

BMJ Open Exercise and incidence of myocardial infarction, stroke, hypertension, type 2 diabetes and site-specific cancers: prospective cohort study of 257 854 adults in South Korea

Youngwon Kim,^{1,2} Stephen Sharp,² Semi Hwang,^{1,3} Sun Ha Jee⁴

To cite: Kim Y, Sharp S, Hwang S, *et al*. Exercise and incidence of myocardial infarction, stroke, hypertension, type 2 diabetes and site-specific cancers: prospective cohort study of 257 854 adults in South Korea. *BMJ Open* 2019;**9**:e025590. doi:10.1136/bmjopen-2018-025590

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-025590>).

Received 23 July 2018
Revised 5 December 2018
Accepted 19 December 2018



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY. Published by BMJ.

¹School of Public Health, The University of Hong Kong Li Ka Shing Faculty of Medicine, Hong Kong, Hong Kong

²MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK

³Department of Epidemiology and Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, South Korea

⁴Epidemiology, Graduate School of Public Health, Yonsei University, Seoul, South Korea

Correspondence to

Dr Sun Ha Jee; jsunha@yuhs.ac

ABSTRACT

Objective The objective of this study was to examine the longitudinal associations of exercise frequency with the incidence of myocardial infarction, stroke, hypertension, type 2 diabetes and 10 different cancer outcomes.

Design A prospective cohort study.

Setting Physical examination data linked with the entire South Korean population's health insurance system: from 2002 to 2015.

Participants 257 854 South Korean adults who provided up to 7 repeat measures of exercise (defined as exercises causing sweat) and confounders.

Primary outcome measures Each disease incidence was defined using both fatal and non-fatal health records (a median follow-up period of 13 years).

Results Compared with no exercise category, the middle categories of exercise frequency (3–4 or 5–6 times/week) showed the lowest risk of myocardial infarction (HR 0.79; 95% CI 0.70 to 0.90), stroke (HR 0.80; 95% CI 0.73 to 0.89), hypertension (HR 0.86; 95% CI 0.85 to 0.88), type 2 diabetes (HR 0.87; 95% CI 0.84 to 0.89), stomach (HR 0.87; 95% CI 0.79 to 0.96), lung (HR 0.80; 95% CI 0.71 to 0.91), liver (HR 0.85; 95% CI 0.75 to 0.98) and head and neck cancers (HR 0.76; 95% CI 0.63 to 0.93); for 1–2 times/week, exhibiting J-shaped associations. There was, in general, little evidence of effect modification by body mass index, smoking, alcohol consumption, family history of disease and sex in these associations.

Conclusions Moderate levels of sweat-inducing exercise showed the lowest risk of myocardial infarction, stroke, hypertension, type 2 diabetes, stomach, lung, liver and head and neck cancers. Public health and lifestyle interventions should, therefore, promote moderate levels of sweat-causing exercise as a behavioural prevention strategy for non-communicable diseases in a wider population of East Asians.

INTRODUCTION

Prevention and control of non-communicable diseases are a contemporary global public health priority. At present, 40 million deaths per year, which accounts for nearly 70% of total deaths globally, are attributable

Strengths and limitations of this study

- This study is the first to investigate the longitudinal associations of exercise with various cardiovascular disease and cancer incidence using a large-scale cohort dataset of South Korean adults (n=257 854) who provided up to seven repeated measures of exercise and all confounders in order to minimise the risk of regression dilution.
- A limitation is that no strong causal inference can be drawn about the exercise-incident disease relationships.
- Findings of this study may not be generalisable to adult populations of other ethnic origins.

to non-communicable diseases.^{1 2} Moreover, the number of deaths due to non-communicable diseases, such as cardiovascular disease (CVD),³ hypertension,⁴ diabetes⁵ and cancer,⁶ has increased dramatically over the past few decades, although age-standardised CVD and cancer rates as well as systolic blood pressure levels⁷ have declined.^{8 9} However, trends in these disease traits have varied across different populations, particularly with less favourable changes observed in East Asian populations compared with Western populations. For example, the prevalence of diabetes¹⁰ has increased more rapidly, while the age-standardised prevalence of CVD³ and systolic blood pressure levels⁷ have fallen less steeply in East Asians in comparison with Westerners.

In addition, adults in East Asia tend to have higher prevalence of physical inactivity,¹¹ which is one of the four target behaviours (including unhealthy diet, tobacco use and harmful use of alcohol) that have been set as the global focus to reduce the risk of non-communicable diseases.¹² The beneficial impacts of increased physical activity on

various non-communicable outcomes have been demonstrated by numerous previous investigations. However, the majority of previous research has been predicated on evidence from Western populations, thereby limiting its application to other populations including East Asians. As such, little is currently known about levels of physical activity including exercise in relation to non-communicable diseases in East Asian populations as compared with Western populations.¹³ Another critical gap in the existing literature is the use of data measured only at a single point in time (ie, baseline), in which case physical activity or exercise levels are assumed to remain constant over time. This methodology, therefore, precludes the fact that individuals' physical activity or exercise levels change with time, and hence may increase the potential for regression dilution.¹⁴ Furthermore, it is well known that temporal changes occur in other traditional behavioural and metabolic risk factors for non-communicable diseases, such as adiposity levels,^{15 16} smoking,¹⁷ glucose levels¹⁸ and total cholesterol levels,¹⁹ exhibiting different patterns of changes between East Asian and Western populations. Nevertheless, no previous research of East Asians or Westerners took into account changes in these risk markers in understanding the relationships between physical activity and non-communicable diseases. Moreover, the dose–response relationship between physical activity and various non-communicable disease outcomes has remained unclear in East Asians. Therefore, the purpose of this research was to explore the dose–response relationships between exercise frequency and various types of incident non-communicable diseases, such as myocardial infarction, stroke, hypertension, type 2 diabetes and site-specific cancers, using a large-scale prospective cohort of South Korean adults with multiple repeated measures of exercise frequency and other risk markers.

