OPEN

Polygenic Score for Physical Activity Is Associated with Multiple Common Diseases

ELINA SILLANPÄÄ[1,2], TEEMU PALVIAINEN[2], SAMULI RIPATTI[2,3,4], URHO M. KUJALA[5], and JAAKKO KAPRIO[2]

[1]Gerontology Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, FINLAND; [2]Institute for Molecular Medicine Finland, HiLIFE, Helsinki, FINLAND; [3]Department of Public Health, University of Helsinki, Helsinki, FINLAND, University of Helsinki; [4]Broad Institute of MIT and Harvard, Cambridge, MA; and [5]Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, FINLAND

ABSTRACT

SILLANPÄÄ, E., T. PALVIAINEN, S. RIPATTI, U. M. KUJALA, and J. KAPRIO. Polygenic Score for Physical Activity Is Associated with Multiple Common Diseases. *Med. Sci. Sports Exerc.*, Vol. 54, No. 2, pp. 280–287, 2022. **Introduction:** Genetic pleiotropy, in which the same genes affect two or more traits, may partially explain the frequently observed associations between high physical activity (PA) and later reduced morbidity or mortality. This study investigated associations between PA polygenic risk scores (PRS) and cardiometabolic diseases among the Finnish population. **Methods:** PRS for device-measured overall PA were adapted to a FinnGen study cohort of 218,792 individuals with genomewide genotyping and extensive digital longitudinal health register data. Associations between PA PRS and body mass index, diseases, and mortality were analyzed with linear and logistic regression models. **Results:** A high PA PRS predicted a lower body mass index ($\beta = -0.025$ kg·m^{-2} per one SD change in PA PRS, SE = 0.013, $P = 1.87 \times 10^{-80}$). The PA PRS also predicted a lower risk for diseases that typically develop later in life or not at all among highly active individuals. A lower disease risk was systematically observed for cardiovascular diseases (odds ratio [OR] per 1 SD change in PA PRS = 0.95, $P = 9.5 \times 10^{-19}$) and, for example, hypertension [OR = 0.93, $P = 2.7 \times 10^{-44}$), type 2 diabetes (OR = 0.91, $P = 4.1 \times 10^{-42}$), and coronary heart disease (OR = 0.95, $P = 1.2 \times 10^{-9}$). Participants with high PA PRS had also lower mortality risk (OR = 0.97, $P = 0.0003$). **Conclusions:** Genetically less active persons are at a higher risk of developing cardiometabolic diseases, which may partly explain the previously observed associations between low PA and higher disease and mortality risk. The same inherited physical fitness and metabolism-related mechanisms may be associated both with PA levels and with cardiometabolic disease risk.

Key Words: GENE, EXERCISE, HERITABILITY, TYPE 2 DIABETES

Noncommunicable diseases, such as cardiovascular disease, type 2 diabetes, and neoplasms, cause a large burden to society, with an estimated 1 billion cases existing worldwide. Therefore, early detection, prevention, and intervention regarding these diseases are fundamental goals in advancing human health and quality of life (1). To date, it is known that individual disease risk is a complex interplay of genetic susceptibility and multiple social, environmental, and policy factors.

Genetic risk estimate can be calculated at the time of birth and is therefore one of the earliest measurable contributors to overall disease risk during person's life span. Genetic contributions to complex traits and diseases, such as physical activity (PA) and cardiometabolic diseases (CMD), are polygenic (i.e., accounted for by a large number of causal variants with very small effects). Polygenic risk scores (PRS) summarize genomewide genotype data into single variables that produce individual-level risk scores regarding genetic liability. PRS already have been produced for several CMD traits (2–4). The single nucleotide polymorphisms and their associated weights summarized into CMD PRS have confirmed the existence of genetic influence on common disease risks previously reported in twin and family studies (2). A couple of studies have already evaluated the value and usability of CMD PRS in

Address for correspondence: Elina Sillanpää, Ph.D., Gerontology Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, P.O. Box 35 (VIV), FIN-40014, Finland; E-mail: elina.sillanpaa@jyu.fi.

Submitted for publication April 2021.
Accepted for publication August 2021.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/21/5402-0280/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000002788

clinical risk prediction. These studies have shown that CMD PRS may have additive value in improving the clinical risk prediction of CMD and that PRS may predict disease onset —especially among high-risk individuals (5). In contrast to CMD, PRS values for lifestyle factors are less frequently used. Multiple twin studies suggest that human behavior is moderately genetically regulated (6), and PRS can be calculated to any heritable trait. In year 2020, we published two PRS for PA and showed their significant out-of-sample predictive values in two independent cohorts with different PA phenotypes (7).

The existence of polygenic influences on both PA and CMD suggests that gene–environment interplay and reverse causality may play a role in associations between PA, CMD, and mortality (8–11). PA has suggested to be a cost-effective strategy for the prevention of CMD (12), as observational epidemiological studies show that high PA levels strongly predict lower disease risk and all-cause and cause-specific mortality (13,14). Clinical trials and field interventional studies have not provided strong evidence for or against a causal role of PA in mortality or cardiac diseases (15). PA has generally beneficial—but relatively modest—effects on selected biological risk factors for disease, such as improvements in blood lipid levels, blood pressure, and glucose metabolism (16,17).

