards models.

**Results** During a median follow-up of 6.8 years, there was a strong linear dose-response association between moderate-to-vigorous-intensity physical activity (MVPA) and incident T2D, even after adjusting for genetic risk. Compared with the least active participants, the HRs (95% CI) for higher levels of MVPA were: 0.63 (0.53 to 0.75) for 5.3–25.9 min/day, 0.41 (0.34 to 0.51) for 26.0–68.4 min/day and 0.26 (0.18 to 0.38) for >68.4 min/day. While no significant multiplicative interaction between physical activity measures and genetic risk was found, we found a significant additive interaction between MVPA and genetic risk score, suggesting larger absolute risk differences by MVPA levels among those with higher genetic risk.

**Conclusion** Participation in physical activity, particularly MVPA, should be promoted especially in those with high genetic risk of T2D. There may be no minimal or maximal threshold for the benefits. This finding can inform future guidelines development and interventions to prevent T2D.

## INTRODUCTION

Diabetes remains, to date, a global public health concern.<sup>1</sup> In 2021, there were 537 million adults living with diabetes worldwide, with a prevalence of 10.5%. The direct health expenditures associated with diabetes during the same year ascended to US\$966 billion.<sup>2</sup> Type 2 diabetes (T2D) accounts for approximately 90% of all types of diabetes mellitus.

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The WHO recommends at least 30 min of moderate-intensity physical activity per day for the prevention of type 2 diabetes (T2D). However, such recommendations almost exclusively rely on studies using self-reported measures of physical activity, which is subject to bias. The dose-response association between device-measured physical activity, particularly light-intensity physical activity, and T2D is rarely explored. Furthermore, while genetics play a major role in the onset of T2D, the interaction between genetic risk and physical activity on T2D is less understood.

## WHAT THIS STUDY ADDS

⇒ First, there was a strong linear dose-response inverse association between accelerometer-measured physical activity and T2D even when adjusting for the genetic risk. The magnitude of the association appeared larger than in studies using self-reported physical activity. Second, this association was stronger for moderate-to-vigorous-intensity physical activity (MVPA), while the association with light-intensity physical activity was weak and inconsistent. Third, we found no multiplicative interaction between physical activity and genetic risk but a significant additive interaction with MVPA. The association between total physical activity/MVPA and T2D was similar across genetic risk strata but the absolute risk reduction from MVPA was the largest for those with high genetic risk.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ In contrast to existing evidence primarily based on self-reported physical activity, our findings imply that there may be no threshold for the 'optimal amount' of physical activity for T2D prevention and the association may be stronger than previously estimated. These findings can inform future guidelines development and interventions to prevent T2D. Physical activity, specifically MVPA, is beneficial, especially in those with high genetic risk, and should be promoted as a priority strategy for T2D prevention.



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Luo M, Yu C, Del Pozo Cruz B, et al. *Br J Sports Med* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bjsports-2022-106653

Physical activity is a first-line strategy for the prevention and management of T2D.<sup>3–4</sup> Health authorities recommended at least 30 min of moderate-intensity physical activity (MPA) on most days of the week or 150 min accumulated per week to prevent the onset of T2D.<sup>1–4</sup> However, such recommendations almost exclusively rely on studies using self-reported measures of physical activity, which is subject to bias. In addition, the role of low-intensity physical activity (LPA) on T2D remains unknown. As a result, the 2018 Physical Activity Guidelines Advisory Committee Scientific Report called for more research on device-based physical activity and the potential health benefits of LPA.<sup>5</sup>

Although the important role genetics plays in the aetiology of T2D is well established, evidence on whether the genetic risk of T2D could be attenuated by physical activity, and whether the protective effects of physical activity on T2D incidence differ by genetic risk remains scarce. Recent advancements in genome-wide association analyses have led to the identification of more than 400 single nucleotide polymorphisms (SNPs) for T2D,<sup>6–7</sup> allowing for in-depth investigation of the interactions between physical activity and genetic risk on incident T2D.

Capitalising on a large sample of adults with available accelerometry, genetic and prospective health data, this study aimed to examine (1) the prospective dose-response relationships between total and intensity-specific physical activity and incident T2D while accounting for genetic risk; and (2) the association between total and intensity-specific physical activity and incident T2D across different levels of genetic risk, and the associations between combinations of genetic risk and physical activity levels and incident T2D.

## METHODS

### Study design and participants

We used data from the UK Biobank (reference number: 63454), a prospective cohort study of over 500 000 participants aged 40–69 years from the UK, with detailed information reported elsewhere.<sup>8</sup> Briefly, between the years 2006 and 2010, baseline assessments were carried out in 22 assessment centres comprising a self-completed touch-screen questionnaire in addition to physical and functional measures and collection of a fresh blood sample. Additionally, between February 2013 and December 2015, a subsample of 103 712 participants wore an accelerometer on their dominant wrist for seven consecutive 24-hour days.<sup>9</sup>

For the current study, we excluded participants with a diagnosis of prevalent diabetes (n=4634; online supplemental table 1), cardiovascular diseases (CVDs) or cancer (n=18 545) before accelerometer measurement. We also excluded those with no genetic data (n=1537) and those with missing values on covariates (n=8119) (online supplemental figure 1).

### Measures

#### Accelerometer-measured physical activity

Raw device-based physical activity data were collected using an accelerometer (Axivity AX3; designed by Open Lab, Newcastle University) and processed following the protocol described elsewhere.<sup>10–11</sup> The Euclidean Norm Minus One (ENMO) in milli-gravity (mg) was used to quantify the acceleration related to the movement registered.<sup>12</sup> A validated tailored machine learning model using balanced random forests with Hidden Markov models was then used to identify different movement behaviours in 30 s time windows, including moderate-to-vigorous-intensity physical activity (MVPA), LPA, sedentary behaviour and sleep.<sup>9</sup> The total volume of physical activity was summarised as the

average ENMO while awake (defined as 06:00 to 22:00), while daily MVPA and LPA were defined as the sum of all minutes spent on waking behaviours at  $\geq 3$  and 1.5–2.9 metabolic equivalent of tasks (METs) in a typical day.