METHODS

Study design and participants

This study is based on data from the National Health Insurance Service—Health Screening (NHIS-HEALS) cohort dataset,²⁰ which is a nationally representative random sample (stratified by 2 groups of sex (males and females), 18 groups of age ranges (less than 1 year, 1–4 years, every 5 years between 5 and 79 years, and more than 80 years), 3 groups of employment status (insured employees, self-employed individuals and medical aid beneficiaries) and 41 groups of income levels (upper 20% for insured employees, lower 20% for insured self-employed individuals and the lowest level for medical aid beneficiaries))²¹ of over 500 000 South Korean adults aged 40–79 years between 2002 and 2003 made available by the NHIS. The NHIS is a single health insurance system in South Korea, which manages and maintains information on the entire South Korean population's healthcare utilisation; it is mandatory for all South Koreans to take part in the national health insurance system. The NHIS is also responsible for maintaining national health examination programmes

involving data from general health examinations of all insured employees, self-employed individuals and medical aid beneficiaries aged over 40 years; it is recommended for them to perform the health examination at least every 2 years. The health examination involves collection of information on body composition, blood profiles, blood pressure, self-reported lifestyles, self-reported physician-diagnosed disease and self-reported family history of disease.

The NHIS-HEALS cohort includes a wide variety of information collected between 2002 and 2015: health examination data and demographic and eligibility data (eg, inpatient and outpatient hospital records, medical bill, health insurance and medical aid beneficiaries). In the present analysis, we used health examination data collected between 2002 and 2008 to define the exercise frequency and all confounders. There was a change in the type of self-report methods in 2009; hence, health examination data collected in or after 2009 were not considered in the analysis due to the inability to harmonise variables. However, we used full follow-up data accrued from 2002 until 2015.

Exposure

The primary exposure variable of this study was exercise frequency, assessed using questionnaires administered during the health examinations. The specific question asked was 'How many times per week do you engage in exercise that causes sweating?' Participants were asked to choose only one of the following 5 possible answers: none, 1–2 times/week, 3–4 times/week, 5–6 times/week and almost every day.

Outcomes

In the present study, we evaluated 14 incident disease outcomes, namely, myocardial infarction; stroke, hypertension; type 2 diabetes mellitus; and stomach, colon, rectum, lung, liver, head and neck, pancreatic, kidney, gall bladder and oesophagus cancers. Participants' inpatient and outpatient hospital records (ie, non-fatal status) and death records (ie, fatal status) obtained through linkage with Statistics Korea were both classified according to the International Classification of Disease (ICD)-10 codes to classify different incidence types (online supplementary table 1). Additionally, blood pressure (eg, systolic ≥ 140 mm Hg, diastolic ≥ 90 mm Hg) and fasting glucose levels (eg, ≥ 126 mg/dL), both of which were measured during physical examinations, were used in conjunction with physicians' diagnosis information and ICD-10 codes to define incident hypertension and type 2 diabetes, respectively. Each incident disease outcome was defined as the first occurrence of either non-fatal or fatal respective disease cases. Incident disease cases were adjudicated using hospital and death records collected through 31 December 2015. The median follow-up was 13.0 years (IQR 10.2–11.3 years).

Other covariates

The following covariates were included as confounders in the analyses: sex, body mass index (weight in kilograms

(kg) divided by height in metres squared (m^2)), systolic blood pressure, fasting glucose, total cholesterol, family history of heart disease, stroke or hypertension (only in models for incident myocardial infarction, stroke or hypertension), family history of diabetes (only in models for type 2 diabetes), family history of cancer (only in models for incident cancer outcomes), smoking status (never, previously, currently) and alcohol consumption (never, 2–3 times/month, 1–2 times/week, ≥ 3 times/week).

Statistical analysis

Analyses were performed to summarise descriptive statistics (eg, means, SD, frequency and proportions) of each covariate and incident disease outcome for all participants and by exercise frequency category. Cox regression with age as the underlying timescale was used to estimate the associations of exercise frequency with each incident disease outcome, with adjustment for all the above-mentioned confounders as well as without any adjustment. HRs along with corresponding 95% CIs were calculated to evaluate relative risk of each incident disease outcome. Data were structured to enable the inclusion of exercise frequency and all confounders from both baseline and up to six repeated measures as time-updated covariates. This approach takes into account changes in exercise frequency as well as each confounder over time in relation to disease incidence. Individuals who reported no exercise served as a reference group for all comparisons. Effect modification by body mass index (<25 , ≥ 25 kg^2/m), smoking status, alcohol consumption, family history of disease and sex was also examined based on Wald tests

of interaction terms in the fully adjusted models for each incident disease outcome. Visual inspections of log–log plots provided support for the assumptions of proportional hazards for all covariates. A sensitivity analysis where incident disease cases occurring during the first 2 years of follow-up were removed was performed to address reverse causality. Analyses were performed in Stata/SE V.14 (StataCorp LP, College Station, Texas, USA).

Patient and public involvement

Neither patients nor members of the public were involved in this study.

RESULTS

Of an initial sample of 512 190 individuals, 74 931 had missing data on at least one of the model covariates, and 179 405 had self-reported physician-diagnosed heart attack, stroke, hypertension (additionally, systolic ≥ 140 mm Hg or diastolic ≥ 90 mm Hg), diabetes (additionally, fasting glucose levels ≥ 126 mg/dL) or cancer at baseline, respectively. Excluding these individuals resulted in a final sample for analysis of 257 854 individuals (figure 1).

Individuals provided up to seven measures of exercise frequency and each confounder (ie, baseline plus six repeated measures). Participants' characteristics at baseline are summarised in table 1. Online supplementary table 2 summarises participants' characteristics at each repeat assessment. Individuals in the categories of 1–2, 3–4 or 5–6 times/week of exercise were slightly younger, but showed higher proportions of family history of disease