The latest animal and human findings challenge the assumption regarding the causal association between higher PA and reduced mortality risk later in life (18–20). When genetic factors are fully controlled, twin studies suggest that PA does not reduce mortality risk (18). It has been suggested that genetic pleiotropy, where the same genes affects two or more characters, may partially explain the frequently observed associations between high PA and reduced mortality risk later in life (8,18), but evidence from human studies is limited.

Genetic confounding occurs when a genetic variant or set of variants causally affect both the risk factors and the outcomes (e.g., variants associated with CMD risk factors, such as PA, also directly affect CMD). This causes challenges in observational epidemiology, as adjusting for genetic confounders is typically insufficient in analysis (21). Another potential source of bias is gene–environment interaction. Although the effect of genetic inheritance on PA is generally poorly understood, it is assumed that individuals with favorable genotypes tend to participate in PA. Studies suggest that these individuals may have inherited better cardiorespiratory fitness (22), derive greater pleasure from PA, and that their personality and other behavioral characteristics makes it easier for them to adopt and follow a physically active lifestyle (23). This gene–environment interaction hinders causal reasoning because the environment or lifestyle experienced by an individual is partly influenced by their genotypes. The PRS construct offers tools to answer confounding and gene–environment challenges in observational research. Individual PRS can be used to explore potential genetic overlap in two or more traits (such as PA and CMD) as well as to predict other traits in a regression model across a study sample.

This study investigated associations between PA PRS, CMD, and mortality among the Finnish population study cohort FinnGen of 218,792 individuals.

METHODS

Study Sample and End Points

The data comprised 218,792 Finnish citizens from FinnGen, Data Freeze 5. The sample included 56.5% women, and the mean age was 59.8 yr (range, 1.5–120.3). FinnGen includes prospective epidemiological cohorts, diseases-based cohorts, and hospital biobank samples (see Table, Supplemental Digital Content 1, List of FinnGen Data Freeze 5 cohorts, http://links.lww.com/MSS/C426). In FinnGen, genome information is combined with national hospital discharge (1968–present), death (1969–present), cancer (1953–present), and the Social insurance Institute of Finland (Kela) medication reimbursement (1995–present) registers. End point definitions were based on the *International Statistical Classification of Diseases and Related Health Problems* (ICD-8, ICD-9, and ICD-10) codes. The ICD codes included in each end point can be revised at FinnGen Web pages (https://www.finngen.fi/en/researchers/clinical-endpoints, DF-5). The quality of the CMD diagnoses in these registers has been extensively validated in several studies (24). For example, health care data included 21,012 major coronary heart disease events, 55,970 hypertension cases, and 29,139 T2D cases (Tables 1–3). We also tested associations with PA PRS and CMD medication end points. Based on the three existing smoking status variables in FinnGen, 40.2% (n = 87,859) had missing smoking status data, whereas 59.8% (n = 130,933) could be classified. Of the latter, 22.9% (29,961) were current smokers, 24% (31,471) were former smokers, 51.1% (66,872) were never smokers, and 2.0% (2629) were noncurrent smokers (it was not known whether they were former smokers or never smokers). Occasional smokers were considered current smokers. Based on this information, we decided to use data from current smokers and never smokers in our analysis.

Genotyping, Quality Control, and Imputation

The FinnGen Study samples were genotyped with various Illumina and custom AxiomGT1 Affymetrix arrays (Thermo Fisher Scientific, Santa Clara, CA; please see http://www.finngen.fi/en/researchers/genotyping and Supplemental Digital Content 2 (see Document, Genotyping and quality control of the FinnGen data, http://links.lww.com/MSS/C427).

Polygenic scoring for PA. The PA PRS, which was recently developed for continuous accelerometer-based overall PA volume (7,25), was adapted to the FinnGen cohort. Briefly, GWAS summary statistics from the UK Biobank for risk score calculation were obtained from the data-sharing repository of the GWAS of PA measured by an accelerometer (Fig. 1) (11). The objective assessment of PA was measured for a 7-d period using an Axivity AX3 wrist-worn triaxial accelerometer in the UK Biobank cohort (n = 103,702). The nonwear time was detected and imputed by the expert working group, resulting in a total PA calculated by averaging all worn and imputed values (11,25). To obtain PRS for PA, we used a

TABLE 1. Association analysis between polygenic score for PA and metabolic end points as well as selected control conditions.