### Genetic risk

Genetic risk was estimated by applying a polygenic risk score (PRS) including 424 selected SNPs identified from genome-wide association analyses.<sup>6–7</sup> The detailed genotyping process, imputation and quality control in the UK Biobank have been described elsewhere.<sup>13</sup> A PRS was calculated using the following formula with each SNP weighted:  $PRS = \beta_1 \times \text{SNP}_1 + \beta_2 \times \text{SNP}_2 + \dots + \beta_n \times \text{SNP}_n$ , where  $\beta_n$  was the relative effect size and  $\text{SNP}_n$  is the risk allele number of each SNP (online supplemental table 2). A higher PRS indicates a higher genetic predisposition to developing T2D. PRSs were categorised by tertile into low, intermediate or high genetic risk.

### Incident T2D

Prevalent diabetes was identified based on the modified Biobank algorithms by Eastwood *et al*,<sup>14</sup> via hospital inpatient and outpatient records, primary care data, death registration, self-reported medical history and medication, as well as biochemical examination for blood glucose (if random  $\text{GLU} \geq 11.1$  mmol/L) and glycated haemoglobin (if  $\text{HbA1c} \geq 6.5\%$ ). Incident T2D was defined as developing T2D (both fatal and nonfatal) after completing an accelerometer assessment up until the censoring date (30 September 2021 for England, 31 July 2021 for Scotland and 28 February 2018 for Wales), ascertained by using the code E11 from the International Classification of Diseases 10th Revision (ICD-10) (online supplemental table 1).<sup>7–15</sup> We calculated the follow-up time from the accelerometer assessment to the time of T2D diagnosis, death, loss to follow-up or censorship, whichever occurred first.

### Covariates

Potential confounders were selected based on an a priori developed acyclic graph (online supplemental figure 2). Sociodemographic characteristics included age (continuous in years), self-reported gender (men; women), ethnicity (white European; non-white European), educational attainment (no qualifications; other qualifications than college/university degree; college or university degree), household income (<18 000; 18 000–30 999; 31 000–51 999; >52 000 £/year), Townsend deprivation index (quartiles), employment status (unemployed; employed; retired) and assessment centres (22 categories). Additional covariates included non-physical activity lifestyle factors: smoking status (never; former; current), alcohol consumption (never or special occasions only; one to three times a month; once or twice a week; three or four times a week; daily or on most days), healthy diet score (quartiles) based on a previously validated dietary index (online supplemental table 3) and body mass index (BMI;  $\text{kg/m}^2$ ).<sup>16</sup> Pre-existing chronic conditions were defined as either a self-reported history of hypertension, dyslipidaemia and depression at the baseline or diagnosed diseases through health records between the time of baseline assessment (the year 2006–2010) and accelerometer measurement (the year 2013–2015). Finally, we adjusted for accelerometer-related variables, including total wear days and seasonality.

### Statistical analysis

Complete case analysis was used in this study. We presented descriptive statistics in counts and percentages for categorical

variables, and means and SD for continuous variables. The continuous dose-response analyses assessed the shape of associations between total/intensity-specific physical activity and incident T2D, with data trimmed at the 5th and 95th percentile of the exposure distribution.<sup>17</sup> Restricted cubic splines were used to allow for potential nonlinearity with three knots placed at the 10th, 50th (reference) and 90th percentiles. We assumed linearity for values of the bottom and top 10%. Departure from linearity was examined by a Wald test.

Total physical activity, MVPA and LPA were then categorised into 4-level variables based on the 10th, 50th and 90th percentiles of data distribution, separately. We conducted multivariable Cox proportional hazards models to estimate the HRs and 95% CIs for incident T2D according to the categories of physical activity exposures, using age as the underlying timescale. The proportional hazards assumption was checked using the Schoenfeld residuals, and no violation was found.

Five sequential models were built: Model 1 accounted for age as the timescale and adjusted for gender and accelerometer-related variables as covariates. Model 2 additionally adjusted for other sociodemographic characteristics and Model 3 further adjusted for non-physical activity lifestyle factors and pre-existing chronic conditions. We further adjusted for PRS, genotyping array and the first 10 principal components of ancestry in Model 4. For the intensity-specific physical activity model, we mutually adjusted for MVPA and LPA in Model 5. We used Model 4 as the main model for total physical activity and Model 5 as the main model for intensity-specific physical activity.

We tested PRS as a potential effect modifier by adding a multiplicative interaction term (total physical activity $\times$ PRS in Model 4; MVPA $\times$ PRS and LPA $\times$ PRS in Model 5 one at a time), followed by stratified analysis by PRS tertiles (low, intermediate and high risk). Considering the different baseline hazards in models for each stratum, we also examined the joint association of physical activity and genetic risk by creating (4 $\times$ 3) mutually exclusive categories combining physical activity levels and PRS with the highest risk combination (ie, highest PRS and lowest physical activity level) as the reference.

We conducted several sensitivity analyses. First, to reduce the risk of reverse causation, we conducted a sensitivity analysis by left truncating the first 2 years of the follow-up period. Second, considering that the current accelerometer processing algorithm (eg, setting 06:00 to 22:00 as waking time) may not accurately classify physical activity for shift workers, we conducted a sensitivity analysis excluding shift workers (n=5111). Third, we reran the models defining awake time as 07:00 to 21:00 and 08:00 to 20:00. Fourth, considering that BMI could be a potential confounder or a mediator, and the current main models treated BMI as a mediator and therefore did not adjust for it, we conducted an additional sensitivity analysis adjusting for BMI as a confounder.<sup>18 19</sup> Lastly, as a post-hoc sensitivity analysis, we performed Aalen's additive hazards models to test the additive interaction effects.<sup>20</sup> The statistical analysis plan was pre-registered on Open Science Framework (Identifier: DOI 10.17605/OSF.IO/9ST7J). We conducted all analyses in SAS (V.9.4) and R (V.3.6.0). Results were reported according to STrengthening the Reporting of OBServational studies in Epidemiology Checklists (online supplemental table 11).

### Equity, diversity and inclusion statement

Our research team included two women and three men, four authors are people of colour. All authors are at the early-stage or mid-stage of their careers. The UK Biobank sample was

population-based (including a broad range of sociodemographic characteristics), but not population-representative, which we acknowledged as a limitation in the Discussion section.

## RESULTS

The final analytical sample included 59 325 participants without diabetes, CVDs or cancer at the time of accelerometer assessment, with valid physical activity, genetic data and all covariates (online supplemental figure 1).