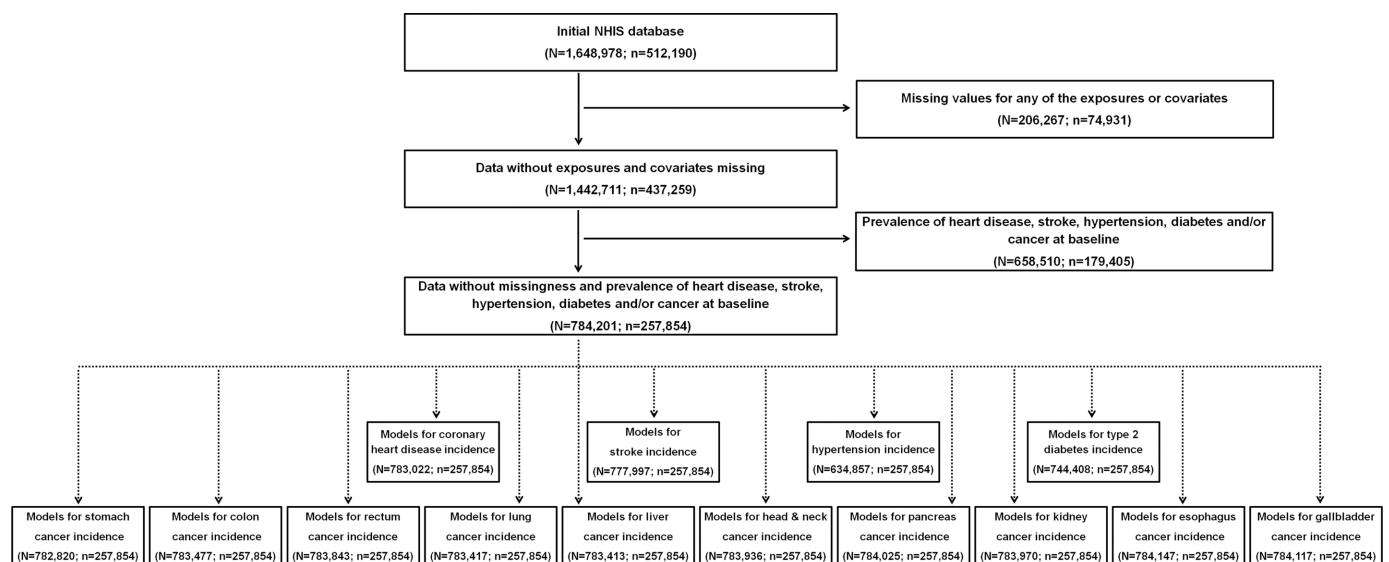


Figure 1 A flow diagram. Note: ‘N’ indicates numbers of total observations (ie, participants who provided repeated measures are treated as separate observations) and ‘n’ indicates numbers of unique participants at baseline. Data without missingness and prevalence of major diseases were used to create final analysis datasets for different incident disease outcomes; while the number of unique participants at baseline is the same for all incident disease outcomes, the total number of observations varied due to the nature of time-updated covariate analyses (ie, censoring of subsequent time-updated covariates when an incident disease case occurs before the end date of repeated measures). Data were obtained from a prospective cohort, which has been established by linking physical examination data of over half a million South Korean adults (2002–2015) with the entire South Korean population’s health insurance system. NHIS, National Health Insurance Service.

Table 1 Characteristics of the participants at baseline

Variables	All (n=257854)	Exercise frequency				
		None (n=148284)	1–2 times/week (n=62923)	3–4 times/week (n=24836)	5–6 times/week (n=6676)	Almost everyday (n=15135)
Sex, %						
Men (%)	50.5	43.8	63.3	57.5	52.9	49.6
Women (%)	49.5	56.2	36.7	42.5	47.1	50.4
Age, years	50.7 (8.7)	51.5 (9.2)	49.0 (7.5)	49.3 (7.5)	49.9 (7.8)	53.5 (9.3)
Body mass index, kg ² /m	23.5 (2.8)	23.3 (2.9)	23.6 (2.7)	23.7 (2.6)	23.7 (2.6)	23.7 (2.7)
Systolic blood pressure, mm Hg	116.8 (11.2)	116.6 (11.3)	117.1 (11.0)	116.8 (11.1)	116.8 (11.2)	117.3 (11.3)
Diastolic blood pressure, mm Hg	73.4 (7.9)	73.1 (8.0)	73.8 (7.8)	73.5 (7.9)	73.4 (8.0)	73.4 (7.9)
Fasting glucose levels, mg/dL	90.2 (12.3)	90.1 (12.4)	90.4 (12.2)	89.9 (11.9)	90.2 (12.3)	90.2 (12.4)
Total cholesterol, mg/dL	197.0 (36.7)	196.5 (37.1)	197.7 (36.2)	197.8 (35.9)	197.8 (36.1)	197.4 (36.8)
Family history of heart disease, stroke or hypertension, %	12.2	10.8	13.9	15.9	16.5	10.9
Family history of cancer, %	14.2	13.1	15.3	17.2	16.6	14.5
Family history of diabetes, %	6.3	5.4	7.2	8.5	9.2	5.7
Smoking status, %						
Never (%)	68.1	71.9	59.1	65.2	68.8	71.7
Previously (%)	8.4	5.8	12.0	12.7	12.5	8.8
Currently (%)	23.6	22.2	28.9	22.1	18.8	19.5
Alcohol consumption, %						
Never (%)	58.4	64.5	47.6	50.5	52.2	59.3
2–3 times/month (%)	16.4	13.8	21.3	19.7	18.8	14.1
1–2 times/week (%)	15.8	12.4	22.0	20.2	18.1	14.5
≥3 times/week (%)	9.5	9.3	9.1	9.5	10.8	12.1
Incident myocardial infarction, n (%)	3047 (1.2)	1741 (1.2)	723 (1.1)	276 (1.1)	88 (1.3)	219 (1.4)
Incident stroke, n (%)	16134 (6.3)	9689 (6.5)	3333 (5.3)	1482 (6.0)	390 (5.8)	1240 (8.2)
Incident hypertension, n (%)	120203 (46.6)	65964 (44.5)	30623 (48.7)	12617 (50.8)	3294 (49.3)	7705 (50.9)
Incident type 2 diabetes, n (%)	50459 (19.6)	27128 (18.3)	10666 (6.5)	5421 (21.8)	1399 (21.0)	3285 (21.7)
Incident stomach cancer, n (%)	4788 (1.9)	2672 (1.8)	13226 (21.0)	489 (2.0)	139 (2.1)	328 (2.2)
Incident colon cancer, n (%)	2711 (1.1)	1424 (1.0)	1160 (1.8)	314 (1.3)	90 (1.3)	191 (1.3)
Incident rectum cancer, n (%)	1494 (0.6)	809 (0.6)	692 (1.1)	154 (0.6)	46 (0.6)	107 (0.7)
Incident lung cancer, n (%)	3601 (1.4)	2138 (1.4)	796 (1.3)	307 (1.2)	85 (1.3)	275 (1.8)
Incident liver cancer, n (%)	2620 (1.0)	1423 (1.0)	680 (1.1)	263 (1.1)	75 (1.1)	179 (1.2)
Incident pancreas cancer, n (%)	864 (0.3)	483 (0.3)	205 (0.3)	92 (0.4)	24 (0.4)	60 (0.4)