Phenotype	*N* Cases/Controls	OR	*P*	β	SE
Obesity	8,908/209,827	0.90	5.8×10^{-20}	–0.1011	0.0108
Type 2 diabetes	29,193/182,573	0.91	4.1×10^{-42}	–0.0906	0.0066
Type 2 diabetes without complications	14,622/183,185	0.91	1.0×10^{-27}	–0.0978	0.0088
Type 2 diabetes with complications	24,133/183,185	0.91	1.9×10^{-42}	–0.0984	0.0071
Type 2 diabetes with peripheral circulatory complications	1,049/183,185	0.91	0.0027	–0.0987	0.0310
Type 2 diabetes medication (all types)	32,897/185,820	0.92	5.3×10^{-40}	–0.0822	0.0061
Type 2 diabetes medication (other than insulin)	28,493/185,895	0.91	2.7×10^{-44}	–0.0938	0.0066
Diabetes, insulin treatment (Kela reimbursement)	29,071/189,721	0.92	2.8×10^{-36}	–0.0820	0.0064
Hypothyroidism (congenital or acquired)	26,342/59,827	0.92	1.8×10^{-29}	–0.0803	0.0085
Hypothyroidism, strict autoimmune	22,997/175,475	0.96	1.9×10^{-8}	–0.0416	0.0072
Disorders of lipoprotein metabolism/other lipidemias	14,010/197,259	0.96	8.3×10^{-6}	–0.0413	0.0090
Hyperlipidemia, other/unspecified	4,535/197,259	0.94	0.0004	–0.0566	0.0152
Nonalcoholic fatty liver disease	894/217,898	0.89	0.0014	–0.1128	0.0335
Cholelithiasis	19,023/195,144	0.93	3.7×10^{-18}	–0.0689	0.0078
Statin medication	68,782/150,010	0.95	3.8×10^{-25}	–0.0559	0.0053
Sleep apnea	16,761/201,194	0.93	3.7×10^{-20}	–0.0762	0.0081
Osteopathies and chondropathies	9,217/209,575	1.00	0.69	–0.0050	0.0107
Osteoporosis	3,203/209,575	0.99	0.62	–0.0104	0.0180
Arthrosis	37,233/147,221	1.00	0.78	0.0020	0.0062
Gonarthrosis	22,796/147,221	0.99	0.12	–0.0128	0.0075
Smoking	29,961/66,872	0.95	2.8×10^{-9}	–0.0507	0.0071

Logistic regression analysis. Model adjusted for age, sex, and 10 genetic principal components of population stratification. Kela, the Social insurance Institute of Finland.

Bayesian approach, accounted for linkage disequilibrium (LDPred) (27), and adjusted for the LD reference panel of unrelated Finnish individuals from the national FINRISK study (n = 27,284) (26). The total number of variants used for risk score calculation in our first analyses was 1,140,182.

Statistical analyses. Associations between PA PRS and body mass index (BMI), common diseases, and mortality were analyzed with linear and logistic regression models adjusted for age, gender, and the 10 principal components of ancestry. An increase in risk was calculated per 1 SD change in PRS. The distribution of 4 SD covered 95% of the population. The number of diseased cases varied between 894 (nonalcoholic fatty level disease) and 111,108 (cardiovascular diseases; Tables 1–3). The false discovery rate was used to correct the *P* values for multiple testing (28), and the significance threshold was set to $P < 0.05$.

Ethics Approval

The patients and control subjects in FinnGen provided their informed consent for biobank research based on the Finnish Biobank Act. Alternatively, older research cohorts, collected before the start of FinnGen (in August 2017), were collected based on study-specific consent and later transferred to the Finnish biobanks after approval by Fimea, the National Supervisory Authority for Welfare and Health. The recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) approved the FinnGen study protocol Nr HUS/990/2017.

The FinnGen study is approved by the Finnish Institute for Health and Welfare (permit nos. THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019, THL/1524/5.05.00/2020, and THL/2364/14.02/2020), Digital and Population Data Service Agency (permit nos. VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit nos. Kela 58/522/2017, Kela 131/522/2018, Kela 70/522/2019, Kela 98/522/2019, Kela 138/522/2019, Kela 2/522/2020, and Kela 16/522/2020), and Statistics Finland (permit nos. TK-53-1041-17 and TK-53-90-20).

The Biobank Access Decisions for FinnGen samples and the data used in FinnGen Data Freeze 6 include the following: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26,

TABLE 2. Association analysis between polygenic score for PA and cardiovascular disease (CVD) end points.

Phenotype	Cases/Controls	OR	*P*	β	SE
CVD, all	111,108/107,684	0.96	9.5×10^{-19}	–0.0445	0.0049
Coronary atherosclerosis	23,363/195,429	0.95	4.4×10^{-11}	–0.0514	0.0076
Ischemic heart diseases	30,952/187,840	0.96	1.1×10^{-10}	–0.0447	0.0068
Angina pectoris	18,168/200,624	0.95	2.5×10^{-10}	–0.0540	0.0083
Myocardial infarction	11,622/207,170	0.96	1.3×10^{-5}	–0.0450	0.0100
Major CVD event	21,012/197,780	0.95	1.2×10^{-9}	–0.0491	0.0079
Hard CVD	29,350/189,442	0.96	4.9×10^{-10}	–0.0441	0.0069
Coronary revascularization	12,271/206,521	0.93	4.7×10^{-13}	–0.0685	0.0099
All-cause heart failure	23,397/194,811	0.94	2.7×10^{-14}	–0.0576	0.0074
Stroke	18,661/162,201	0.95	6.4×10^{-8}	–0.0466	0.0084
Hypertension	55,955/162,837	0.93	2.7×10^{-44}	–0.0777	0.0055
Antihypertensive medication	107,287/111,505	0.94	2.4×10^{-34}	–0.0616	0.0050

Logistic regression analysis. Model adjusted for age, sex, and 10 genetic principal components of population stratification.