Table 1 presents the baseline characteristics of the analytical sample by levels of total physical activity. Overall, 44.0% of the participants were men; 97.1% were of white European descent; 46.6% had a college or university degree, over 60% were employed and a similar proportion had a household income of over 31 000 £/year. The prevalence of current smoking, poor diet, overweight and obesity, hypertension, dyslipidaemia and depression was the highest among the group with the lowest level of total physical activity. Additional information on the baseline characteristics of the sample by different levels of MVPA and LPA are displayed in online supplemental tables 4 and 5, respectively.

### Physical activity and incident T2D

We observed 884 incident T2D cases during a median follow-up period of 6.8 years (IQR=6.3–7.3). Table 2 shows the HRs and 95% CIs for the association between total physical activity/MVPA/LPA and incident T2D.

For overall physical activity, compared with the least active participants, those in the most active group had an 80% lower risk of incident T2D in Model 1. The association was attenuated after the covariates were introduced. In the final model with PRS adjusted (Model 4), the inverse relationship between total physical activity and incident T2D remained significant with the most active group having a 68% lower risk of incident T2D. When examining intensity-specific physical activity, MVPA had a strong association with T2D (HR=0.26; 95% CI=0.18 to 0.38 in the most active group), even after adjusting for LPA and all other covariates. However, for LPA, we only found a statistically significant association with incident T2D at the highest level (top 10%). The dose-response analysis revealed linear relationships between all physical activity exposures and incident T2D (figure 1).

### Stratified and joint analysis

There is a strong positive association between genetic risk score and incident T2D (HR=1.41; 95% CI=1.16 to 1.71 for the intermediate risk group and HR=2.43; 95% CI=2.04 to 2.90 for the high genetic risk group, as compared with low genetic risk group) (online supplemental table 6). Stratified by genetic risk, within each category, total physical activity and MVPA were consistently associated with incident T2D, but LPA was not (table 3). For example, compared with those with the lowest level of total physical activity, participants with the highest level had a 66%, 76% and 61% lower risk of T2D in the high, intermediate and low genetic risk groups. There was no significant multiplicative interaction based on p value and stratified analyses. However, we found a significant additive interaction with MVPA (p=0.011), but not with total physical activity (p=0.113) or LPA (p=0.524). Specifically, we observed absolute risk differences of 3.8, 2.1 and 1.0 cases per 100 000 person-years between the most and least active groups based on MVPA in those with high, intermediate and low genetic risk, respectively (online supplemental figure 3). Within each genetic risk

**Table 1** Baseline characteristics of the participants by categories of the accelerometer-measured total volume of physical activity (n=59 325)

	Total sample	Total volume of physical activity (mg)*				P value
		P <sub>0</sub> -P <sub>10</sub> (≤26.1)	P <sub>10</sub> -P <sub>50</sub> (26.2-38.0)	P <sub>50</sub> -P <sub>90</sub> (38.1-54.3)	P <sub>90</sub> -P <sub>100</sub> (>54.3)	
Sample size (n)	59 325	5933	23 730	23 729	5933	
Follow-up in years, median (IQR)	6.8 (6.3-7.3)	6.7 (6.2-7.3)	6.8 (6.3-7.3)	6.9 (6.3-7.4)	6.9 (6.3-7.4)	
<i>Sociodemographic characteristics</i>						
Age at accelerometer measurement, years, mean (SD)	61.1 (7.8)	64.4 (7.5)	62.2 (7.7)	60.2 (7.7)	57.5 (7.4)	<0.001
Men, n (%)	26 098 (44.0)	3089 (52.1)	10 547 (44.4)	9777 (41.2)	2685 (45.3)	<0.001
White European, n (%)	57 577 (97.1)	5789 (97.6)	23 065 (97.2)	22 989 (96.9)	5734 (96.6)	0.004
College or university degree, n (%)	27 621 (46.6)	2532 (42.7)	11 182 (47.1)	11 136 (46.9)	2771 (46.7)	<0.001
Townsend deprivation index, n (%)						
Quartile 1	14 830 (25.0)	1344 (22.7)	5895 (24.8)	6090 (25.7)	1501 (25.3)	<0.001
Quartile 2	14 834 (25.0)	1461 (24.6)	5936 (25.0)	5949 (25.1)	1488 (25.1)	
Quartile 3	14 828 (25.0)	1464 (24.7)	5917 (24.9)	5984 (25.2)	1463 (24.7)	
Quartile 4	14 833 (25.0)	1664 (28.0)	5982 (25.2)	5706 (24.0)	1481 (25.0)	
Household income, £/year, n (%)						
<18 000	7794 (13.1)	1189 (20.0)	3302 (13.9)	2684 (11.3)	619 (10.4)	<0.001
18 000-30 999	13 573 (22.9)	1602 (27.0)	5617 (23.7)	5149 (21.7)	1205 (20.3)	
31 000-51 999	17 425 (29.4)	1639 (27.6)	6911 (29.1)	7094 (29.9)	1781 (30.0)	
>52 000	20 533 (34.6)	1503 (25.3)	7900 (33.3)	8802 (37.1)	2328 (39.2)	
Employed, n (%)	38 487 (64.9)	2979 (50.2)	14 555 (61.3)	16 423 (69.2)	4530 (76.4)	<0.001
<i>Non-PA lifestyle factors</i>						
Current smoking, n (%)	4023 (6.8)	640 (10.8)	1629 (6.9)	1430 (6.0)	324 (5.5)	<0.001
Alcohol consumption frequency, n (%)						
Daily or most daily	13 689 (23.1)	1322 (22.3)	5544 (23.4)	5511 (23.2)	1312 (22.1)	<0.001
Three or four times a week	15 979 (26.9)	1306 (22.0)	6265 (26.4)	6683 (28.2)	1725 (29.1)	
Once or twice a week	14 987 (25.3)	1405 (23.7)	5939 (25.0)	6085 (25.6)	1558 (26.3)	
One to three times a month	6479 (10.9)	770 (13.0)	2640 (11.1)	2475 (10.4)	594 (10.0)	
Never or special occasions only	8191 (13.8)	1130 (19.0)	3342 (14.1)	2975 (12.5)	744 (12.5)	
Healthy diet score, mean (SD) †	3.1 (1.4)	2.9 (1.4)	3.1 (1.4)	3.2 (1.4)	3.3 (1.5)	<0.001
Body mass index, kg/m <sup>2</sup> , n (%)						
Underweight (<18.5)	336 (0.6)	17 (0.3)	109 (0.5)	146 (0.6)	64 (1.1)	<0.001
Normal (18.5-24.9)	24 446 (41.2)	1475 (24.9)	8362 (35.2)	11 142 (47.0)	3467 (58.4)	
Overweight (25.0-29.9)	24 402 (41.1)	2500 (42.1)	10 380 (43.7)	9522 (40.1)	2000 (33.7)	
Obese (≥30.0)	10 141 (17.1)	1941 (32.7)	4879 (20.6)	2919 (12.3)	402 (6.8)	
<i>Pre-existing chronic conditions</i>						
Hypertension, n (%)	27 765 (46.8)	3641 (61.4)	12 022 (50.7)	9972 (42.0)	2130 (35.9)	<0.001
Dyslipidaemia, n (%)	26 101 (44.0)	3490 (58.8)	11 360 (47.9)	9435 (39.8)	1816 (30.6)	<0.001
Depression, n (%)	6836 (11.5)	899 (15.2)	2819 (11.9)	2531 (10.7)	587 (9.9)	<0.001
Sources of T2D diagnosis, n (%)						
Not diagnosed	58 441 (98.5)	5728 (96.5)	23 321 (98.3)	23 493 (99.0)	5899 (99.4)	<0.001
Death registration, self-reported or others	105 (0.2)	21 (0.4)	38 (0.2)	41 (0.2)	5 (0.1)	
Hospital inpatient record	779 (1.3)	184 (3.1)	371 (1.6)	195 (0.8)	29 (0.5)	
<i>Accelerometer-related variables</i>						
Total wear days, mean (SD)	6.7 (0.7)	6.7 (0.7)	6.7 (0.7)	6.7 (0.7)	6.7 (0.8)	0.059
The total volume of physical activity, mg, mean (SD)	39.5 (11.9)	22.4 (3.1)	32.6 (3.3)	44.6 (4.5)	63.7 (10.3)	<0.001
MVPA, min/day, mean (SD)	32.6 (27.3)	13.4 (12.5)	24.7 (18.9)	38.3 (26.5)	61.2 (38.7)	<0.001
LPA, min/day, mean (SD)	319.0 (94.1)	205.7 (59.3)	287.5 (68.1)	356.1 (80.3)	409.4 (101.2)	<0.001
PRS, mean (SD)	414.3 (12.1)	413.8 (12.0)	414.3 (12.1)	414.3 (12.2)	414.4 (12.1)	0.008