Continued

Table 1 Continued

Variables	All (n=257854)	Exercise frequency				
		None (n=148284)	1–2 times/week (n=62923)	3–4 times/week (n=24836)	5–6 times/week (n=6676)	Almost everyday (n=15135)
Incident head and neck cancer, n (%)	656 (0.3)	377 (0.3)	144 (0.2)	73 (0.3)	15 (0.2)	47 (0.3)
Incident kidney cancer, n (%)	589 (0.2)	301 (0.2)	153 (0.2)	75 (0.3)	16 (0.2)	44 (0.3)
Incident gallbladder cancer, n (%)	400 (0.2)	219 (0.1)	83 (0.1)	43 (0.2)	16 (0.2)	39 (0.3)
Incident oesophagus cancer, n (%)	352 (0.1)	214 (0.1)	75 (0.1)	29 (0.1)	6 (0.09)	28 (0.2)
Median follow-up period, years (IQR)	13.0 (12.2, 13.3)	13.0 (12.2, 13.3)	13.0 (12.2, 13.3)	13.0 (12.2, 13.3)	12.7 (12.2, 13.3)	12.6 (12.2, 13.3)

Values presented are means unless indicated as an 'n'. Values in parentheses are SD unless otherwise indicated. Data were obtained from a prospective cohort, which has been established by linking physical examination data of over half a million South Korean adults (2002–2015) with the entire South Korean population's health insurance system.

and lower proportions of never smoking or drinking alcohol, compared with those in the categories of none or almost every day of exercise. Across the seven time points (online supplementary figure 1), the proportion of individuals who reported no exercise decreased while the proportion who reported 1–2 or 3–4 times/week of exercise increased; there were no noticeable changes for the categories of 5–6 times/week or almost every day of exercise.

Overall, J-shaped associations were found between exercise frequency and incident myocardial infarction, stroke, hypertension and type 2 diabetes. HRs for these diseases were lowest in the middle categories of exercise frequency (eg, 3–4 or 5–6 times/week) (figure 2). There were no associations for the most frequent exercise category (eg, almost every day) with the incidence of myocardial infarction, stroke and type 2 diabetes.

J-shaped associations were also found for incident stomach, lung, liver and head and neck cancers (figure 3). Higher exercise frequencies (eg, 1–2, 3–4 times/week and almost every day) were associated with lower hazards of incident stomach cancer. No statistical significance was observed for incident colon, rectum, pancreas, kidney, gallbladder and oesophagus cancers. Crude event rates per 100 000 person years in the middle categories of exercise frequency were relatively lower for incident rectum, and oesophagus cancers, but higher for incident pancreas, kidney and gallbladder cancers. Cox regression models with no adjustment for confounders (online supplementary figure 2) and a sensitivity analysis (online supplementary figure 3) in which incident cases occurring in the first 2 years of follow-up were removed both revealed nearly identical patterns of associations as the main analyses.

Figure 4 shows comparisons of results that showed statistical significance for multiplicative interaction terms between exercise frequency and each incident disease outcome. Strong J-shaped associations for incident

hypertension were identified at each level of body mass index. J-shaped associations of exercise frequency with incident hypertension were strong only in the more favourable levels of smoking (eg, never, previously) and alcohol consumption (eg, never, 2–3 times/month, 1–2 times/week); no or weak associations were identified in the most harmful level of smoking (eg, current smokers) and alcohol consumption (eg, ≥ 3 times/week). J-shaped associations were evident at all levels of family history of CVD and sex for incident hypertension, and sex for incident type 2 diabetes. Exercise frequency was associated with incident lung cancer in non-obese individuals, but there was no evidence of association in obese individuals. All comparisons stratified by each potential effect modifier are presented in online supplementary figures 4 and 5.

DISCUSSION

This is the first investigation examining the prospective associations of exercise with various incident non-communicable disease outcomes using multiple repeated measures of covariates in East Asian populations. We identified J-shaped associations of sweat-inducing exercise with incident myocardial infarction, stroke, hypertension, type 2 diabetes, stomach, lung, liver and head and neck cancers, with the greatest benefits being observed in the middle categories of exercise frequency (eg, 3–4 or 5–6 times/week): 1–2 times/week for head and neck cancer. These findings provide two important clinical and public health implications. First, prevention and management of non-communicable diseases in East Asians may benefit considerably from employing an exercise promotion approach in the context of combined non-communicable disease prevention. Mechanism research indicates that CVD and type 2 diabetes have similar biological pathways relating to exercise,^{22 23} so an integrated prevention approach can be applied to control and manage