TABLE 3. Association analysis between polygenic score for PA, dementia end points, and death.

Phenotype	Cases/Controls	OR	*P*	β	SE
Dementia, all	7,284/209,487	1.01	0.4743	0.0108	0.0128
Alzheimer disease	3,899/214,893	1.05	0.0112	0.0462	0.0170
Alzheimer's disease, atypical or mixed	800/214,893	1.09	0.0262	0.0863	0.0361
Alzheimer's disease, early onset	587/111,471	1.08	0.1199	0.0724	0.0419
Alzheimer's disease, late onset	2,670/111,471	1.02	0.3794	0.0218	0.0216
Vascular dementia	859/211,300	0.99	0.7712	–0.0120	0.0346
Any death	15,152/203,640	0.97	0.0003	–0.0335	0.0088

Logistic regression analysis. Model adjusted for age, sex, and 10 genetic principal components of population stratification.

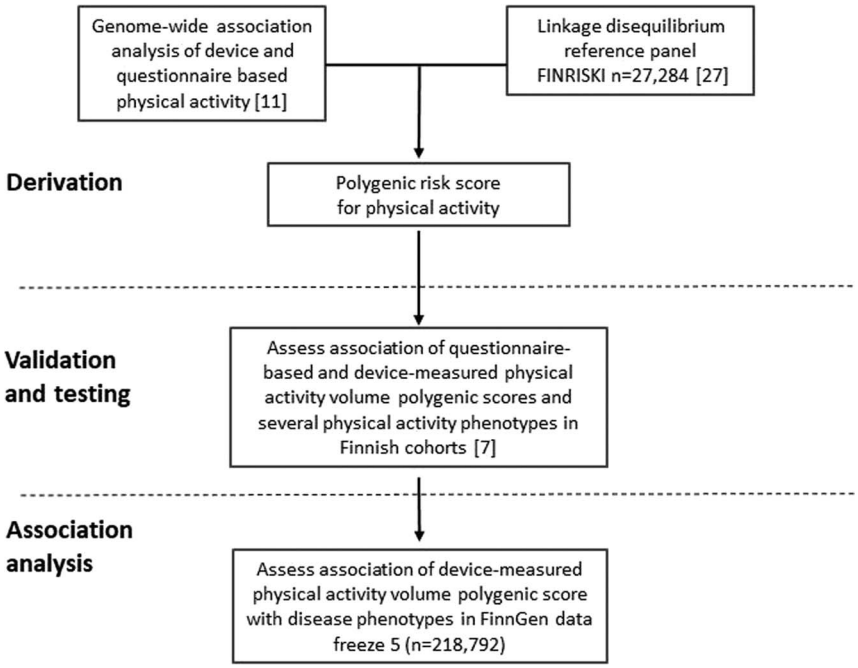

FIGURE 1—Study design and workflow. PRS for device-based measures of overall PA volume was derived from recent GWA study (11) and a linkage disequilibrium reference panel of 27,284 unrelated Finnish individuals (26). Out-of-sample predictive value was tested using two independent Finnish cohorts and several PA phenotypes (7). Association analysis was conducted in a FinnGen cohort of 218,792 Finnish participants. The clinical end points used in the analysis were derived from Finnish nationwide digital health registers.

BB2020_1, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, Auria Biobank AB17–5154, Biobank Borealis of Northern Finland_2017_1013, Biobank of Eastern Finland 1186/2018, Finnish Clinical Biobank Tampere MH0004, Central Finland Biobank 1–2017, and Terveystalo Biobank STB 2018001. The study conducted in accordance with the declaration of Helsinki.

RESULTS

We observed that a high PA PRS, a genetic inheritance that supports higher volumes for PA, systematically predicted a lower risk for diseases that typically develop later in life or not at all among highly active individuals.

Polygenic risk, BMI, and obesity. High PA PRS predicted lower BMI (β = −0.025 kg·m^{-2} per SD of PA PRS, SE = 0.013, $P = 1.87 \times 10^{-80}$, n = 160,334) and body weight (β = −0.876 kg per SD of PA PRS, SE = 0.041, $P = 3.69 \times 10^{-102}$, n = 164,964). Higher PA PRS also predicted a decreased risk for obesity (Table 1).

Polygenic risk, type 2 diabetes, and other metabolic diseases. Genetically active participants were at a significantly lower risk for developing type 2 diabetes (odds ratio [OR] per SD of PA PRS = 0.91, $P = 4.1 \times 10^{-42}$). Similar OR values were observed in both uncomplicated type 2 diabetes cases and cases where peripheral complications were present (Table 1).

Higher PA PRS values were also associated with lower odds of diabetes medication and insulin treatment end points (based on Kela reimbursement). Genetically active participants also exhibited lower risk for several diseases related to fat metabolism, such as disorders of lipoprotein metabolism, lipidemias, nonalcoholic fatty liver disease, and use of statin medication (OR = 0.89 to 0.96). Lower odds for hyperthyroidism were also observed among the genetically active participants (OR = 0.96, $P = 1.8 \times 10^{-29}$). However, associations between PA PRS and bone metabolism phenotypes (arthrosis, osteoporosis, etc.) were not found. High PA PRS value was associated with lower odds for smoking.