\*Based on the 10th, 50th and 90th percentiles of total physical activity distribution.

†Health diet score was calculated based on self-reported servings of fruits, vegetables, whole grains, vegetable oil, fish, dairy, refined grains, unprocessed meats, processed meats and sugar-sweetened beverages. More details can be found in online supplemental table 3.

LPA, light-intensity physical activity; mg, milligravity; MVPA, moderate-to-vigorous-intensity physical activity; P, percentiles; PRS, polygenic risk score; T2D, type 2 diabetes.

subgroup, we observed stronger associations between MVPA and T2D compared with those between LPA and T2D.

Figure 2 shows the association between joint physical activity/PRS categories and T2D. Overall, results based on the total physical activity and MVPA showed similar patterns. In both cases, high genetic risk and the highest total physical activity/

MVPA combination were associated with a lower risk of incident T2D than low genetic risk and the lowest total physical activity/MVPA combination. In contrast, associations between LPA/PRS combinations and T2D were less clear—those with high genetic risk and the lowest three LPA categories had similar HRs and all the other combinations had similarly smaller HRs.

**Table 2** Association between accelerometer-measured total and intensity-specific physical activity and incident T2D (n=59 325)

	Events/number of participants	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 3 HR (95% CI)	Model 4 HR (95% CI)	Model 5 HR (95% CI)
The total volume of physical activity (mg)						
P <sub>0</sub> –P <sub>10</sub> (≤26.1)	205/5728	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)
P <sub>10</sub> –P <sub>50</sub> (26.2–38.0)	409/23 321	<b>0.54 (0.46 to 0.64)</b>	<b>0.59 (0.50 to 0.70)</b>	<b>0.71 (0.60 to 0.84)</b>	<b>0.68 (0.57 to 0.80)</b>	–
P <sub>50</sub> –P <sub>90</sub> (38.1–54.3)	236/23 493	<b>0.33 (0.28 to 0.41)</b>	<b>0.37 (0.31 to 0.45)</b>	<b>0.51 (0.42 to 0.62)</b>	<b>0.49 (0.40 to 0.60)</b>	–
P <sub>90</sub> –P <sub>100</sub> (>54.3)	34/5899	<b>0.20 (0.14 to 0.29)</b>	<b>0.22 (0.15 to 0.32)</b>	<b>0.35 (0.24 to 0.50)</b>	<b>0.32 (0.22 to 0.47)</b>	–
MVPA (min/day)						
P <sub>0</sub> –P <sub>10</sub> (≤5.2)	193/5758	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)
P <sub>10</sub> –P <sub>50</sub> (5.3–25.9)	418/23 316	<b>0.50 (0.42 to 0.59)</b>	<b>0.54 (0.45 to 0.64)</b>	<b>0.63 (0.53 to 0.75)</b>	<b>0.62 (0.52 to 0.74)</b>	<b>0.63 (0.53 to 0.75)</b>
P <sub>50</sub> –P <sub>90</sub> (26.0–68.4)	237/23 470	<b>0.27 (0.22 to 0.32)</b>	<b>0.31 (0.25 to 0.37)</b>	<b>0.41 (0.33 to 0.50)</b>	<b>0.40 (0.33 to 0.49)</b>	<b>0.41 (0.34 to 0.51)</b>
P <sub>90</sub> –P <sub>100</sub> (>68.4)	36/5897	<b>0.15 (0.10 to 0.21)</b>	<b>0.17 (0.12 to 0.25)</b>	<b>0.26 (0.18 to 0.38)</b>	<b>0.25 (0.18 to 0.37)</b>	<b>0.26 (0.18 to 0.38)</b>
LPA (min/day)						
P <sub>0</sub> –P <sub>10</sub> (≤202)	135/5798	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)
P <sub>10</sub> –P <sub>50</sub> (202.1–314)	379/23 368	<b>0.73 (0.60 to 0.89)</b>	<b>0.75 (0.62 to 0.92)</b>	0.84 (0.69 to 1.02)	<b>0.81 (0.66 to 0.98)</b>	0.89 (0.73 to 1.09)
P <sub>50</sub> –P <sub>90</sub> (314.1–444)	315/23 398	<b>0.64 (0.52 to 0.78)</b>	<b>0.64 (0.52 to 0.79)</b>	<b>0.79 (0.64 to 0.97)</b>	<b>0.76 (0.62 to 0.93)</b>	0.87 (0.70 to 1.07)
P <sub>90</sub> –P <sub>100</sub> (>444)	55/5877	<b>0.48 (0.35 to 0.66)</b>	<b>0.45 (0.33 to 0.62)</b>	<b>0.58 (0.42 to 0.80)</b>	<b>0.57 (0.41 to 0.78)</b>	<b>0.64 (0.47 to 0.89)</b>