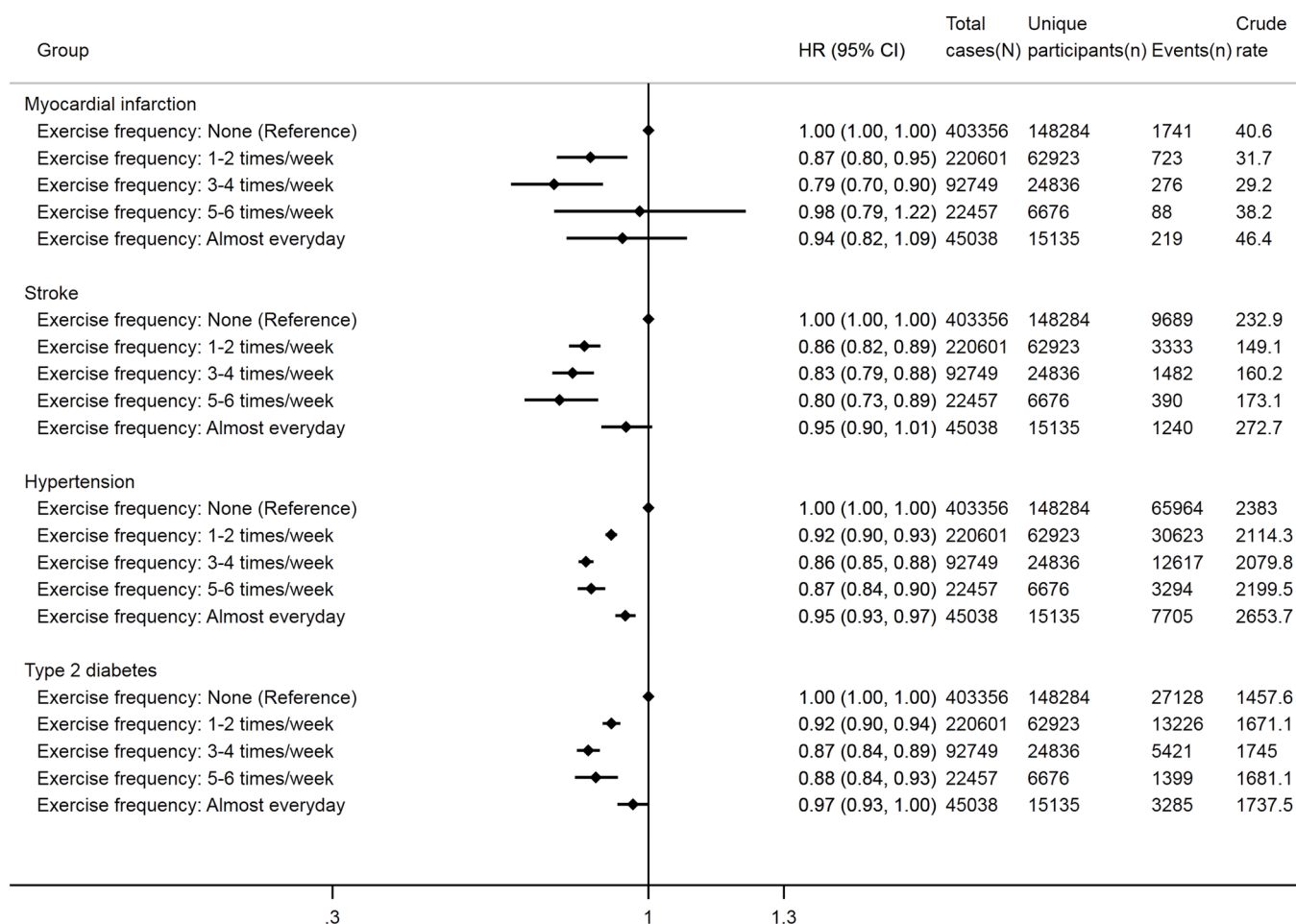


Figure 2 Associations of exercise frequency with various incident cardiovascular disease outcomes. Cox regression models with age as the underlying timescale were adjusted for sex, body mass index, systolic blood pressure, fasting glucose levels, total cholesterol levels, family history of heart disease, stroke or hypertension (in models for myocardial infarction, stroke and hypertension) or diabetes (in models for type 2 diabetes), smoking status and alcohol consumption. Crude rates are per 100 000 person years. 'N' indicates numbers of total observations (ie, participants who provided repeated measures are treated as separate observations) and 'n' indicates numbers of unique participants at baseline. Data were obtained from a prospective cohort, which has been established by linking physical examination data of over half a million Korean adults (2002–2015) with the entire South Korean population's health insurance system.

these two diseases at a minimum.⁵ Moreover, regular participation in exercise can induce favourable changes in intermediate cardiometabolic risk markers,²⁴ which are important predictors of typical non-communicable diseases. Hence, promoting exercise has great potential to act as an integrative behavioural strategy for preventing and controlling various non-communicable diseases simultaneously in East Asian populations.

Second, individuals who engage in exercise 3–4 or 5–6 times/week, rather than every day, may be able to reduce their risk of developing myocardial infarction, stroke, hypertension, type 2 diabetes, stomach, lung and liver cancers: 1–2 times/week for head and neck cancer. Similar J-shaped associations between high intensity exercise (eg, running) and CVD risk have also been reported in previous cohort studies of Western^{25 26} and

Japanese adults.²⁷ Nevertheless, the present study as well as previous research^{25–27} found that the risk of developing cardiovascular events in individuals who had the highest level of exercise was not noticeably higher compared with those who had the lowest level of exercise. No previous research in East Asians has found such J-shaped relationships between exercise or physical activity and other incident disease outcomes such as hypertension,^{28–32} diabetes^{33–40} and different type of cancers.^{41–47} However, previous meta-analyses of cohort studies comprising predominantly Westerners found leisure-time physical activity to have curvilinear (but not J-shaped) associations with the incidence of type 2 diabetes,⁴⁸ and linear associations with the incidence of hypertension⁴⁹ and various site-specific cancers (liver, lung, head and neck, kidney, colon, rectal, bladder, gastric cardia, breast, endometrial,