Polygenic risk and cardiovascular diseases. High PA PRS systematically associated with smaller cardiovascular disease (CVD) risk (Table 2). The genetically active participants had fewer overall CVD (OR = 0.96, $P = 9.5 \times 10^{-19}$). They also had a lower risk of ischemic heart disease (OR = 0.96, $P = 1.1 \times 10^{-10}$), stroke (OR = 0.95, $P = 6.4 \times 10^{-8}$), and hypertension (OR = 0.93, $P = 2.7 \times 10^{-44}$), and they also used fewer antihypertensive medications (OR = 0.94, $P = 2.4 \times 10^{-34}$).

Polygenic risk, mortality, and dementia. In FinnGen cohort, which included 15,152 deaths, a 1 SD increase in the PA PRS was found to be associated with lower odds for all-cause mortality (OR = 0.97, P = 0.0003). The risk of Alzheimer's disease was, however, increased (OR = 1.05, P = 0.0112), although the number of cases in the FinnGen data was rather modest (n = 3,899). The PA PRS value was not associated with vascular dementia (Table 3).

DISCUSSION

We adapted a polygenic score for device-based overall PA volume (7) and showed that genetically less physically active

persons are at higher risk of developing several CMD and phenotypes when compared with persons with a genetic predisposition for high PA. Furthermore, the risk of all-cause mortality was higher among genetically less active individuals. Our results suggest that genetic pleiotropy, i.e., the same genes affecting both PA behavior and CMD risk, may partly explain the associations between low PA and higher disease and mortality risk that have frequently reported in literature. In addition, possibly because genetically active persons tend to live longer, it was observed that they are at higher risk of developing Alzheimer's disease.

There is a substantial amount of literature on the associations between higher PA and a lower risk of common CMD. Moreover, PA has been reported to reduce disease risk in a dose–response manner (29). However, adjusting for genetic confounders in these designs has been impossible except for in twin studies. PA is a multifactorial behavioral and physiological trait influenced by hundreds—if not thousands—of genes that exhibit variation due to hundreds of thousands of genetic variants, most of which are single-nucleotide variants. To our knowledge, this is the first study that has used PA PRS to demonstrate that a genetic predisposition for PA is associated with several cardiometabolic traits. The results are in line with a few twin studies, which have shown that genetic inheritance strongly mediates the association between PA and disease risk and mortality (30,31) and genetically informative studies suggesting a bidirectional relationship with PA and adiposity and CMD (11,32). Our study design cannot distinguish between horizontal pleiotropy and vertical pleiotropy, i.e., the mediating role of PA in association between PA genotype and CMD. This does not essentially change our conclusion on that genetic predisposition may partly explain the association seen between the traits in observational follow-ups. However, our findings should not be interpreted as not to recommend exercise to patients as exercise therapy provides many health benefits (33).

The PA PRS, derived from over million single nucleotide polymorphisms, does not reflect a genetic predisposition to single underlying mechanism but rather the combined influence of multiple pathways. Genetic risk can lead to activity behavior through multiple distinct social-behavioral factors (34). In addition, complex biological systems regulated by genetics, but also influenced by external environment and epigenetic regulation, also play a significant role in regulating PA levels (35). These systems—for example, the brain, the cardiorespiratory system, and the muscles—also interact within the human body. The optimal function of these organs and systems that participate in energy production predicts higher aerobic capacity among individuals representing all age-groups. Aerobic capacity is a highly heritable phenotype according to twin studies (22). By contrast, many pathological conditions related to these functions, such as impaired heart and liver function, vascular structure, and circulatory and metabolic systems, encompass in CMD. It is therefore reasonable to suggest that shared genetic factors that regulate both PA and CMD may have a physiological basis.

Recent studies using newly discovered genetic variants for CMD as well as novel methods of generating PRS that use genomewide variation rather than only a handful of genomewide significant variants have shown the polygenic nature of CMD and better performance for predicting disease onset (5). Improved risk prediction over commonly used clinical estimates has been especially found among high-risk individuals representing genetic extremities (36). The role that PRS will play in clinical care is currently unclear, but it has been constantly suggested that lifestyle interventions, including PA, might be one treatment option for high-risk CMD individuals (36). Yet, it is not known how high-risk individuals accept and respond to lifestyle treatment. In terms of BMI, it has been shown that the effects of an unhealthy diet, PA, and sedentary behavior on BMI are pronounced in those with a genetic predisposition for a high BMI (37,38). We observed here that the prevalence of CMD increased linearly with decreasing PA PRS. We hypothesize that it is likely that individuals who have high genetic or clinical risk for CMD may also have a "low activity genotype." This suggests that it can be challenging to intervene with formerly inactive individuals to prevent cardiometabolic diseases that typically develop over decades. The first lifestyle studies that have informed participants about their genetic risk for CMD have resulted in conflicting findings. Information about genetic risk for type 2 diabetes did not result in increases in PA even in short term intervention (39). On the contrary, Web page communication about genetic and traditional atherosclerotic cardiovascular disease risk motivated 42.6% of the high-risk individuals to improve their health behavior, which resulted in clinically significant improvement of lipid profiles and lower systolic blood pressure during a 1.5-yr follow-up (40). The number of individuals that took some actions to improve their health during the follow-up was significantly higher compared with the individuals who were at average or lower risk (33.5%). It was observed that both traditional and genetic risk contributed to the change in health behavior independently. It must also to be noted that exercise training intervention effects on intermediate CMD risk factors are highly individual and suggested to originate from genetic diversities (17,41). Future studies investigating lifestyle interventions among individuals representing genetic extremities may reveal how genetic inheritance affects intervention responses or lifestyle modifications.