The analyses were all based on multivariable Cox proportional hazards models. Bold face values indicate  $p < 0.05$ .

Model 1: adjusted for age as the underlying timescale, adjusted for gender, total wear time and seasonality.

Model 2: adjusted for variables in Model 1+ ethnicity, educational attainment, household income, Townsend deprivation index, employment status and assessment centres.

Model 3: adjusted for variables in Model 2+ smoking status, alcohol consumption, healthy diet score, hypertension, dyslipidaemia and depression.

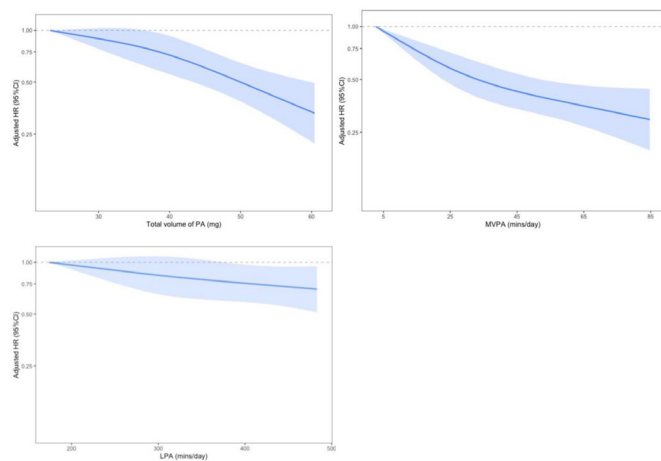
Model 4: adjusted for variables in Model 3+ PRS, genotyping array, and the first 10 principal components of ancestry.

Model 5: adjusted for variables in Model 4+ LPA or MVPA (MVPA and LPA were mutually adjusted, but not for the total volume of physical activity).

LPA, light-intensity physical activity; mg, milligravity; MVPA, moderate-to-vigorous-intensity physical activity; P, percentiles; PRS, polygenic risk score; T2D, type 2 diabetes.

### Sensitivity analysis

Left truncating the first 2 years of follow-up did not appreciably change the associations between total physical activity/MVPA and incident T2D (online supplemental table 6), but the



**Figure 1** Restricted cubic spline for the association between accelerometer-measured total and intensity-specific physical activity and incident T2D. The analyses were all based on multivariable Cox proportional hazards models adjusted for age as the underlying timescale, gender, total wear time, seasonality, ethnicity, educational attainment, household income, Townsend deprivation index, employment status, assessment centres, smoking status, alcohol consumption, healthy diet score, hypertension, dyslipidaemia, depression, PRS, genotyping array, the first 10 principal components of ancestry, and LPA or MVPA (MVPA and LPA were mutually adjusted, but not for total volume of physical activity). \*The y-axis is plotted on a log scale. Solid line is the estimate (HR) and the shade area is the 95% CI. LPA, light-intensity physical activity; mg, milligravity; MVPA, moderate-to-vigorous-intensity physical activity; PA, physical activity; PRS, polygenic risk score; T2D, type 2 diabetes.

highest level of LPA was no longer associated with incident T2D. Excluding shift workers (n=5 111) and redefining awake time as 07:00 to 21:00 or 08:00 to 20:00 did not substantially alter any estimates (online supplemental tables 7–9). Furthermore, the relationships between total/intensity-specific physical activity and incident T2D were attenuated after adjusting for BMI, and no inverse relationship was found between LPA and T2D incidence after accounting for BMI (online supplemental table 10).

### DISCUSSION

Based on the largest accelerometry dataset paired with the most comprehensive genetic risk indicators to our knowledge, our study showed that higher levels of total physical activity and particularly MVPA were strongly associated with a lower risk of developing T2D, independent of genetic risk. Considering that physical inactivity is a modifiable risk factor in most cases, our findings provide insights for informing public health recommendations to prevent T2D. Particularly, we found a linear dose-response association with a considerably larger magnitude than studies using self-reported physical activity. The association was the strongest when physical activity was operationalised as MVPA, followed by total physical activity, and the association between LPA and T2D was the least consistent and weakest. Finally, although we did not find a significant effect modification by genetic risk on a multiplicative scale, we found a significant additive interaction indicating a larger absolute risk reduction among those with the highest genetic risk.

Our results confirm previously observed protective effects of physical activity on the onset of T2D, regardless of genetic risk. In our study, 68.4 min/day of MVPA (ie, 90th percentile) was associated with a 74% lower risk of developing T2D when compared with participants who spent less than 5.2 min in MVPA (ie, 10th percentile). This association is of a larger magnitude than previous studies based on self-reported physical activity,<sup>3</sup> which was also observed for mortality outcomes.<sup>21</sup> This is likely

**Table 3** Associations between accelerometer-measured total and intensity-specific physical activity and incident T2D stratified by different levels of genetic risk