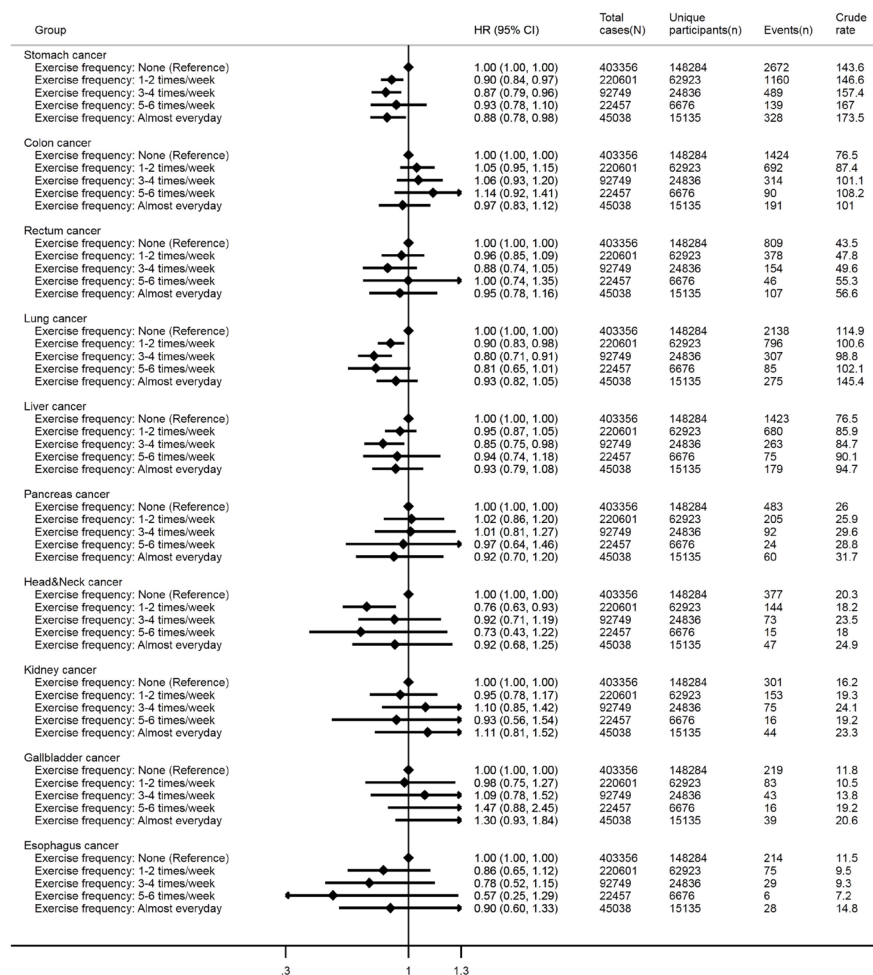


Figure 3 Associations of exercise frequency with various incident cancer outcomes. Cox regression models with age as the underlying timescale were adjusted for sex, body mass index, systolic blood pressure, fasting glucose levels, total cholesterol levels, family history of cancer, smoking status and alcohol consumption. Crude rates are per 100 000 person years. ‘N’ indicates numbers of total observations (ie, participants who provided repeated measures are treated as separate observations) and ‘n’ indicates numbers of unique participants at baseline. Data were obtained from a prospective cohort, which has been established by linking physical examination data of over half a million South Korean adults (2002–2015) with the entire South Korean population’s health insurance system.

myeloid leukaemia, myeloma, oesophageal adenocarcinoma).⁵⁰ While additional research is needed to confirm the J-shaped associations of exercise with various incident diseases in other samples of East Asians, findings of this research provide a strong rationale for the development and implementation of public health policies and clinical trials aimed at promoting a moderate level of sweat-causing exercise to minimise the risk of myocardial infarction, stroke, hypertension, type 2 diabetes, stomach, lung, liver and head and neck cancers.

Another finding of this research is that associations of sweat-inducing exercise with hypertension were modified by body mass index, smoking, alcohol consumption, family history of CVD and sex: lung cancer by body mass index and type 2 diabetes by sex. Notably, exercise frequency was not associated with hypertension in individuals who are smokers or drinking alcohol ≥ 3 times/week (except for 3–4 times/week of exercise). This observation provides some evidence that the harmful impacts of

smoking or binge drinking on hypertension^{51–53} may not be offset completely by exercise. This, in turn, appears to advocate for the need for implementing a combined hypertension prevention strategy targeting promotion of exercise in conjunction with smoking cessation and reductions in alcohol consumption in East Asians.¹³ For lung cancer, the null associations in individuals with body mass index ≥ 25 may be indicative of potential residual confounding through reported bias in smoking behaviours. Nonetheless, there was little evidence for effect modification for other disease comparisons, highlighting the importance of promoting exercise for the prevention of various non-communicable diseases in individuals at different levels of body mass index, smoking, alcohol consumption, family history of disease and sex.

This study has several notable strengths. First, we used data from a large prospective cohort study in which exercise and other risk markers were assessed on multiple occasions (up to seven times). Nearly 84% and 5% of the

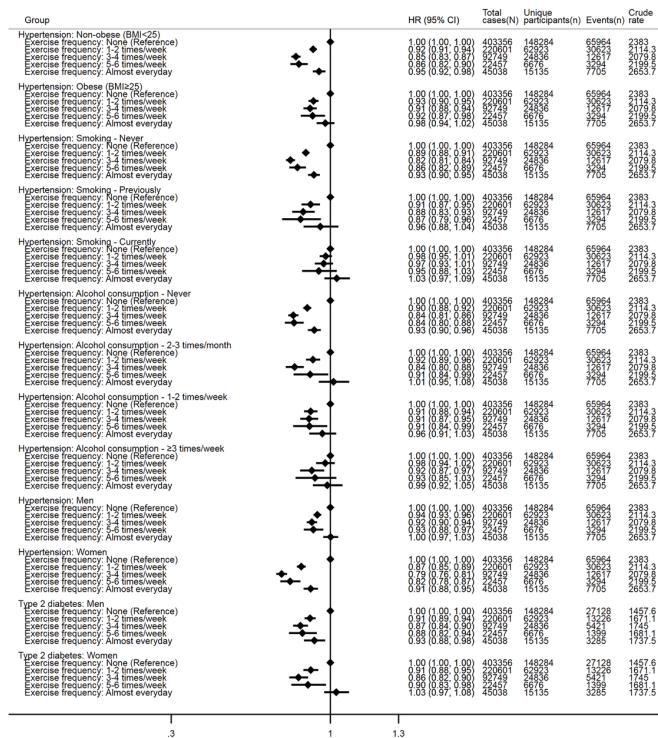


Figure 4 Results from assessment of effect modification of sex, body mass index (BMI), smoking, alcohol consumption and family history of disease in the associations between exercise frequency and each incident disease outcome. Only associations for which the multiplicative interaction terms were statistically significant are presented. Cox regression models with age as the underlying timescale were adjusted for sex (not in models for effect modification by sex), BMI (not in models for effect modification by BMI), systolic blood pressure, fasting glucose levels, total cholesterol levels, family history (not in models for effect modification by family history of respective disease) of heart disease/stroke/hypertension (in models for myocardial infarction, stroke and hypertension), diabetes (in models for type 2 diabetes) or cancer (in models for each cancer), smoking status (not in models for effect modification by smoking status) and alcohol consumption (not in models for effect modification by alcohol consumption). Crude rates are per 100 000 person years. ‘N’ indicates numbers of total observations (ie, participants who provided repeated measures are treated as separate observations) and ‘n’ indicates numbers of unique participants at baseline. P values for multiplicative interactions with exercise frequency are as follows; outcome—hypertension: BMI (0.049), smoking (<0.001), alcohol consumption (<0.001), family history of CVD (0.029) and sex (<0.001); outcome—lung cancer: BMI (0.016), and outcome—type 2 diabetes: sex (0.012). Data were obtained from a prospective cohort, which has been established by linking physical examination data of over half a million South Korean adults (2002–2015) with the entire South Korean population’s health insurance system.