In our study, a higher PA PRS, which suggests a genetically active genotype, was associated with a higher risk for developing Alzheimer's disease. This association was evident especially in atypical or mixed Alzheimer's diseases diagnosis, but associations with all dementia diagnoses or vascular dementia were not found. In general, PA has typically been suggested to reduce the risk of Alzheimer's disease and its brain-related complications (42). A healthy lifestyle has been found to be associated with lower dementia risk—independent of one's genetic risk level (43). However, some recent findings are better in line with our results. For example, a 28-yr follow-up study by Sabia et al. (44) found no evidence about the neuroprotective effect of PA and states that earlier findings

showing a lower risk of dementia in physically active people may be attributable to reverse causation. The risk of Alzheimer's disease increases strongly with increasing age, but it has been probably overall decreasing in recent years as the overall health of the elderly population has improved globally (45). In the Finnish population, 15%–20% of those over 85 yr of age have received an Alzheimer's diagnosis. We suggest that the association between PA PRS and Alzheimer's disease may partially be explained by improved survival of genetically physically active individuals because of the lower risk of early-onset cardiovascular death. However, a formal competing-risk analysis is needed to demonstrate whether this association is robust. Another potential explanation to our findings about increased Alzheimer risk among genetically active individuals is the genetic overlap of PA and neurodegenerative diseases. The potential causal association between different PA estimates and Alzheimer's disease has previously been tested by Doherty et al. (11) using comprehensive Mendelian randomization design and UK biobank data. These results did not suggest causal associations.

Thus far, PRS values have been established for various traits and diseases, but their development for behavioral traits has been limited (46,47). Moreover, earlier studies have not assessed how behavioral PRS predict future major health events—except with regard to alcohol consumption (48). In this study, we investigated associations between the polygenic score for PA and the multiple register-derived disease end points in a large population-based sample (n = 218,792). The Finnish personal identity numbers, unique for each individual and used in all national digital health care registers, enabled linking the genotyped FinnGen study participants into different health registers. This, in turn, allowed us to test the PA PRS with multiple validated ICD end points simultaneously. Our study provides novel information regarding associations between PA and common diseases and contributes to the interpretation of sports and exercise science studies.

The polygenic score for PA used in this study was derived using the UK biobank population representing European ancestry. UK biobank subjects are volunteers and somewhat healthier compared with the general population (49). However, in the original UK biobank GWAS data, there were many individuals with chronic diseases who were not excluded when constructing the PA PRS. It is known that chronic diseases are a common reason for one's reduction in PA (50); thus, the constructed PA PRS may include genetic variants that are primarily predictors of chronic diseases. This needs to be taken into account when interpreting our results. PA PRS value was validated using Finnish cohorts (7), and our current association analyses were performed among Finnish citizens. In general, Finns differ from other Europeans primarily in terms of the frequency of less common and rare variants because of genetic isolation and bottlenecks (51). This different genetic ancestry may limit the generalizability of the PA PRS to Finns. Although the PA PRS value was strongly associated with cardiometabolic traits in our study, the UK Biobank data used to derive PRS for PA may have resulted in an underestimation of the associations between PRS for PA and CMD. By contrast, a fraction of FinnGen individuals were collected through hospital biobanks, which may lead to the overestimation of risk (5). The PRS for PA and its associations with CMD needs to also be tested in non-European samples.

Today, PRS for common diseases and lifestyle behavior can be simultaneously and inexpensively calculated for individuals at birth. The usefulness as well as potential harms of this knowledge needs to be carefully considered before this information is routinely used in health care, i.e., for screening subjects who are at an extreme risk of developing a disease. It is important to consider how absolute and relative risks are assessed and communicated at different stages of life. In addition, there is a need for studies investigating the potential genetic overlap between disease risk and health behavior in longitudinal settings as well as to increase knowledge regarding how potential interventions and treatments work depending on individuals' polygenic risk. Like any other genetic risk information, PRS values are not deterministic. This is demonstrated by studies of monozygotic twins, who have the same genomic sequence and hence identical polygenic scores for any trait or disease. Yet, even the identical twin pairs are more discordant than concordant for all common diseases, indicating the roles of the environment, lifestyle, and chance.

CONCLUSIONS

In conclusion, in this study, a higher PA PRS value was associated with a lower risk for several CMD and all-cause mortality. These findings highlight the fact that shared genetic factors may modify both health-related behavior, such as PA, and disease risk. The practical applications of polygenic risk information in disease screening as well as for guiding lifestyle and medical interventions remain to be investigated in further studies. An understanding of the individuals' genetic predisposition to diseases as well as lifestyle factors may help destigmatize individuals who cannot fulfill public lifestyle recommendations.