	High genetic risk		Intermediate genetic risk		Low genetic risk	
	n with/without incident T2D	HR (95% CI)	n with/without incident T2D	HR (95% CI)	n with/without incident T2D	HR (95% CI)
The total volume of physical activity (mg)						
P <sub>0</sub> -P <sub>10</sub> (≤26.1)	93/1801	1.00 (ref.)	63/1891	1.00 (ref.)	49/2036	1.00 (ref.)
P <sub>10</sub> -P <sub>50</sub> (26.2-38.0)	221/7679	<b>0.76 (0.59 to 0.97)</b>	113/7813	<b>0.62 (0.45 to 0.86)</b>	75/7829	<b>0.61 (0.42 to 0.89)</b>
P <sub>50</sub> -P <sub>90</sub> (38.1-54.3)	126/7847	<b>0.53 (0.40 to 0.70)</b>	67/7882	<b>0.45 (0.32 to 0.65)</b>	43/7764	<b>0.44 (0.29 to 0.68)</b>
P <sub>90</sub> -P <sub>100</sub> (>54.3)	19/1989	<b>0.34 (0.21 to 0.57)</b>	7/1939	<b>0.24 (0.11 to 0.54)</b>	8/1971	<b>0.39 (0.18 to 0.85)</b>
MVPA (min/day)						
P <sub>0</sub> -P <sub>10</sub> (≤5.2)	95/1908	1.00 (ref.)	61/1923	1.00 (ref.)	37/1927	1.00 (ref.)
P <sub>10</sub> -P <sub>50</sub> (5.3-25.9)	224/7720	<b>0.64 (0.50 to 0.82)</b>	114/7809	<b>0.55 (0.40 to 0.77)</b>	80/7787	<b>0.64 (0.42 to 0.95)</b>
P <sub>50</sub> -P <sub>90</sub> (26.0-68.4)	118/7719	<b>0.38 (0.29 to 0.51)</b>	65/7913	<b>0.35 (0.24 to 0.50)</b>	54/7838	<b>0.52 (0.33 to 0.81)</b>
P <sub>90</sub> -P <sub>100</sub> (>68.4)	22/1969	<b>0.30 (0.19 to 0.49)</b>	10/1880	<b>0.23 (0.11 to 0.46)</b>	4/2048	<b>0.16 (0.06 to 0.46)</b>
LPA (min/day)						
P <sub>0</sub> -P <sub>10</sub> (≤202)	63/1808	1.00 (ref.)	42/1976	1.00 (ref.)	30/2014	1.00 (ref.)
P <sub>10</sub> -P <sub>50</sub> (202.1-314)	194/7696	0.84 (0.63 to 1.12)	112/7794	0.83 (0.57 to 1.19)	73/7878	0.75 (0.49 to 1.16)
P <sub>50</sub> -P <sub>90</sub> (314.1-444)	179/7808	0.87 (0.64 to 1.16)	80/7789	0.68 (0.46 to 1.01)	56/7801	0.64 (0.40 to 1.02)
P <sub>90</sub> -P <sub>100</sub> (>444)	23/2004	<b>0.46 (0.28 to 0.75)</b>	16/1966	0.60 (0.33 to 1.08)	16/1907	0.76 (0.40 to 1.42)
Effect modification on multiplicative scale total physical activity×PRS: p value=0.343.						
Effect modification on multiplicative scale MVPA×PRS: p value=0.671.						
Effect modification on multiplicative scale LPA×PRS: p value=0.662.						
The analyses were all based on multivariable Cox proportional hazards models adjusted for age as the underlying timescale, gender, total wear time, seasonality, ethnicity, educational attainment, household income, Townsend deprivation index, employment status, assessment centres, smoking status, alcohol consumption, healthy diet score, hypertension, dyslipidaemia, depression, PRS, genotyping array, the first 10 principal components of ancestry, and LPA or MVPA (MVPA and LPA were mutually adjusted, but not for the total volume of PA). Bold face values indicate p<0.05.						
LPA, light-intensity physical activity; mg, milligravity; MVPA, moderate-to-vigorous-intensity physical activity; P, percentiles; PA, physical activity; PRS, polygenic risk score; T2D, type 2 diabetes.						

explained by the measurement errors in self-reported physical activity that tend to bias the association towards the null.<sup>22 23</sup>

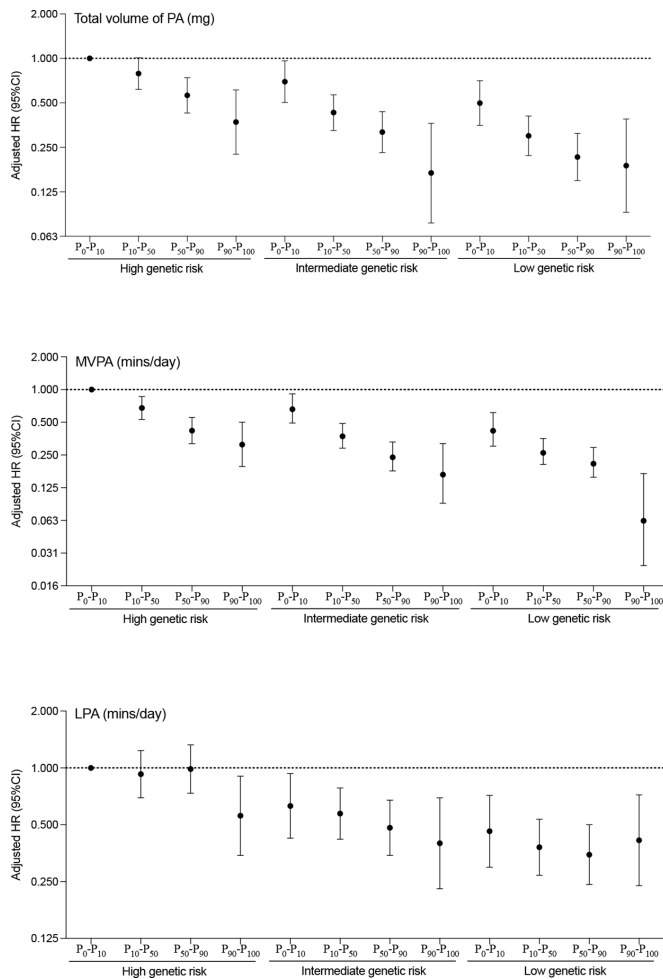
Remarkably, a linear dose-response pattern with no minimal or maximal threshold was identified. This finding contrasts previous evidence supporting a curvilinear dose-response association whereby the association became weaker at higher levels of physical activity.<sup>5</sup> For example, Kyu and colleagues' dose-response meta-analysis including 55 prospective cohort studies with self-reported physical activity found a clear threshold effect,<sup>24</sup> where the largest risk reduction was observed when physical activity was increased from 600 to 3600 MET-min/week (equivalent to approximately from 150 to 900 min/week of MPA). Beyond this threshold, increasing levels of physical activity were associated with little improvements in diabetes outcomes. However, reporting biases, particularly at the higher levels of physical activity, may have potentially distorted the observed associations. Our results indicate that individuals should be encouraged to be as physically active as possible to maximise the benefits.