full participants provided one and six repeated measures of all covariates, respectively. Compelling evidence indicates that the risk of regression dilution can be reduced using repeated measures of exposure and confounders.¹⁴ Moreover, we examined the dose–response relationship

of exercise frequency with a wide variety of specific types of incident non-communicable disease outcomes simultaneously using inpatient and outpatient diagnosis data as well as mortality data. The large sample size (n=257 854) is another strength.

This study has some limitations. Findings of this study may not be generalisable to adult populations of other ethnic origins. Due to the observational nature of this research, no strong causal inference can be drawn about the exercise-incident disease relationships. In addition, the accuracy of hospital admission records is uncertain, although the accuracy of death records from Statistics Korea was found to be 92% in previous research.⁵³ No information about medication use was available in the cohort data, so we could not use it as a potential confounder and another condition when defining disease status (eg, hypertension, type 2 diabetes) at both baseline and follow-up. Furthermore, no exercise duration was assessed; hence, inference was made purely based on exercise frequency. Moreover, ICD-10 codes for sex-specific cancers (eg, prostate and breast cancers) were masked due to the data management policy set forth by the NHIS, so it was not possible to examine such cancers in the present study. The lack of data on diet, which is another behavioural risk marker for non-communicable diseases,¹² is another limitation. Moreover, a sizeable proportion (n=74 931; 14.6%) of individuals were excluded due to the missing information on the covariates. Another limitation is that the measurement methods to assess the covariates were not standardised across the different medical institutes participating in the NHIS-HEALS cohort.

CONCLUSION

Individuals who engaged in sweat-inducing exercise around 3–6 times/week (as opposed to every day) generally had the lowest risk of developing myocardial infarction, stroke, hypertension, type 2 diabetes, stomach, lung, liver and head and neck cancers. These findings were generally applicable to different subpopulations as stratified by body mass index, smoking, alcohol consumption, family history of disease and sex. Public health and lifestyle interventions should promote a moderate level of sweat-inducing exercise as a behavioural strategy for prevention and control of non-communicable diseases in a wider population of East Asians.

Contributors YK designed this study, performed statistical analysis and drafted an initial version of the manuscript. SS, S-mH and SHJ all contributed to conceptualising the study idea and developing the analytical plans, and provided assistance with statistical analysis. All authors critically reviewed, approved the final version of the manuscript and agreed to be responsible for all facets of this work.