The authors thank all affiliated biobanks and the FinnGen participants and researchers for their valuable contribution to science and this work. Contributors of FinnGen are listed in Supplemental digital content 3 (see Document, Contributors of FinnGen, http://links.lww.com/MSS/C428).

The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie Inc., AstraZeneca UK Ltd., Biogen MA Inc., Celgene Corporation, Celgene International II Sàrl, Genentech Inc., Merck Sharp & Dohme Corp., Pfizer Inc., GlaxoSmithKline Intellectual Property Development Ltd., Sanofi US Services Inc., Maze Therapeutics Inc., Janssen Biotech Inc, and Novartis AG. The following biobanks are acknowledged for the project samples: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.ppshp.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki.fi/en), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaalle/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta), and Terveystalo Biobank (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/). All Finnish Biobanks are members of BBMRI.fi infrastructure (www.bbmri.fi) and FinBB (https://finbb.fi/).

E. S. was supported by the Juho Vainio Foundation and the Päivikki and Sakari Sohlberg Foundation. S. R. was supported by the Academy of the Finland Center of Excellence in Complex Disease Genetics (grant nos. 312062 and 336820), the Finnish Foundation for Cardiovascular

Research, the Sigrid Juselius Foundation, and the University of Helsinki HiLIFE Fellow and Grand Challenge grants.

The funders had no roles in the study design, collection, analysis, or interpretations of the data or in the publication process. The authors have no conflicts of interest to report. All authors declare that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

REFERENCES

1. Lopez AD, Murray CC. The global burden of disease, 1990–2020. *Nat Med*. 1998;4(11):1241–3.
2. Grotzinger AD, Rhemtulla M, de Vlaming R, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav*. 2019;3(5):513–25.
3. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518(7538):197–206.
4. Scott RA, Scott LJ, Magi R, et al. An expanded genome-wide association study of type 2 diabetes in Europeans. *Diabetes*. 2017;66(11): 2888–902.
5. Mars N, Koskela JT, Ripatti P, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med*. 2020;26(4):549–57.
6. Polderman TJ, Benyamin B, de Leeuw CA, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet*. 2015;47(7):702–9.
7. Kujala UM, Palviainen T, Pesonen P, et al. Polygenic risk scores and physical activity. *Med Sci Sports Exerc*. 2020;52(7):1518–24.
8. Kujala UM. Is physical activity a cause of longevity? It is not as straightforward as some would believe. A critical analysis. *Br J Sports Med*. 2018;52(14):914–8.
9. Kujala UM. Physical activity, genes, and lifetime predisposition to chronic disease. *Eur Rev Aging Phys Act*. 2011;8(1):31–6.
10. Kujala UM, Marti P, Kaprio J, Hernelahti M, Tikkanen H, Sarna S. Occurrence of chronic disease in former top-level athletes. Predominance of benefits, risks or selection effects? *Sports Med*. 2003;33(8):553–61.
11. Doherty A, Smith-Byrne K, Ferreira T, et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat Commun*. 2018;9(1):5257.
12. World Health Organization. *Global Action Plan on Physical Activity 2018–2030: More Active People for a Healthier World*. Geneva: World Health Organization; 2018.
13. Warburton DER, Bredin SSD. Health benefits of physical activity: a systematic review of current systematic reviews. *Curr Opin Cardiol*. 2017;32(5):541–56.
14. 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.
15. Ekelund U, Brown WJ, Steene-Johannessen J, et al. Do the associations of sedentary behaviour with cardiovascular disease mortality and cancer mortality differ by physical activity level? A systematic review and harmonised meta-analysis of data from 850 060 participants. *Br J Sports Med*. 2019;53(14):886–94.
16. Kujala U. Summary of the effects of exercise therapy in noncommunicable diseases: clinically relevant evidence from meta-analyses of randomized controlled trials. *medRxiv*. 2021.02.11.21251608, 2021; doi: https://doi.org/10.1101/2021.02.11.21251608.
17. Ahtiainen JP, Sallinen J, Hakkinen K, Sillanpaa E. Inter-individual variation in response to resistance training in cardiometabolic health indicators. *Scand J Med Sci Sports*. 2020;30(6):1040–53.
18. Karvinen S, Waller K, Silvennoinen M, et al. Physical activity in adulthood: genes and mortality. *Sci Rep*. 2015;5:18259.
19. Koch LG, Britton SL. Rat models of exercise for the study of complex disease. *Methods Mol Biol*. 2019;2018:309–17.
20. Bahls M, Leitzmann MF, Karch A, et al. Physical activity, sedentary behavior and risk of coronary artery disease, myocardial infarction and ischemic stroke: a two-sample Mendelian randomization study. *Clin Res Cardiol*. 2021;110(10):1564–73.
21. Pingault JB, O'Reilly PF, Schoeler T, Ploubidis GB, Rijsdijk F, Dudbridge F. Using genetic data to strengthen causal inference in observational research. *Nat Rev Genet*. 2018;19(9):566–80.
22. Zadro JR, Shirley D, Andrade TB, Scurrah KJ, Bauman A, Ferreira PH. The beneficial effects of physical activity: is it down to your genes? A systematic review and meta-analysis of twin and family studies. *Sports Med Open*. 2017;3(1):4,9.
23. Aaltonen S, Kujala UM, Kaprio J. Factors behind leisure-time physical activity behavior based on Finnish twin studies: the role of genetic and environmental influences and the role of motives. *Biomed Res Int*. 2014;2014:931820.
24. Sund R. Quality of the Finnish hospital discharge register: a systematic review. *Scand J Public Health*. 2012;40(6):505–15.
25. 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(2):e0169649.
26. Borodulin K, Tolonen H, Jousilahti P, et al. Cohort profile: The National FINRISK Study. *Int J Epidemiol*. 2018;47(3):696–696i.
27. Vilhjalmsson BJ, Yang J, Finucane HK, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet*. 2015;97(4):576–92.
28. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J Clin Epidemiol*. 2014;67(8):850–7.
29. 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.
30. Kujala UM, Kaprio J, Sarna S, Koskenvuo M. Future hospital care in a population-based series of twin pairs discordant for physical activity behavior. *Am J Public Health*. 1999;89(12):1869–72.
31. Kujala UM, Kaprio J, Koskenvuo M. Modifiable risk factors as predictors of all-cause mortality: the roles of genetics and childhood environment. *Am J Epidemiol*. 2002;156(11):985–93.
32. Klimentidis YC, Raichlen DA, Bea J, et al. Genome-wide association study of habitual physical activity in over 377,000 UK biobank participants identifies multiple variants including CADM2 and APOE. *Int J Obes (Lond)*. 2018;42(6):1161–76.
33. Kujala UM. Evidence on the effects of exercise therapy in the treatment of chronic disease. *Br J Sports Med*. 2009;43(8):550–5.
34. Stephan Y, Sutin AR, Luchetti M, Caille P, Terracciano A. An examination of potential mediators of the relationship between polygenic scores of BMI and waist circumference and phenotypic adiposity. *Psychol Health*. 2020;35(9):1151–61.
35. Lightfoot JT, De Geus EJC, Booth FW, et al. Biological/genetic regulation of physical activity level: consensus from GenBioPAC. *Med Sci Sports Exerc*. 2018;50(4):863–73.
36. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50(9):1219–24.
37. Qi Q, Li Y, Chomistek AK, et al. Television watching, leisure time physical activity, and the genetic predisposition in relation to body mass index in women and men. *Circulation*. 2012;126(15):1821–7.
38. Tyrrell J, Wood AR, Ames RM, et al. Gene–obesogenic environment interactions in the UK biobank study. *Int J Epidemiol*. 2017;46(2): 559–75.
39. Godino JG, van Sluijs EM, Marteau TM, Sutton S, Sharp SJ, Griffin SJ. Lifestyle advice combined with personalized estimates of genetic or phenotypic risk of type 2 diabetes, and objectively measured physical activity: a randomized controlled trial. *PLoS Med*. 2016;13(11):e1002185.