When examining intensity-specific physical activity, the association between MVPA and incident T2D was strong and significant at every level when compared with the bottom 10% MVPA level (≤5.2 min/week). In contrast, only participants in the top 10% LPA level (over 7.4 hours/day) had a statistically significantly lower risk of developing T2D and the association became non-significant in several sensitivity analyses. Due to the difficulty of measuring LPA using self-reported instruments, the evidence base for LPA and health outcomes is less developed than that for MVPA. The large amount of LPA required to reduce the risk of T2D may, at least in part, be explained by the interaction of the insulin system with different intensities of physical activity. A position statement by the American Diabetes

Association pointed out that 20 min of MVPA can help elevate glucose uptake and enhance insulin action, while far more LPA is needed to achieve a similar effect.<sup>4</sup> Evidence from our study implies that, when possible, it may be more effective for T2D prevention programmes to focus on promoting MVPA.

We did not observe a multiplicative interaction between physical activity and genetic risk using 424 SNPs among the UK Biobank participants. However, we found a significant additive interaction with MVPA, suggesting the largest absolute risk reduction in those with a high genetic predisposition to T2D. Interestingly, our results showed that participants with a high genetic risk and the highest total physical activity and MVPA (ie, upper 10th percentile, at least 68.4 min/day of MVPA) had a lower risk of incident T2D than those with a low genetic risk and the lowest total physical activity/MVPA (ie, bottom 10th percentile, less than 5.2 min/day of MVPA). These results underscore the importance of being physically active for T2D prevention, particularly in those with high genetic risk. Our finding is in line with two studies using 19 SNPs in Poland and 49 SNPs in nine European countries.<sup>25 26</sup> In contrast, Klimentidis and his colleagues found that the protective effect of physical activity in those with higher genetic risk was weaker compared with those with lower genetic risk using a PRS comprised of 65 SNPs in the USA.<sup>27</sup> The difference between studies is likely due to different measurements of physical activity (ie, our study used accelerometers while previous studies used self-reported physical activity) and genetic risk (our study captured more T2D-related SNPs than previous studies).<sup>28</sup>

To our knowledge, this investigation is the first population-based prospective cohort study to explore the relationship between accelerometer-measured physical activity and incident T2D while accounting for genetic risk. Our study has several



**Figure 2** Joint association between accelerometer-measured total and intensity-specific physical activity and polygenic risk score on incident T2D. The analyses were all based on multivariable Cox proportional hazards models adjusted for age as the underlying timescale, gender, total wear day, seasonality, ethnicity, educational attainment, household income, Townsend deprivation index, employment status, assessment centres, smoking status, alcohol consumption, healthy diet score, sources of T2D diagnose, hypertension, dyslipidaemia, depression, genotyping array, the first 10 principal components of ancestry, and LPA or MVPA (MVPA and LPA were mutually adjusted, but not for total volume of physical activity). A polygenic risk score was constructed with 424 selected SNPs genome-wide significantly associated with T2D, and was further categorised into low, intermediate and high genetic risk groups by tertile. \*The y-axis is plotted on a log scale. LPA, light-intensity physical activity; mg, milligravity; MVPA, moderate-to-vigorous-intensity physical activity; P, percentiles; PA, physical activity; SNPs, single nucleotide polymorphisms; T2D, type 2 diabetes.

strengths. First, while most previous studies used self-reported physical activity measures, which are prone to recall and social desirability biases, we used accelerometers that monitor physical activity of different intensities continuously through 24-hour days. Second, we used machine learning approaches to quantify the volume of intensity-specific physical activity instead of using pre-defined cut points, which are subject to misclassifications, and have shown substantial agreement with wearable camera ground-truth data.<sup>29,30</sup> Furthermore, the machine learning model used in the current study was trained and validated in a realistic

free-living environment, which covered a wider variety of activities than in previous lab-based studies.<sup>31</sup> Third, we included 424 selected SNPs for the genetic risk of T2D, which is the most comprehensive genetic study on physical activity and T2D available to date.

This study is subject to limitations. Physical activity was only measured at a single time point which could not capture long-term patterns of physical activity. Additionally, participants of the UK Biobank study are not nationally representative of the general UK population; however, a previous study found that the association between risk factors and health outcomes derived from the study may be generalisable.<sup>32</sup> Lastly, since the accelerometry data were collected between 2013 and 2015, the follow-up time is relatively short.

In conclusion, our results reinforced the importance of physical activity for T2D prevention, particularly for those with a high genetic risk. Future guideline development should consider physical activity intensity and the dose-response nature of the association. The promotion of physical activity, particularly MVPA, should be a priority strategy for T2D prevention.

#### Author affiliations

<sup>1</sup>Prevention Research Collaboration, Sydney School of Public Health, The University of Sydney, Sydney, New South Wales, Australia

<sup>2</sup>Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia

<sup>3</sup>School of Public Health, School of Medicine, Shanghai Jiao Tong University, Shanghai, People's Republic of China

<sup>4</sup>Centre for Active and Healthy Ageing, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

<sup>5</sup>Faculty of Education, University of Cádiz, Cádiz, Spain

<sup>6</sup>Biomedical Research and Innovation Institute of Cádiz (INIbICA) Research Unit, Puerta del Mar University Hospital, University of Cádiz, Cádiz, Spain

<sup>7</sup>Department of Nutrition and Food Hygiene, Hubei Key Laboratory of Food Nutrition and Safety, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, People's Republic of China

**Twitter** Mengyun Luo @SusanLuo1 and Ding Ding @DrMelodyDing

**Acknowledgements** We gratefully acknowledge the commitment and dedication of the participants of the UK Biobank.

**Contributors** ML, LC and DD conceptualised and designed the study. ML analysed the data with the help of CY, BDPC, LC and DD, and drafted the study with DD. All the authors reviewed the results, provided input during critical revisions and approved the submitted manuscript. All authors have final responsibility for the decision to submit for publication.