Funding This work was supported by the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (Grant 1631020 to SHJ), The Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (Grant H114C2686 to SHJ) and the Medical Research Council (Grant MC_UU_12015/1 to SS). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This research was approved by the Institutional Review Board (4-2017-0051) of the Yonsei University's Severance Hospital in Republic of Korea.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data sharing is not applicable because no informed consent for data sharing was obtained from the participants.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388:1459-544.
- Lakerveld J, Bot SD, van der Ploeg HP, *et al.* The effects of a lifestyle intervention on leisure-time sedentary behaviors in adults at risk: the Hoorn Prevention Study, a randomized controlled trial. *Prev Med* 2013;57:351-6.
- Roth GA, Johnson C, Abajobir A, *et al.* Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;70:1-25.
- Forouzanfar MH, Liu P, Roth GA, *et al.* Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm Hg, 1990-2015. *JAMA* 2017;317:165-82.
- World Health Organization. *World Health Organization Global report on diabetes*. Geneva: World Health Organization, 2016.
- Fitzmaurice C, Allen C, Barber RM, *et al.* Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA Oncol* 2017;3:524-48.
- Danaei G, Finucane MM, Lin JK, *et al.* National, regional, and global trends in systolic blood pressure since 1980: systematic analysis of health examination surveys and epidemiological studies with 786 country-years and 5.4 million participants. *Lancet* 2011;377:568-77.
- Roth GA, Huffman MD, Moran AE, *et al.* Global and regional patterns in cardiovascular mortality from 1990 to 2013. *Circulation* 2015;132:1667-78.
- GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017;390:E38.
- NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. *Lancet* 2016;387:1513-30.
- Ku PW, Fox KR, McKenna J, *et al.* Prevalence of leisure-time physical activity in Taiwanese adults: results of four national surveys, 2000-2004. *Prev Med* 2006;43:454-7.
- World Health Organization. *Global status report on noncommunicable diseases 2010*. Geneva: World Health Organization, 2011.
- Kokubo Y. Prevention of hypertension and cardiovascular diseases: a comparison of lifestyle factors in Westerners and East Asians. *Hypertension* 2014;63:655-60.
- Hutcheon JA, Chioleri A, Hanley JA. Random measurement error and regression dilution bias. *BMJ* 2010;340:c2289.
- Ng M, Fleming T, Robinson M, *et al.* Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014;384:766-81.
- Lear SA, Humphries KH, Kohli S, *et al.* Visceral adipose tissue accumulation differs according to ethnic background: results of the Multicultural Community Health Assessment Trial (M-CHAT). *Am J Clin Nutr* 2007;86:353-9.
- Chan JC, Malik V, Jia W, *et al.* Diabetes in Asia: epidemiology, risk factors, and pathophysiology. *JAMA* 2009;301:2129-40.
- Nguyen HN, Fujiyoshi A, Abbott RD, *et al.* Epidemiology of cardiovascular risk factors in Asian countries. *Circ J* 2013;77:2851-9.
- Seong SC, Kim YY, Park SK, *et al.* Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. *BMJ Open* 2017;7:e016640.
- Lee J, Lee JS, Park SH, *et al.* Cohort profile: the National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. *Int J Epidemiol* 2017;46:e15.
- Pedersen BK. Anti-inflammatory effects of exercise: role in diabetes and cardiovascular disease. *Eur J Clin Invest* 2017;47:600-11.
- Stern MP. Diabetes and cardiovascular disease. The "common soil" hypothesis. *Diabetes* 1995;44:369-74.
- WHO. *Global recommendations on physical activity for health*. Geneva: WHO, 2010.
- Lee DC, Pate RR, Lavie CJ, *et al.* Leisure-time running reduces all-cause and cardiovascular mortality risk. *J Am Coll Cardiol* 2014;64:472-81.
- Tanasescu M, Leitzmann MF, Rimm EB, *et al.* Exercise type and intensity in relation to coronary heart disease in men. *JAMA* 2002;288:1994-2000.
- Kubota Y, Iso H, Yamagishi K, *et al.* Daily total physical activity and incident stroke: the Japan Public Health Center-based prospective study. *Stroke* 2017;48:1730-6.
- Lu Y, Lu M, Dai H, *et al.* Lifestyle and risk of hypertension: follow-up of a young pre-hypertensive cohort. *Int J Med Sci* 2015;12:605-12.
- Sun Z, Zheng L, Detrano R, *et al.* Incidence and predictors of hypertension among rural Chinese adults: results from Liaoning province. *Ann Fam Med* 2010;8:19-24.
- Gu D, Wildman RP, Wu X, *et al.* Incidence and predictors of hypertension over 8 years among Chinese men and women. *J Hypertens* 2007;25:517-23.
- Lee S, Kuk JL, Katzmarzyk PT, *et al.* Cardiorespiratory fitness attenuates metabolic risk independent of abdominal subcutaneous and visceral fat in men. *Diabetes Care* 2005;28:895-901.
- Nakanishi N, Suzuki K. Daily life activity and the risk of developing hypertension in middle-aged Japanese men. *Arch Intern Med* 2005;165:214-20.
- Fan S, Chen J, Huang J, *et al.* Physical activity level and incident type 2 diabetes among Chinese adults. *Med Sci Sports Exerc* 2015;47:751-6.
- Joseph JJ, Echouffo-Tcheugui JB, Golden SH, *et al.* Physical activity, sedentary behaviors and the incidence of type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA). *BMJ Open Diabetes Res Care* 2016;4:e000185.
- Xu F, Ware RS, Tse LA, *et al.* Joint associations of physical activity and hypertension with the development of type 2 diabetes among urban men and women in Mainland China. *PLoS One* 2014;9:e88719.
- Xu F, Wang Y, Ware RS, *et al.* Joint impact of physical activity and family history on the development of diabetes among urban adults in Mainland China: a pooled analysis of community-based prospective cohort studies. *Asia Pac J Public Health* 2015;27:NP372-81.
- Shi L, Shu XO, Li H, *et al.* Physical activity, smoking, and alcohol consumption in association with incidence of type 2 diabetes among middle-aged and elderly Chinese men. *PLoS One* 2013;8:e77919.
- Villegas R, Shu XO, Li H, *et al.* Physical activity and the incidence of type 2 diabetes in the Shanghai women's health study. *Int J Epidemiol* 2006;35:1553-62.
- Lee DC, Park I, Jun TW, *et al.* Physical activity and body mass index and their associations with the development of type 2 diabetes in Korean men. *Am J Epidemiol* 2012;176:43-51.
- Honda T, Kuwahara K, Nakagawa T, *et al.* Leisure-time, occupational, and commuting physical activity and risk of type 2 diabetes in Japanese workers: a cohort study. *BMC Public Health* 2015;15:1004.
- Inoue M, Yamamoto S, Kurahashi N, *et al.* Daily total physical activity level and total cancer risk in men and women: results from a large-scale population-based cohort study in Japan. *Am J Epidemiol* 2008;168:391-403.
- Lee KJ, Inoue M, Otani T, *et al.* Physical activity and risk of colorectal cancer in Japanese men and women: the Japan Public Health Center-based prospective study. *Cancer Causes Control* 2007;18:199-209.
- Luo J, Iwasaki M, Inoue M, *et al.* Body mass index, physical activity and the risk of pancreatic cancer in relation to smoking status and history of diabetes: a large-scale population-based cohort study in Japan—the JPHC study. *Cancer Causes Control* 2007;18:603-12.
- Suzuki S, Kojima M, Tokudome S, *et al.* Effect of physical activity on breast cancer risk: findings of the Japan collaborative cohort study. *Cancer Epidemiol Biomarkers Prev* 2008;17:3396-401.
- Eaglehouse YL, Koh WP, Wang R, *et al.* Physical activity, sedentary time, and risk of colorectal cancer: the Singapore Chinese Health Study. *Eur J Cancer Prev* 2017;26:469-75.
- Mok Y, Jeon C, Lee GJ, *et al.* Physical activity level and colorectal cancer mortality. *Asia Pac J Public Health* 2016;28:638-47.
- Yun YH, Lim MK, Won YJ, *et al.* Dietary preference, physical activity, and cancer risk in men: national health insurance corporation study. *BMC Cancer* 2008;8:366.

47. 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.
48. Liu X, Zhang D, Liu Y, *et al.* Dose-response association between physical activity and incident hypertension: a systematic review and meta-analysis of cohort studies. *Hypertension* 2017;69:813.
49. Moore SC, Lee IM, Weiderpass E, *et al.* Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults. *JAMA Intern Med* 2016;176:816–25.
50. Stranges S, Wu T, Dorn JM, *et al.* Relationship of alcohol drinking pattern to risk of hypertension: a population-based study. *Hypertension* 2004;44:813–9.
51. Niskanen L, Laaksonen DE, Nyyssönen K, *et al.* Inflammation, abdominal obesity, and smoking as predictors of hypertension. *Hypertension* 2004;44:859–65.
52. Halperin RO, Gaziano JM, Sesso HD. Smoking and the risk of incident hypertension in middle-aged and older men. *Am J Hypertens* 2008;21:148–52.
53. Won T, Kang B, Im T, *et al.* The study of accuracy of death statistics. *J Korean Soc Emerg Med* 2007;18:256–62.