40. Widén E, Partanen J, Hotakainen K, et al. Communicating polygenic and non-genetic risk for atherosclerotic cardiovascular disease—an observational follow-up study. *medRxiv*. 2020.09.18.20197137, 2020; https://doi.org/10.1101/2020.09.18.20197137.
41. Karavirta L, Hakkinen K, Kauhanen A, et al. Individual responses to combined endurance and strength training in older adults. *Med Sci Sports Exerc*. 2011;43(3):484–90.
42. Brown BM, Peiffer JJ, Martins RN. Multiple effects of physical activity on molecular and cognitive signs of brain aging: can exercise slow neurodegeneration and delay Alzheimer's disease?*Mol Psychiatry*. 2013;18(8):864–74.
43. Lourida I, Hannon E, Littlejohns TJ, et al. Association of lifestyle and genetic risk with incidence of dementia. *JAMA*. 2019;322(5): 430–7.
44. Sabia S, Dugravot A, Dartigues JF, et al. Physical activity, cognitive decline, and risk of dementia: 28 year follow-up of Whitehall II cohort study. *BMJ*. 2017;357:j2709.
45. Roehr S, Pabst A, Luck T, Riedel-Heller SG. Is dementia incidence declining in high-income countries? A systematic review and meta-analysis. *Clin Epidemiol*. 2018;10:1233–47.
46. Vink JM, Hottenga JJ, de Geus EJ, et al. Polygenic risk scores for smoking: predictors for alcohol and cannabis use?*Addiction*. 2014; 109(7):1141–51.
47. Belsky DW, Moffitt TE, Baker TB, et al. Polygenic risk and the developmental progression to heavy, persistent smoking and nicotine dependence: evidence from a 4-decade longitudinal study. *JAMA Psychiat*. 2013;70(5):534–42.
48. Kiiskinen T, Mars NJ, Palviainen T, et al. Genomic prediction of alcohol-related morbidity and mortality. *Transl Psychiatry*. 2020;10(1):23.
49. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *Am J Epidemiol*. 2017;186(9): 1026–34.
50. Kujala UM, Hautasaari P, Vaha-Ypya H, et al. Chronic diseases and objectively monitored physical activity profile among aged individuals—a cross-sectional twin cohort study. *Ann Med*. 2019; 51(1):78–87.
51. Lim ET, Wurtz P, Havulinna AS, et al. Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet*. 2014;10(7):e1004494.