**Funding** DD is supported by a National Health and Medical Research Council Investigator Grant (# 2009254). LC is supported by the Young Elite Scientists Sponsorship Program by China Association for Science and Technology (grant number: YESS20210143). BDPC is supported by the Government of Andalusia, Research Talent Recruitment Program (EMERGIA #2020/00158).

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by North West Multi-centre Research Ethics Committee. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** The data analyzed in this study are available from the UK Biobank website with approved access. Data could be obtained upon direct application to the UK Biobank Study.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines,

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Mengyun Luo <http://orcid.org/0000-0002-6482-7299>

Chenhao Yu <http://orcid.org/0000-0002-4759-1042>

Borja Del Pozo Cruz <http://orcid.org/0000-0003-3944-2212>

Liangkai Chen <http://orcid.org/0000-0002-5227-4236>

Ding Ding <http://orcid.org/0000-0001-9850-9224>

#### REFERENCES

- World Health Organization. Diabetes. Available: <https://www.who.int/news-room/fact-sheets/detail/diabetes> [Accessed 27 Oct 2022].
- International Diabetes Federation. IDF diabetes atlas. Available: <https://diabetesatlas.org> [Accessed 27 Oct 2022].
- Smith AD, Crippa A, Woodcock J, et al. Physical activity and incident type 2 diabetes mellitus: a systematic review and dose-response meta-analysis of prospective cohort studies. *Diabetologia* 2016;59:2527–45.
- Colberg SR, Sigal RJ, Yardley JE, et al. Physical activity/exercise and diabetes: a position statement of the American diabetes association. *Diabetes Care* 2016;39:2065–79.
- 2018 Physical Activity Guidelines Advisory Committee. *2018 Physical activity guidelines advisory committee scientific report*. Washington, DC: U.S. Department of Health and Human Services, 2018.
- Vujkovic M, Keaton JM, Lynch JA, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet* 2020;52:680–91.
- Zhuang P, Liu X, Li Y, et al. Effect of diet quality and genetic predisposition on hemoglobin A1c and type 2 diabetes risk: gene-diet interaction analysis of 357,419 individuals. *Diabetes Care* 2021;44:2470–9.
- Sudlow C, Gallacher J, Allen N, 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:e1001779.
- Doherty A, Jackson D, Hammerla N, et al. Large scale population assessment of physical activity using wrist worn accelerometers: the UK Biobank study. *PLoS One* 2017;12:e0169649.
- DigitalInteraction/openmovement. n.d. Available: <https://github.com/digitalinteraction/openmovement>
- Physical Activity Expert Working Group for UK Biobank. Physical activity monitor (accelerometer) version 1.0. Available: <https://biobank.ndph.ox.ac.uk/ukb/docs/PhysicalActivityMonitor.pdf> [Accessed 27 Oct 2022].
- Additional physical activity metrics. Available: [https://docs.movisens.com/Algorithms/physical\\_activity\\_metrics/#movement-acceleration-intensity-movementacceleration](https://docs.movisens.com/Algorithms/physical_activity_metrics/#movement-acceleration-intensity-movementacceleration) [Accessed 27 Oct 2022].
- Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9.
- Eastwood SV, Mathur R, Atkinson M, et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PLoS One* 2016;11:e0162388.
- UK Biobank. Category. 2000. Available: <https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=2000> [Accessed 27 Oct 2022].
- Mozaffarian D. Dietary and policy priorities for cardiovascular disease, diabetes, and obesity: a comprehensive review. *Circulation* 2016;133:187–225.
- Harrell FE. Chapter 2 General aspects of fitting regression models. In: *Regression modeling strategies*. Springer International Publishing, 2016: 13–44.
- Díez-Fernández A, Sánchez-López M, Mora-Rodríguez R, et al. Obesity as a mediator of the influence of cardiorespiratory fitness on cardiometabolic risk: a mediation analysis. *Diabetes Care* 2014;37:855–62.
- García-Hermoso A, Martínez-Vizcaino V, Recio-Rodríguez JI, et al. Abdominal obesity as a mediator of the influence of physical activity on insulin resistance in Spanish adults. *Prev Med* 2016;82:59–64.
- VanderWeele TJ, Knol MJ. A tutorial on interaction. *Epidemiol Methods* 2014;3:33–72.
- Ekelund U, Tarp J, Steene-Johannessen J, et al. Dose-Response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ* 2019;366:l4570.
- Mahmood S, Nguyen NH, Bassett JK, et al. A quantitative bias analysis to estimate measurement error-related attenuation of the association between self-reported physical activity and colorectal cancer risk. *Int J Epidemiol* 2020;49:153–61.
- Ekelund U, Dalene KE, Tarp J, et al. Physical activity and mortality: what is the dose response and how big is the effect? *Br J Sports Med* 2020;54:1125–6.
- Kyu HH, Bachman VF, Alexander LT, et al. Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. *BMJ* 2016;354:i3857.
- Szczerbiński Ł, Gościak J, Bauer W, et al. Efficacy of family history, genetic risk score, and physical activity in assessing the prevalence of type 2 diabetes. *Pol Arch Intern Med* 2019;129:442–50.
- Langenberg C, Sharp SJ, Franks PW, et al. Gene-lifestyle interaction and type 2 diabetes: the EPIC interact case-cohort study. *PLoS Med* 2014;11:e1001647.
- Klimentidis YC, Chen Z, Arora A, et al. Association of physical activity with lower type 2 diabetes incidence is weaker among individuals at high genetic risk. *Diabetologia* 2014;57:2530–4.
- Douglas GM, Bielawski JP, Langille MGI. Re-evaluating the relationship between missing heritability and the microbiome. *Microbiome* 2020;8:87.
- Ahmadi MN, Trost SG. Device-based measurement of physical activity in pre-schoolers: comparison of machine learning and cut point methods. *PLoS One* 2022;17:e0266970.
- Clark CCT, Barnes CM, Stratton G, et al. A review of emerging analytical techniques for objective physical activity measurement in humans. *Sports Med* 2017;47:439–47.
- Willetts M, Hollowell S, Aslett L, et al. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci Rep* 2018;8:7961.
- Batty GD, Gale CR, Kivimäki M, et al. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ* 2020;368:m131.